UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA

RODRIGO LOURO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE
TRANSCRITOS HUMANOS: novas famílias de
pequenas GTPases e novos Longos RNAs
Intrônicos não-codificantes**

São Paulo
Data do Depósito na SPG:
24/10/2006

RODRIGO LOURO

**IDENTIFICAÇÃO E CARACTERIZAÇÃO DE
TRANSCRITOS HUMANOS: novas famílias de
pequenas GTPases e novos Longos RNAs
Intrônicos não-codificantes**

*Tese apresentada ao Instituto de Química da
Universidade de São Paulo para obtenção do
Título de Doutor em Bioquímica*

Orientador: Prof. Dr. Sergio Verjovski-Almeida

São Paulo
2006

Ficha Catalográfica
Elaborada pela Divisão de Biblioteca e
Documentação do Conjunto das Químicas da USP.

L892i Louro, Rodrigo
Identificação e caracterização de transcritos humanos: novas famílias de pequenas GTPases e novos longos RNAs intrônicos não-codificantes / Rodrigo Louro. -- São Paulo, 2006. 202p.

Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica.
Orientador: Verjovski-Almeida, Sergio

1. Expressão gênica 2. Regulação gênica 3. Controle gênico
I. T. II. Verjovski-Almeida, Sergio, orientador
574.88 CDD

“Identificação e caracterização de transcritos humanos: novas famílias de pequenas GTPases e novos Longos RNAs intrônicos não-codificantes”

RODRIGO LOURO

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutorado em Ciências – Área: Bioquímica.

Aprovado(a) por:

Prof. Dr. SERGIO VERJOVSKI DE ALMEIDA
(Orientador e Presidente)

Profa. Dra. SUELY LOPES GOMES
IQ - USP

Profa. Dra. VILMA REGINA MARTINS
Ludwig

Profa. Dra. ENILZA MARIA ESPREAFICO
FM - RP - USP

Prof. Dr. MARCELO RIBEIRO DA SILVA BRIONES
UNIFESP

SÃO PAULO
27 DE NOVEMBRO 2006

*Aos meus queridos pais, Percival e Izabel, que com amor extremo e incondicional
proporcionam aos seus filhos TUDO para transformar sonhos em realidade.*

Devo tudo que sou a vocês.

AGRADECIMENTOS

Ao Prof. Dr. Sergio Verjovski-Almeida, e ao Prof. Dr. Eduardo Reis, pelo apoio, incentivo e orientação constantes desde o início de minhas atividades científicas. Ambos são os responsáveis pela minha formação como pesquisador.

À Profa. Dra. Adele Caterino-de-Araújo, minha madrinha, por servir como modelo de dedicação e sucesso nessa difícil, porém apaixonante, profissão.

Ao Prof. Dr. Bayardo Baptista, que abriu meus olhos para a beleza dos fenômenos bioquímicos, e que tem me inspirado a não deixar a importante função de ensinar em segundo plano.

À Profa. Dra. Aline Maria da Silva pelas importantes colaborações nesse trabalho, pela utilização dos equipamentos de seu laboratório, e pela fundamental orientação para formatação dessa tese. Agradeço também à Profa. Dra. Mari Cleide Sogayar, pelas críticas ao trabalho de análise de expressão gênica após estímulo por andrógeno.

Aos parceiros próximos de projeto (Helder Nakaya, Paulo Paiva Amaral e Fernanda Festa) com quem compartilhei diversos momentos de entusiasmo e de decepção proporcionados pela ciência nesses últimos anos.

Aos grandes amigos gaúchos, Tarik e Ângela, conquistados durante a pós-graduação, pelos novos conhecimentos (científicos e culturais) adquiridos, e por serem ótimos companheiros de trabalho, de cafezinhos, de cervejinhas e de poker.

À Ana Paula Lopes Vidal que tem facilitado (e muito) minha relação com a FAPESP e com as empresas que fabricam e vendem reagentes para pesquisa. Sem ela a rotina de laboratório seria muito mais estressante.

Aos alunos de iniciação científica que contribuíram para a obtenção dos resultados apresentados aqui, Rafael, Yuri, Junio, Vinícius e Bárbara Camargo.

Aos demais colegas de laboratório (Glauber, Thiago, Renato, Jefferson, Camila, Giulliana, Kátia, Felipe, Lauren, Rogério, Ana C. Tahira, Ana C. Ayupe, Ana Cláudia, Denise, Adriana, Julio e Ricardo) e do departamento de bioquímica (em especial, Vanessa Falcão, Camila Carrião, e Osmar), que tive a sorte de conhecer e de dividir boas experiências tanto no trabalho quanto na vida “real”.

Ao Dr. Luiz Otavio Penalva, professor assistente da UTHSCSA, e seu grupo de pesquisa (Deepa, Patrícia Sanchez e Suzanne Burns) pela oportunidade de aprender novas técnicas e conceitos, de superar antigos temores, e pela agradável recepção no “inóspito” estado do Texas.

À Carina Victoria e ao Carlos Alexandre, cuja amizade que se iniciou durante a “inocente” época de graduação, passou pela conturbada fase da pós-graduação e deve prosseguir pela vida toda. Obrigado pelo apoio nos momentos difíceis, pela manutenção das tradições, e pelas constantes discussões científicas livres de embasamento experimental.

À minha querida irmã Carolina Louro, ao meu grande irmão Alexandre Miranda, aos demais amigos não-biólogos mais próximos (Paulão, Paulo “Feijão”, entre outros), aos meus demais familiares, em especial ao meu padrinho Carlos, às minhas tias Marisa e Marilene e à minha falecida avó Amélia que admiram meu trabalho mesmo achando que fazer pesquisa é a maior moleza.

À FAPESP pelo apoio financeiro ao meu projeto (02/02950-1) e aos projetos temáticos relacionados (99/07390-0 e 02/13283-6).

Aos médicos clínicos das instituições (Hospital do Câncer/RJ, Hospital Sírio-Libanês/SP e Hospital das Clínicas da USP) que as cederam as preciosas amostras de tecido que viabilizaram parte dos estudos apresentados nessa Tese

E finalmente, ao Morihei Ueshiba (*O-Sensei*), criador e fundador do Aikido. Praticando essa fantástica arte marcial consegui atingir um novo equilíbrio em minha vida, que evitou a concretização literal e integral da epígrafe contida na página seguinte.

"I have put my heart and soul into this work...and lost my mind in the process."

Vincent Van Gogh

RESUMO

Louro, R. **Identificação e caracterização de transcritos humanos: novas famílias gênicas de pequenas GTPases e novos Longos RNAs Intrônicos não-codificantes.** 2006. 202 f. Tese (Doutorado) - Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

Terminado o sequenciamento do genoma humano, as atenções se voltaram para a determinação do conjunto completo de transcritos humanos. Diversos trabalhos sugerem que enquanto apenas uma pequena fração de mRNAs codificantes para proteína não é conhecida, existe um grande número de RNAs não-codificantes (ncRNAs) ainda não caracterizados. Nesse contexto, o presente trabalho visou explorar as informações de expressão gênica contidas em ESTs para identificar e caracterizar novos transcritos humanos. A busca genômica por membros de famílias gênicas relacionadas com câncer levou a identificação de novas pequenas GTPases, destacando uma subfamília que deve apresentar função supressora tumoral em próstata. Uma classe de ncRNAs longos, sem *splicing*, expressos antisenso a partir de regiões intrônicas foi descrita utilizando plataformas de *microarrays*, construídas pelo grupo, enriquecidas com seqüências sem anotação. O perfil de expressão de 23 ncRNAs intrônicos estava significativamente correlacionado com o grau de diferenciação de tumores de próstata (*Gleason Score*), e pode ser utilizado como candidato a marcador molecular de prognóstico. Um total de 39 ncRNAs intrônicos responderam à estimulação por andrógeno, apontando para um mecanismo regulatório da expressão intrônica por sinais fisiológicos hormonais. A biogênese da expressão intrônica parece ser complexa, pois uma fração não é transcrita pela RNA Polimerase II. A transcrição intrônica estava correlacionada com uso de exons em células tratadas com andrógeno.

Assinaturas de expressão intrônica conservadas em tecidos humanos e de camundongos, e interações de transcritos intrônicos com proteínas regulatórias foram observadas. Este trabalho contribui com novas e originais evidências que dão apoio ao papel postulado para esses ncRNAs no controle fino do programa transcricional humano.

Palavras-chave: Análise de expressão gênica, Pequenas GTPases monoméricas, Câncer de próstata, RNAs não-codificantes, Introns, Andrógeno, Controle da expressão gênica.

ABSTRACT

Louro, R. **Identification and characterization of human transcripts: novel small GTPase gene families and novel Long Intronic non-coding RNAs.** 2006. 202 p. Thesis (Doctoral) - Biochemistry Graduate Program. Instituto de Química. Universidade de São Paulo.

With the completion of the human genome sequence, attention has shifted towards determining the complete set of human transcripts. Multiple lines of evidence suggest that while only a small fraction of protein-coding mRNAs remains to be described, there is a huge amount of uncharacterized non-coding RNAs (ncRNAs). In this context, the present work sought to explore the gene expression information provided by ESTs to identify and characterize new human transcripts. A genomic-wide search for cancer related gene family members identified novel small GTPase genes, and highlighted an uncharacterized subfamily that may have a tumor suppressor role in prostate cancer. A class of long unspliced ncRNAs, expressed antisense from introns of protein-coding genes was described using custom-designed microarray platforms enriched with unannotated sequences. The expression profile of 23 intronic ncRNAs was significantly correlated to the degree of prostate tumor differentiation (Gleason Score), and could be used as a candidate prognostic molecular marker. A total of 39 intronic ncRNAs were responsive to androgen stimulation, pointing to a mechanism of intronic expression regulation by physiological hormone signals. Intronic ncRNA biogenesis seems to be complex, since a fraction of them is not transcribed by RNA Polymerase II. Intronic transcription was correlated to exon usage in androgen treated cells. Tissue expression signatures of intronic transcription were conserved in human and mouse, and intronic transcripts were found to interact with regulatory

proteins. This work provides new and original contributions that support the postulated role of ncRNAs in the fine tuning of the human transcriptional program.

Keywords: Gene expression analysis, Small monomeric GTPases, Prostate cancer, Non-coding RNAs, Introns, Androgen, Gene expression control.

LISTA DE ABREVIATURAS E SIGLAS

ARE	Elemento de Resposta a Andr�geno
ATCC	<i>American Type Culture Collection</i>
BLAST	<i>Basic Local Alignment Search Tool</i>
bp	Pares de bases
CAGE	Coopera��o para An�lise da Express�o G�nica
CBC	<i>Cap Binding Complex</i>
CGAP	<i>Cancer Genome Anatomy Project</i>
ChIP	Imunoprecipita��o da cromatina
C_T	<i>Cycle Threshold</i>
ESE	<i>Exonic Splicing Enhancer</i>
EST	<i>Expressed Sequence Tag</i>
FANTOM	<i>Functional Annotation of the Mouse Genome</i>
GEO	<i>Gene Expression Omnibus</i>
GO	<i>Gene Ontology</i>
GS	<i>Gleason Score</i>
GTF	<i>General Transcription Factor</i>
HCGP	<i>Human Cancer Genome Project</i>
IgG	Imunoglobulina G
IPTG	<i>isopropyl-beta-D-thiogalactopyranoside</i>
ISE	<i>Intronic Splicing Enhancer</i>
ISS	<i>Intronic Splicing Silencer</i>
kb	Quilobase
LCR	<i>Locus Control Region</i>
MPSS	<i>Massively Parallel Signature Sequencing</i>
mRNA	RNA mensageiro
NCBI	<i>National Center for Biotechnology Information</i>
ncRNA	RNA n�o-codificante para prote�nas
nt	Nucle�tidos
ORF	<i>Open reading frame</i>
OSRT	<i>Orientation Specific Reverse Transcription</i>
PCR	Rea��o em Cadeia da Polimerase
PIC	<i>Preiniciation Complex</i>
PMSF	Fluoreto de fenilmetilsulfonil
qPCR	PCR quantitativo

RACE	<i>Rapid Amplification of cDNA Ends</i>
RBP	<i>RNA Binding Protein</i>
RefSeq	<i>Reference Sequence</i>
RNAP	<i>RNA Polimerase</i>
RNP	<i>Ribonucleoprotein</i>
SAGE	<i>Serial Analysis of Gene Expression</i>
SAM	<i>Significance Analysis of Microarrays</i>
SDS	<i>Dodecilsulfato de sódio</i>
TF	<i>Transcription Factor</i>
TFBS	<i>Transcriptional Factor Binding Site</i>

SUMÁRIO

1 INTRODUÇÃO	18
1.1 Conteúdo transcricional humano	18
1.2 Os RNAs não-codificantes	24
1.3 Transcrição gênica e seus mecanismos de regulação	29
1.3.1 Fatores de transcrição	32
1.3.2 Modificações pós-transcricionais.....	35
1.3.4 Proteínas ligadoras de RNA.....	38
1.3.5 Envolvimento de RNAs não-codificantes	39
1.4 Genômica comparativa e seqüências não-codificantes	47
1.5 Expressão gênica e câncer.....	49
1.5.1 Parálogos e famílias gênicas	49
1.5.2 Marcadores moleculares de diagnóstico e prognóstico em câncer de próstata	54
2 OBJETIVOS E RELEVÂNCIA	58
3 MATERIAIS E MÉTODOS	61
3.1 Análises Informáticas.....	61
3.1.1 Identificação <i>in silico</i> de parálogos de genes relacionados com câncer ...	61
3.1.2 Análises dos produtos protéicos hipotéticos de <i>RASL11A</i> e <i>ARL-9</i>	62
3.1.3 Análises Estatísticas para experimentos de expressão gênica com <i>microarrays</i>	63
3.1.5 Análises de enriquecimento de categorias de ontologia gênica (GO)	67
3.2 Amostras de tecido e linhagens celulares.....	68
3.2.1 Amostras de tecidos humanos	68
3.2.2 Amostras de tecidos de camundongos	69
3.3 Extração de RNA	71
3.3.1 Extração de RNA poli-(A)+	71
3.3.2 Extração de RNA total.....	72
3.3.3 Extração de RNA co-imunoprecipitado com proteínas ligadoras de RNA..	72
3.3.4 Tratamento com DNase livre de RNase.....	73
3.3.5 Quantificação e verificação da qualidade dos RNAs extraídos	74

3.4 Abordagens de PCR para amplificação de transcritos completos	75
3.4.1 RACE-PCR	75
3.4.2 PCR com oligonucleotídeos iniciadores específicos	76
3.5 Estudo de expressão gênica por <i>microarrays</i> de cDNA.....	77
3.5.1 Plataforma de <i>microarrays</i> de cDNA do IQUSP	77
3.5.2 Obtenção de alvos fluorescentes	79
3.5.3 Hibridizações em lâminas de <i>microarrays</i> de cDNA.....	82
3.5.4 Desenho dos experimentos com <i>microarrays</i> de cDNA	83
3.5.5 Obtenção de imagens e valores de intensidade comparáveis	85
3.6 Hibridizações em <i>Northern Blots</i>	88
3.6.1 Detecção de mensagens utilizando sondas de fita dupla.....	89
3.6.2 Detecção de mensagens utilizando sondas de fita simples	90
3.7 Análise de expressão gênica por PCR	92
3.7.1 Determinação da orientação de mensagens utilizando cDNA transcrito reversamente com oligonucleotídeos iniciadores específicos.....	92
3.7.2 Avaliação da abundância relativa de mensagens utilizando PCR quantitativo.....	94
3.8 Experimentos de Imunoprecipitação	95
3.8.1 Ensaio de imunoprecipitação da cromatina.....	95
3.8.2 Imunoprecipitação de proteínas ligadoras de RNA (RBPs).....	96
3.9 Sistemas de expressão episomal	98
3.9.1 Construção do sistema de expressão de Rasl11a/b em procarioto	98
3.9.2 Indução da expressão de Rasl11b em bactéria	100
3.9.3 Detecção de Rasl11b expresso em bactéria com anticorpo anti-histidinas.	101
3.9.4 Sistema de expressão de proteínas ligadoras de RNA em células mamífero	102
3.9.5 Detecção das RBPs expressas em 293T com anticorpo contra proteína A.	103
3.9.6 Sistema de expressão de transcritos intrônicos em células de mamífero	104
3.10 Experimentos para obtenção de anticorpos.....	106
3.10.1 Preparo de antígenos de Rasl11b para imunização de coelhos.	106
3.10.2 Imunização de coelhos com corpos de inclusão de Rasl11b.	107

3.11 Análise de expressão gênica por <i>microarrays</i> de oligonucleotídeos	108
3.11.1 Plataforma de <i>microarrays</i> de oligonucleotídeos.....	108
3.11.2 Marcação de alvos fluorescentes por amplificação de RNA	110
3.11.3 Hibridização dos experimentos com lâminas de <i>microarrays</i> de oligonucleotídeos	110
3.11.4 Obtenção de imagens e valores de intensidade comparáveis	111
4 RESULTADOS.....	112
4.1 Identificação e caracterização de membros representativos de novas famílias de pequenas GTPases monoméricas.....	112
4.1.1 <i>RASL11A</i> e <i>ARL-9</i> , dois novos genes codificantes para pequenas GTPases monoméricas.....	112
4.1.2 <i>RASL11A</i> e <i>ARL-9</i> constituem novas subfamílias de pequenas GTPases.	114
4.1.3 Produtos protéicos codificados por <i>RASL11A</i> e <i>ARL9</i> são muito conservados entre os eucariotos.	120
4.1.4 Níveis do transcrito <i>RASL11A</i> estão diminuídos em tumores de próstata.	122
4.1.5 Sistema de expressão de proteínas Rasl11 em procariotos.	124
4.1.6 Expressão de Rasl11b recombinante em corpos de inclusão para imunização de coelhos.....	126
4.2 Transcritos intrônicos marcadores de prognóstico em câncer de próstata....	128
4.2.1 Transcritos intrônicos são expressos em próstata	129
4.2.2 Níveis de transcritos intrônicos não codificantes são correlacionados ao grau de diferenciação tumoral em próstata	130
4.2.3 Caracterização de transcritos intrônicos relacionados com câncer.....	133
4.2.4 Validação por PCR quantitativo em tempo real das mensagens intrônicas de <i>RASSF1</i>	137
4.3 Transcritos intrônicos regulados pelo estímulo de andrógeno.....	139
4.3.1 Identificação de transcritos intrônicos regulados por andrógeno	139
4.3.2 RNAs intrônicos regulados por andrógeno mapeiam em genes envolvidos em transdução de sinais	143
4.3.3 Caracterização dos RNAs intrônicos regulados por andrógeno.....	143
4.3.4 Efeitos do andrógeno nos níveis dos RNAs intrônicos e dos transcritos codificantes de proteína correspondentes.....	147

4.3.5 Aumento da ligação de AR a um motivo ARE a 5' do <i>asMYO5A</i>	151
4.4 Transcritos intrônicos sintênicos com baixa conservação expressos diferencialmente em tecidos humanos e de camundongos.	153
4.4.1 Detecção de transcritos intrônicos em tecidos humanos e de camundongos.....	154
4.4.2 Expressão diferencial de transcritos intrônicos em tecidos humanos e de camundongos.....	155
4.4.3 Validação da existência e orientação transcricional de mensagens intrônicas sintênicas diferencialmente expressas em tecidos humanos e de camundongos.....	156
4.5 Interação de transcritos intrônicos com proteínas ligadoras de RNA.	159
4.5.1 Expressão das proteínas ligadoras de RNA em células 293T	159
4.5.2 Identificação, utilizando <i>microarrays</i> de cDNA, de RNAs longos intrônicos não-codificantes associados a proteínas ligadoras de RNA	160
4.6 Análise de expressão de RNAs intrônicos em células LNCaP tratadas com α - amanitina	161
4.6.1 Uma grande fração de RNAs intrônicos aparentemente não é transcrita pela RNAP II	162
4.6 Sistema para modulação da expressão de <i>asRASSF1</i> e <i>asPPHLN1</i> em células humanas	164
5 DISCUSSÃO	165
5.1 Identificação e caracterização parcial de <i>RASL11A</i> e <i>ARL-9</i> , membros representativos de novas famílias de pequenas GTPases monoméricas.	165
5.2 Transcritos intrônicos marcadores de prognóstico em câncer de próstata....	168
5.3 Transcritos intrônicos regulados pelo estímulo de andrógeno.....	171
5.4 Transcritos intrônicos diferencialmente expressos em tecidos humanos e de camundongos.	175
5.5 Interação de transcritos intrônicos com proteínas ligadoras de RNA.	178
5.6 Análise de expressão de RNAs intrônicos em células LNCaP tratadas com α - amanitina	179
6 CONCLUSÕES E CONSIDERAÇÕES FINAIS	181
7 REFERÊNCIAS BIBLIOGRÁFICAS.....	184
8 LISTA DE ANEXOS.....	201

1 INTRODUÇÃO

1.1 Conteúdo transcricional humano

Com a conclusão do projeto Genoma Humano (*Human Genome Project*) no começo de 2001 e a publicação simultânea dos rascunhos da sequência de bases completa que compõe nosso código genético, produzidos pelas iniciativas pública (Lander *et al.*, 2001) e privada (Venter *et al.*, 2001), tornou-se tanto factível quanto desejável a determinação do conjunto completo de transcritos humanos. Contudo, rapidamente ficou evidente que a simples utilização de algoritmos *in silico* para buscas de regiões exônicas na sequência completa de DNA genômico era uma abordagem difícil e bastante limitada, além de ser totalmente enviesada, pois definia apenas as regiões codificantes para proteína.

Assim, diversas estratégias complementares têm sido utilizadas por pesquisadores no mundo todo para atingir os objetivos tanto de identificação do repertório transcricional humano total, quanto de caracterização individual das mensagens que o compõe. Nesse contexto foi fundamental o aperfeiçoamento de técnicas de análise de expressão gênica em larga escala. Abordagens de estudo por *microarrays* de DNA (Shoemaker *et al.*, 2001; Kapranov *et al.*, 2002; Rinn *et al.*, 2003; Bertone *et al.*, 2004; Kampa *et al.*, 2004), assim como através de sequenciamento de fragmentos expressos (Williamson, 1999; Dias-Neto *et al.*, 2000; Strausberg *et al.*, 2000; Camargo *et al.*, 2001; Saha *et al.*, 2002; Ota *et al.*, 2004; Jongeneel *et al.*, 2005) utilizando amostras de diversos tecidos, têm sido largamente empregadas.

Análises de estruturas gênicas realizadas a partir da sequência completa do genoma humano revelaram que os introns compreendem em média 95% dos genes codificantes de proteína, os exons internos são relativamente pequenos, e os genes

aparentemente possuem muitas formas alternativas, criando dificuldades para os programas de predição baseados na busca de fases abertas de leitura (*Open Reading Frames* – ORFs) devido a pequena razão sinal-ruído (Lander *et al.*, 2001; Venter *et al.*, 2001). Dessa maneira utilizando abordagens complexas, compostas de análises *in silico* e evidências experimentais, foram estimadas cerca de 32.000 (Lander *et al.*, 2001) e 27.000 (Venter *et al.*, 2001) mensagens codificantes para proteína no genoma humano. Aproximadamente 700 RNAs não-codificantes (ncRNAs) estruturais clássicos, incluindo tRNAs, rRNAs, snRNAs e snoRNAs, e o *Xist* foram preditos no genoma através de buscas por similaridade (Lander *et al.*, 2001). Porém, estimativas anteriores baseadas no mapeamento de seqüências expressas (*Expressed Sequence Tags* – ESTs) apontavam para um número total de mensagens substancialmente maior, que variava desde 35.000 (Ewing & Green, 2000) até 120.000 (Liang *et al.*, 2000) transcritos.

De fato, nos esforços pioneiros onde foram utilizadas plataformas de *microarrays* contendo sondas para toda a seqüência genômica dos cromossomos 21 e 22, foi observado um número cerca de 10 vezes maior de regiões transcricionalmente ativas do que o predito anteriormente pelo mapeamento de regiões exônicas de genes humanos conhecidos (Kapranov *et al.*, 2002; Rinn *et al.*, 2003; Kampa *et al.*, 2004). Frações comparáveis de atividade transcricional foram detectadas nos cromossomos 21 e 22 a partir de regiões exônicas e de regiões intrônicas de genes funcionalmente anotados (Rinn *et al.*, 2003; Kampa *et al.*, 2004).

Ainda foi constatado que cerca da metade dos transcritos provenientes das regiões intrônicas são expressos antisenso em relação ao bem caracterizado intron do gene codificante para proteína de mesmo *locus*, tanto nas análises focadas nos cromossomos 21 e 22 (Rinn *et al.*, 2003; Kampa *et al.*, 2004) quanto em um estudo

com todo o genoma humano (Bertone *et al.*, 2004). Nesse último estudo apenas 13.889 regiões humanas transcricionalmente ativas foram detectadas devido principalmente à utilização de uma abordagem de análise que privilegiou a especificidade em sacrifício da sensibilidade (Bertone *et al.*, 2004). Entretanto, Bertone e colaboradores já puderam gerar um rascunho do mapeamento da expressão proveniente do genoma humano completo, através da observação de muitos transcritos conhecidos ou preditos, e de 10.595 novas seqüências, das quais 3.095 (22% do total) eram intrônicas, sendo cerca de metade dessas (1.529), provenientes de transcrição antisenso em relação ao gene codificante para proteína de mesmo *locus* (Bertone *et al.*, 2004).

Evidências similares de atividade transcricional antisenso ubíqua a partir de regiões intrônicas no genoma humano, assim como no genoma de camundongo, foram também obtidas por metodologias baseadas em sequenciamento de fragmentos expressos. Essas abordagens são ainda mais abrangentes, pois possuem como vantagens a possibilidade de detecção de transcritos provenientes de regiões repetitivas, e de avaliação da orientação das mensagens com menor probabilidade de artefatos. Análises de mapeamento de porções 3' de transcritos utilizando LongSAGE, uma modificação de SAGE (*Serial Analysis of Gene Expression*) (Saha *et al.*, 2002), de porções 5' (Ota *et al.*, 2004), e de MPSS (*Massively Parallel Signature Sequencing*) (Jongeneel *et al.*, 2005), a partir de amostras humanas; assim como análises semelhantes realizadas pelo projeto de anotação funcional do genoma de camundongo (*Functional Annotation of the Mouse Genome* - FANTOM) (Okazaki *et al.*, 2002; Carninci *et al.*, 2005), confirmaram as observações iniciais.

No projeto FANTOM, através de sequenciamento em larga escala e anotação manual de cDNAs completos foram caracterizadas 102.281 mensagens, incluindo 32.129 codificantes para proteína (2.222 para proteínas novas) e 34.030 ncRNAs (Carninci *et al.*, 2005). Ainda nesse projeto os pesquisadores também concluíram que o número total de transcritos seria pelo menos 10 vezes maior que o previsto pelas estimativas de predição informática de genes no genoma de camundongos (Carninci *et al.*, 2005), e notaram que apenas 40% dos ncRNAs já haviam sido identificados em coleções anteriores de cDNA FANTOM (Okazaki *et al.*, 2002), sugerindo que essa quantidade ainda poderia ser maior.

As análises de seqüências expressas humanas têm confirmado as observações que apontam para um elevado nível de diversidade transcricional em células de mamíferos. Os estudos de expressão por LongSAGE indicam cerca de 15.000 regiões 3' humanas ainda sem caracterização, sugerindo também que o número de novas isoformas ou mesmo de novos transcritos deva ser na ordem de milhares (Saha *et al.*, 2002). Depois do sequenciamento da região 5' de cerca de 1.000.000 de clones de cDNA, e da remoção dos referentes a genes conhecidos, Ota e colaboradores encontraram 7.000 novos transcritos humanos completos a partir de diversos tecidos e linhagens celulares humanos (Ota *et al.*, 2004). Apenas metade dessas mensagens parecem codificar para proteínas, sugerindo que tanto mRNAs quanto ncRNAs ainda permanecem a ser caracterizados (Ota *et al.*, 2004). Mais importante, a análise por MPSS realizada em uma coleção de 32 tecidos humanos, apontou para cerca de 65% de assinaturas transcripcionais (total de 142.872) que não possuem sobreposição com mensagens anotadas (Jongeneel *et al.*, 2005). Dessas, 53.402 (37%) mapeiam em regiões intergênicas ou intrônicas de *loci* gênicos anotados, não descritas anteriormente como sendo expressas

individualmente (Jongeneel *et al.*, 2005). Aproximadamente metade (29.661 = 55,5%) dessas assinaturas de MPSS possuem orientação antisense, mapeando na fita oposta de transcritos já conhecidos (Jongeneel *et al.*, 2005).

Além da expressão de mensagens a partir de porções intrônicas e intergênicas, a observação de ocorrência natural e simultânea de transcrição senso e antisense também tem sido observada de forma muito mais comum do que antecipado, tanto por estratégias de hibridização em *microarrays* quanto por análises computacionais de seqüências expressas, no genoma humano e de camundongo (Shendure & Church, 2002; Kiyosawa *et al.*, 2003; Yelin *et al.*, 2003; Chen *et al.*, 2004; Kampa *et al.*, 2004). Utilizando bancos de EST e de seqüências completas foram identificados 5.880 *clusters* de transcritos humanos que contêm mensagens antisense sobrepostas, sendo algumas dessas validadas por *Northern Blot* (Yelin *et al.*, 2003; Chen *et al.*, 2004). De 60.000 clones de cDNA de camundongo, ~2.400 pares senso–antisense foram detectados (Kiyosawa *et al.*, 2003). E em experimentos utilizando *tiling arrays* foram detectados 11% de sondas para exons conhecidos com transcrição na fita oposta (Kampa *et al.*, 2004). Transcrição, tanto de ncRNAs quanto de mRNAs, a partir da fita oposta de *locus* codificantes para proteína já anotados, produzem RNAs que podem hibridizar com o DNA ou RNA e interferir com a transcrição, tradução ou estabilidade do produto senso (Reis *et al.*, 2005¹; Mattick & Makunin, 2006), como será discutido mais adiante, no tópico apropriado.

Uma análise direcionada da extensão da transcrição intrônica em humanos foi realizada apenas recentemente, através de uma abordagem de “clusterização” e mapeamento de mais de 5 milhões ESTs depositadas em bancos de dados públicos

¹ Reis *et al.* 2005 (Artigo publicado). ANEXO 1 - incluso no CD-ROM suplementar a esta Tese.

(Nakaya *et al.*, 2006)². Aproximadamente 55 mil unidades transcricionais (representadas por *clusters* de ESTs), sem evidência de *splicing*, totalmente intrônicas e 12.600 parcialmente intrônicas (com sobreposição com um éxon) e com tamanhos médios 4 vezes maiores que os tamanhos médios de exons foram identificadas (Nakaya *et al.*, 2006). Uma análise de potencial codificante indicou que cerca de 90% desses transcritos intrônicos provavelmente correspondem a ncRNAs (Nakaya *et al.*, 2006). Ainda, regiões intrônicas transcricionalmente ativas foram observadas em 74 % dos genes humanos RefSeq conhecidos, ressaltando a abundância dos ncRNAs intrônicos e apontando para alguma importância funcional (Nakaya *et al.*, 2006).

Estimativas baseadas nos trabalhos discutidos até aqui indicam que até 60% do genoma humano é transcrito, porém essa porcentagem ainda pode estar sendo subestimada devido ao número limitado de células e estados de diferenciação analisados (Willingham & Gingeras, 2006). Esse enorme número de transcritos aponta para uma arquitetura transcricional em humanos extremamente complexa, que inclui formas alternativas de *splicing* para a imensa maioria de mensagens codificantes de proteína, uma imensa quantidade de ncRNAs, e transcrição antisenso disseminada (Figura 1) que provavelmente possui importantes funções biológicas mas ainda precisam ser determinadas (Willingham & Gingeras, 2006).

Finalmente, uma diversidade transcricional como essa pode responder como quantidades relativamente similares de genes codificantes para proteína estimadas para moscas (13.985; *BDGP release 4*), para vermes (21.009; *Wormbase release 150*), e para humanos (23.341; *NCBI release 36*) resultam nos fenótipos extremamente diferentes observados nessas espécies (Willingham & Gingeras,

² Nakaya *et al.* 2006 (Artigo submetido). ANEXO 2 - incluso no CD-ROM suplementar a esta Tese.

2006). Assim, a identificação sistemática e a caracterização das diversas classes de mensagens humanas expressas continuam sendo importantes desafios, que quando superados resultarão na ampliação da compreensão da biologia de nossa espécie.

Figura 1 – **Representação gráfica da complexidade transcricional humana.** As caixas brancas representam seqüências não-codificantes e as caixas azuis as regiões codificantes. A utilização de exons no transcrito maduro é representada pelas linhas diagonais. Losangos verdes representam snoRNAs e os triângulos laranja os microRNAs. Em (A) estão representados transcritos codificantes e não-codificantes para proteína e as regiões de transcrição de snoRNAs e microRNAs. Em (B) transcritos antisense com exons entrelaçados. Em (C) transcritos totalmente intrônicos senso e antisense. Em (D) transcritos parcialmente intrônicos. Figura adaptada de Mattick, J. S. & Makunin, I. V. (2006). Non-coding RNA. *Hum Mol Genet*, **15** suppl 1, R17-29.

1.2 Os RNAs não-codificantes

Potencialmente existem muitas explicações para a atividade transcricional fora de regiões exônicas conhecidas: seriam novas mensagens, ou novos exons alternativos codificantes para proteína, RNAs não-codificantes (ncRNAs), ou apenas simples ruído biológico. Apesar do conjunto central de genes curados e bem caracterizados (*RefSeq*) ser concordante com a maioria das predições para o número de genes humanos, a comparação das estimativas iniciais realizadas pelas diferentes montagens do genoma humano (Lander *et al.*, 2001; Venter *et al.*, 2001) indicava pouca sobreposição de genes não-*RefSeq* (Hogenesch *et al.*, 2001), sugerindo um maior número de mensagens codificantes para proteína. De fato alguns trabalhos, em especial baseados em sequenciamento de cDNA, indicam que ainda existem genes humanos codificantes para proteína a serem descobertos, na maioria mensagens com baixos níveis de expressão, específicas de tecidos ou

estágios de desenvolvimento, ou ainda com propriedades de seqüência atípicas que dificultaram sua detecção (Ota *et al.*, 2004). Porém a observação que de um conjunto de 21.243 clusters não-redundantes de transcritos, menos de 9% eram novos e apresentavam um potencial codificante significativo (Ota *et al.*, 2004), indica uma tendência clara de queda na taxa de descobrimento desse tipo de mensagem, sugerindo que dificilmente a grande maioria dos transcritos humanos sejam codificantes para proteína. Na realidade, atualmente já existem fortes indícios que a expressão de ncRNAs, provenientes principalmente de regiões humanas intrônicas com orientação transcricional senso ou antisense em relação ao mRNA do mesmo *locus* gênico e de regiões intergênicas ativas, é um fenômeno bastante geral (Mattick & Makunin, 2006).

Um importante argumento a favor da expressão funcional de ncRNAs foi provido por estudos que combinaram ensaios de imunoprecipitação da cromatina (*Chromatin Immunoprecipitation Assays* - ChIP) e *tiling arrays*. Essa estratégia permitiu um mapeamento aleatório, e de alta resolução ao longo do genoma humano, de sítios de ligação para fatores de transcrição (*Transcription Factor Binding Sites* – TFBS) e para outros complexos proteicos envolvidos na transcrição. Foi encontrada ligação destes fatores não só na região promotora a montante de genes codificantes para proteína, mas também em regiões intrônicas e intergênicas do genoma, o que deu apoio a existência disseminada de transcrição (Martone *et al.*, 2003; Cawley *et al.*, 2004; Euskirchen *et al.*, 2004; Kim *et al.*, 2005). Os TFBS para RelA/p65, membro da família NF- κ B, foram mapeados no cromossomo 22, sendo que 22% dos sítios encontravam-se a mais de 50kb distantes de genes codificantes para proteína, e 40% em regiões intrônicas de genes anotados (Martone *et al.*, 2003). O mapeamento dos TFBS para Sp1, cMyc, e p53 nos cromossomos 21 e 22

revelou que cerca da metade desses está correlacionada com a presença de ncRNAs, sendo que uma fração considerável desses transcritos respondeu a um tratamento com ácido retinóico (Cawley *et al.*, 2004). Um número significativo de sítios de ligação de CREB no cromossomo 22 foi também encontrado em regiões sem anotação, mas transcricionalmente ativas, como introns e seqüências intergênicas (Euskirchen *et al.*, 2004). Ainda, o mapeamento no genoma completo do sítios de ligação para o complexo de pré-iniciação da RNA polimerase II em fibroblastos humanos identificou cerca de 10.000 promotores ativos, sendo aproximadamente 13% correspondentes a regiões intrônicas ou intergênicas sem anotação (Kim *et al.*, 2005).

Outros fortes argumentos são obtidos a partir das evidências de relação funcional de ncRNAs com importantes processos celulares. Até recentemente, em uma época anterior a essa revolução genômica que foi parcialmente ilustrada no tópico anterior (item 1.1), um número restrito de ncRNAs foi descrito como participando de funções relativamente genéricas nas células. Os RNAs ribossomais (rRNAs) e RNAs transportadores (tRNAs) atuando na tradução de RNAs mensageiros (mRNAs), os pequenos RNAs nucleares (snRNAs) envolvidos em *splicing*, e os pequenos RNAs nucleolares (snoRNAs) participando principalmente na modificação dos rRNAs (Mattick & Makunin, 2006). Aparentemente a grande maioria desses outros ncRNAs desempenham papéis definidos no programa transcricional humano, assim como apresentado nos próximos parágrafos. Contudo, a expressão desse tipo de transcrito ocorre em níveis mais baixos (2 a 3 vezes menores) quando comparados com os níveis dos transcritos codificantes de proteína conhecidos (Kampa *et al.*, 2004), sendo essa a razão mais provável para sua ampla existência ter sido reconhecida apenas a partir do aperfeiçoamento das técnicas de análise

transcricional em larga escala. Provavelmente essa característica também responda pela escassez ainda hoje de informações sobre a funcionalidade dessa classe de transcritos.

Dentre os ncRNAs melhor caracterizados encontram-se os microRNAs, descobertos na década anterior em *C. elegans*, sendo atualmente reconhecidos como uma grande família conservada de RNAs regulatórios com 22 nucleotídeos gerados a partir de percussores maiores, e que aparentemente são essenciais para uma variedade de processos celulares (Bartel, 2004; Esquela-Kerscher & Slack, 2006). Estimativas apontam para a existência de centenas de microRNAs no genoma humano, e essas análises computacionais ainda sugerem que mais de 20%–30% dos mRNAs codificantes para proteína humanos sejam regulados por essa classe de ncRNAs. Ilustrando a importância desses transcritos existe a observação que assinaturas de expressão de microRNAs, específicas de tecidos humanos, têm sido utilizadas para classificação molecular de câncer (Lu *et al.*, 2005), refletindo de maneira surpreendentemente mais acurada os estados de diferenciação tumoral do que perfis gerados com mRNAs.

Até mesmo os snoRNAs, que compreendem uma classe de ncRNAs que apresentam tamanho de 60 a 300 nucleotídeos, descritos inicialmente como funcionais apenas na modificação de rRNAs durante a biogênese dos ribossomos, hoje são vistos como mensagens mais participativas no programa transcricional de mamíferos tendo como alvos outros pequenos ncRNAs e até mesmo mRNAs (Kiss, 2002). Apesar da maioria delas, por serem derivados de introns de mRNAs, serem transcritas concomitantemente com outras mensagens (Kiss, 2002), aparentemente algumas são derivadas de unidades transcricionais independentes, pois apresentam caps de metilguanossina em suas porções 5' (Tycowski *et al.*, 2004). Uma fração

desses snoRNAs também exibe especificidade de tecidos e regulação durante o desenvolvimento (Cavaille *et al.*, 2000; Cavaille *et al.*, 2001; Rogelj & Giese, 2004), o que é um forte indicativo de papéis regulatórios. Funções regulatórias para snoRNA também são evidenciadas pelo seu papel no desenvolvimento da síndrome de Prader–Willi (Kishore & Stamm, 2006).

Outra classe de pequenos ncRNAs que apresenta potenciais papéis funcionais específicos no desenvolvimento foi recentemente descrita. A subfamília Piwi de proteínas Argonauta tem sido demonstrada como importante no desenvolvimento de linhagens germinativas (Cox *et al.*, 1998). RNAs com tamanhos de 26–31 nucleotídeos foram identificados como associados a essas proteínas em amostras de testículos murinos e foram denominados piRNAs (Girard *et al.*, 2006). Mais uma vez a especificidade tecidual é um ótimo indício de funcionalidade.

Existem ainda milhares de outros ncRNAs pouco caracterizados. Experimentos utilizando tanto *tiling arrays* genômicos, quanto RACE (*Rapid Amplification of cDNA Ends*) têm demonstrado que alguns ncRNAs provenientes de regiões intrônicas são mensagens longas, contendo de 400 a 2.000 bases (Rinn *et al.*, 2003; Reis *et al.*, 2004; Kapranov *et al.*, 2005). Enquanto a existência de vários ncRNAs longos já foi descrita em mamíferos (Okazaki *et al.*, 2002; Yelin *et al.*, 2003), poucos têm sido caracterizados em detalhes, sendo o RNA de ~15 kb, denominado *Xist* e que está envolvido na inativação do cromossomo X, uma óbvia exceção (Penny *et al.*, 1996). O potencial papel funcional de longos ncRNAs é exemplificado também pela associação do transcrito com cerca de 7 kb, *MALAT-1*, com câncer de pulmão metastático (Ji *et al.*, 2003), e pelo poder de classificação de amostras de tumores de próstata, com diferentes graus de diferenciação histológica, observado para o perfil de expressão de um conjunto de 23 longos RNAs intrônicos não-

codificantes (Reis *et al.*, 2004)³ As poucas funções específicas dessa e das demais classes de ncRNAs apresentadas aqui serão discutidas posteriormente no tópico apropriado.

Entretanto, é fato que pouco conhecimento encontra-se disponível com relação aos papéis biológicos precisos da enorme quantidade de RNAs, predominantemente sem potencial codificante de proteína, transcritos nas células humanas (Johnson *et al.*, 2005; Willingham & Gingeras, 2006). O mesmo pode ser dito com relação ao mecanismo celular de biogênese desses transcritos. Diversas abordagens de investigação indicam que os RNAs intrônicos estão envolvidos em processos de controle transcricional e pós-transcricional da expressão gênica (discutido no tópico 1.3.4). Assim, têm se postulado que sua transcrição controlada constitui um avanço fundamental na complexa rede de regulação da expressão gênica, fundamental para a ontogenia e evolução dos eucariotos (Mattick & Makunin, 2006).

1.3 Transcrição gênica e seus mecanismos de regulação

Tradicionalmente é reconhecido que os humanos, assim como todos os eucariotos, possuem três RNA polimerases dependentes de DNA (RNAP I, II e III), as quais apresentam particularidades e especificidades funcionais, e que seriam participantes centrais na produção de todos RNAs codificados pelo genoma nuclear (Howe, 2002). Uma quarta RNAP nuclear, que consiste de uma isoforma da RNAP mitocondrial humana, denominada spRNAP IV, foi descrita muito recentemente e também é responsável pela transcrição de uma população RNAs codificados pelo DNA do núcleo da célula (Kravchenko *et al.*, 2005). As RNAPs são facilmente

³ Reis *et al.* 2004 (Artigo publicado). ANEXO 3 - incluso no CD-ROM suplementar a esta Tese. Os resultados desse trabalho são apropriadamente apresentados e discutidos nessa Tese.

discriminadas a partir de sua sensibilidade ao octapeptídeo α -amanitina, que não afeta a atividade da RNAP I, inibe fortemente a RNAP II, apenas diminui a atividade da RNAP III (Jacob *et al.*, 1970; Kedinger *et al.*, 1970; Lindell *et al.*, 1970), e aparentemente até estimula a transcrição realizada pela spRNA IV (Kravchenko *et al.*, 2005).

A RNAP I e a RNAP III são em conjunto responsáveis por aproximadamente 80% da síntese de RNA em células em crescimento, sendo os rRNAs; e alguns pequenos RNAs como tRNAs, U6 snRNA, 7SL RNA, assim como o 5S rRNA, respectivamente transcritos por essas enzimas (Paule & White, 2000). Contudo, acredita-se que a RNAP II é a RNA polimerase responsável pela transcrição da grande maioria das regiões genômicas humanas, sendo sintetizados por ela os precursores dos RNAs mensageiros (pré-mRNAs), destinados a serem traduzidos em proteína no citoplasma pelo complexo ribossômico (Howe, 2002). Devido ao compartilhamento de características estruturais, como *splicing*, adição de *cap* 5' e poliadenilação 3', observado em parte do enorme repertório de ncRNAs e que são comuns aos mRNAs (Numata *et al.*, 2003), e pela correlação da existência de ncRNAs com a presença no genoma de sítios de ligação de proteínas e complexos protéicos necessários a transcrição pela RNAP II (Martone *et al.*, 2003; Cawley *et al.*, 2004; Kim *et al.*, 2005), assumiu-se que a expressão de ncRNAs também deva ser função da RNAP II. Entretanto, para mRNAs já existem fortes evidências de transcrição também pela recém-descoberta spRNAP IV (Kravchenko *et al.*, 2005). O envolvimento da spRNAP IV com a transcrição de ncRNAs ainda não foi estudado.

A transcrição é um processo cíclico com diversos passos intimamente regulados. Nos seus estágios iniciais o complexo de pré-iniciação (*Preinitiation Complex* - PIC) da RNA polimerase se liga ao promotor e a síntese de uma cadeia

de ribonucleotídeos começa (Lee & Young, 2000). A iniciação é rapidamente seguida pela liberação do promotor, passo onde os contatos entre a polimerase e a região promotora são rompidos (Dvir, 2002). Essa etapa é intermediária entre a elongação, onde a maior parte do transcrito é sintetizado, e a terminação onde acontece a liberação da polimerase do molde de DNA e do transcrito nascente, e inicia-se um novo ciclo pela enzima em outra região genômica (Howe, 2002).

O controle preciso do processo de transcrição, através da expressão gênica, é fundamental para fazer com que as informações contidas no genoma gerem os eventos biológicos, desde a proliferação celular até à diferenciação e desenvolvimento (Mata *et al.*, 2005). As células precisam integrar informações intrínsecas e ambientais e coordenar múltiplos mecanismos regulatórios, em diversos níveis, para poder exercer suas funções corretamente, sendo que erros ou defeitos nesse processo levam invariavelmente ao surgimento de doenças.

A etapa onde se acredita que ocorra a maior parte regulação da expressão gênica é a iniciação da transcrição. Os genes transcritos pela RNAP II contêm tipicamente duas famílias distintas de elementos regulatórios no DNA, que atuam em *cis*: um promotor, composto por uma região central e elementos regulatórios proximais; e outros elementos distais, que podem ser *enhancers*, *silencers*, *insulators*, ou regiões de controle no *locus* (*Locus Control Regions* - LCR) (Figura 2) (Maston *et al.*, 2006). Esses elementos em *cis* no DNA contêm sinais de reconhecimento e ligação para fatores de transcrição que atuam em *trans*, que funcionam tanto como repressores quanto estimuladores da transcrição (Maston *et al.*, 2006). Aparentemente cerca de 1.850 fatores de transcrição (TF) ligadores de DNA existem (Venter *et al.*, 2001), e é a presença desses múltiplos elementos e a

ligação combinatória desses TF que são responsáveis pela geração de padrões de expressão específicos (Maston *et al.*, 2006).

Figura 2 – **Esquema de uma região gênica regulatória típica.** O promotor é composto por uma região central e elementos regulatórios proximais. Outros elementos regulatórios distais podem ser *enhancers*, *silencers*, *insulators*, ou regiões de controle no *locus* (*Locus Control Regions* - LCR). Figura adaptada de Maston, G. A. *et al.* (2006). *Transcriptional Regulatory Elements in the Human Genome. Annu Rev Genomics Hum Genet*, 7:29–59.

1.3.1 Fatores de transcrição

Fatores proteicos envolvidos na acurada transcrição gênica pela RNAP II podem ser classificados em três grupos: fatores de transcrição gerais (*General Transcription Factors* - GTFs), ativadores, e coativadores (Figura 3) (Maston *et al.*, 2006). Os GTFs, que incluem a própria RNAP II e uma variedade de componentes auxiliares como TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, e TFIIH, e um complexo protéico denominado Mediador, se ligam a região central do promotor de maneira ordenada para formação do PIC, que direciona a maquinaria para o local de iniciação da transcrição (Maston *et al.*, 2006).

Após a ligação dos GTFs a transcrição já pode ocorrer em níveis basais, mas sua taxa é bastante aumentada após a ligação dos ativadores. Os ativadores são proteínas ligadoras de DNA que possuem especificidade de sequência e suas

regiões de reconhecimento ficam próximas à porção central do promotor (Ptashne & Gann, 1997). Comumente denominados apenas como fatores de transcrição, os ativadores existem em diversas famílias que discriminam domínios específicos de ligação no DNA, por exemplo: ativadores do tipo *zinc finger* ricos em cisteínas, *homeobox*, HLH (*helix-loop-helix*), *zipper* de leucina básico (bZIP), *forkhead*, ETS, e POU (Pit-Oct-Unc) (Pabo & Sauer, 1992). Os ativadores podem atuar aumentando a formação do PIC, na iniciação e alongação da transcrição, e ainda podem recrutar atividades que modificam a estrutura da cromatina facilitando a interação da maquinaria transcritcional com o promotor (Maston *et al.*, 2006).

Figura 3 – **Maquinaria transcritcional eucariótica.** Grupos de fatores envolvidos na transcrição pela RNAP II: fatores de transcrição gerais, ativadores, e coativadores. Figura adaptada de Maston, G. A. *et al.* (2006). Transcriptional Regulatory Elements in the Human Genome. *Annu Rev Genomics Hum Genet*, 7:29–59.

Os sítios no DNA para ligação dos ativadores, denominados de sítios de ligação para fatores de transcrição (*Transcription Factor Binding Sites* - TFBSs) são em geral pequenos com 6–12 bp, a especificidade é normalmente devido a não mais que 4–6 posições, em geral são bastante degenerados e descritos por uma seqüência consenso (Maston *et al.*, 2006). Essas diferenças nos TFBSs para cada ativador determinam a afinidade da ligação, o que tem muito impacto em sua atividade regulatória na taxa de transcrição (Maston *et al.*, 2006). A atividade dos ativadores pode ainda ser modulada pelo terceiro grupo de fatores protéicos envolvidos na transcrição gênica, os coativadores. Esses fatores não tem propriedades de ligação ao DNA, apresentando a propriedade de ligação direta aos ativadores (Lonard & O'malley, 2005).

Um dos exemplos de fatores de transcrição mais estudados são os receptores nucleares de hormônios, que são proteínas ativadoras que dependem de ligantes e incluem receptores para andrógenos, estrógenos, glicocorticóides, mineralocorticóides, progesterona e vitamina D (Claessens & Gewirth, 2004). O controle transcricional realizado pelos receptores nucleares envolve a ligação do hormônio, direcionamento para o núcleo e interação do complexo formado com regiões específicas do DNA denominadas elementos de resposta hormonal (Claessens & Gewirth, 2004). Tanto a seqüência de reconhecimento do elemento de resposta, quanto as vias reguladas são, em diferentes níveis, específicas para cada hormônio e seu receptor nuclear (Claessens & Gewirth, 2004). Em especial o andrógeno, e seu receptor nuclear, tem sido muito estudado devido sua implicação central na regulação de diversos processos fisiológicos em células normais e neoplásicas de próstata. Esse hormônio é reconhecido por promover divisão e proliferação celular, modular morte celular programada, quiescência celular e regular

o metabolismo (Isaacs *et al.*, 1992; Geck *et al.*, 1997; Swinnen & Verhoeven, 1998; Maffini *et al.*, 2002). A região de ligação do complexo andrógeno/receptor de andrógeno ao DNA é identificada pelo consenso palindrômico AGAACAnnnTGTTCT, denominado de elemento de resposta ao andrógeno (*Androgen Response Element – ARE*) (Shaffer *et al.*, 2004). A figura 4 apresenta um esquema ilustrativo desse processo.

Figura 4 – **Hormônio andrógeno e o seu receptor nuclear.** Diagrama esquemático do controle da transcrição gênica realizado pelo fator de transcrição formado pelo complexo andrógeno/receptor de andrógeno. O local de ligação ao DNA é identificado pelo consenso palindrômico AGAACAnnnTGTTCT, denominado de elemento de resposta ao andrógeno (ARE).

1.3.2 Modificações pós-transcricionais

A maturação dos transcritos primários da RNAP II no núcleo requer diversos eventos enzimáticos que incluem o processamento das porções terminais 5' e 3', e a remoção dos introns (Howe, 2002). E apesar desses eventos regulatórios serem classicamente denominados como modificações pós-transcricionais, tem sido demonstrado um acoplamento físico entre esses processos e a síntese da cadeia de ribonucleotídeos pela RNAP II, o que responde em parte pela especificidade dessas alterações nos mRNAs (Howe, 2002).

Um desses processos de maturação é a adição do *cap* 5', que estabiliza o mRNA contra degradação exonucleolítica 5'-3', promove a retirada dos introns e o processamento da região 3' terminal, facilita o transporte ao citoplasma, e auxilia na tradução (Shatkin & Manley, 2000). Essa modificação envolve 3 atividades enzimáticas: uma trifosfatase 5' que remove o fosfato γ do primeiro nucleotídeo transcrito, uma guanililtransferase que adiciona uma guanosina através de uma ligação trifosfato 5'-5', e uma 7 metiltransferase que modifica a guanina terminal pela adição de um radical metil (Shatkin & Manley, 2000). As proteínas envolvidas nessas reações formam o complexo de ligação do *cap* (*Cap Binding Complex* - CBC), que mantêm sua associação ao mRNA durante a transcrição, processamento e transporte para o citoplasma (Howe, 2002). De fato, essa modificação ocorre no pré-mRNA nascente logo após 25–30 bases do RNA terem sido polimerizadas e depois que a região 5' terminal já estiver fora do sítio de ligação de RNA da RNAP II (Howe, 2002).

Outra modificação é a remoção dos introns, o *splicing*, que é catalizado por um complexo ribonucleoproteico (RNP) denominado spliceossomo (Howe, 2002). Brevemente, a formação do complexo é iniciada pelo reconhecimento no pré-mRNA nascente de pequenos elementos em *cis* conservados, os *splice sites* e o *branchpoint*, localizados próximos ao final dos introns (Howe, 2002). Após o reconhecimento desses sinais, os *splice sites* 5' e 3' interagem através de RNPs, e diversos outros fatores (proteicos e RNPs) são adicionados de maneira ordenada formando o spliceossomo que após várias alterações conformacionais e reações químicas realiza o *splicing* (Howe, 2002). Essas etapas de reconhecimento envolvem ainda outras proteínas reguladoras que apresentam capacidade de se ligar ao pré-mRNA para regular a correta utilização de exons, os fatores de *splicing*,

o que possibilita diversas combinações de uso de exons (Howe, 2002). O uso de exons alternativos surge quando diferentes *splice sites* são reconhecidos e pareados para gerar transcritos maduros com distintos conteúdos de informação (Howe, 2002). Realmente, os *splicing* alternativos provêm um importante controle adicional na expressão gênica, permitindo as células tanto modular os níveis das proteínas expressas, quanto produzir múltiplas isoformas de proteínas a partir do mesmo *locus* gênico (Smith & Valcarcel, 2000). Mais uma vez esse processos ocorrem ainda durante a transcrição, o que parece ser fundamental para a acurácia e eficiência dessa modificação (Howe, 2002). A figura 5 ilustra e sumariza os processos descritos nesse parágrafo.

Figura 5 – **Processamento dos pré-mRNAs pelo spliceossomo.** (A) Reconhecimento e pareamento dos *splice sites* e do *branchpoint* (consenso para leveduras mostrado) através de interações com ribonucleoproteínas (RNPs) (em tons de azul e preto) para formação do spliceossomo. (B) Ligação de fatores de *splicing* (em amarelo, verde e roxo) ao pré-mRNA para regular a utilização de exons, permitindo os diversos padrões de *splicing*. Figura adaptada de Howe, K. J. (2002). RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochim Biophys Acta*, **1577**(2):308-324.

A modificação das regiões 3' terminais dos pré-mRNAs nascentes ocorre por um mecanismo de múltiplos passos que é iniciado pelo reconhecimento do sítio de poliadenilação por fatores de processamento (Howe, 2002). Duas atividades

químicas, mediadas por proteínas evolutivamente conservadas, são necessárias: quebra endonucleolítica do transcrito nascente e adição de uma cauda de poli-(A) por uma polimerase na região 3' clivada (Howe, 2002). O processo se inicia pelo reconhecimento dos sinais de poliadenilação (em geral AAUAAA), e pela clivagem de alguns nucleotídeos após essa região para criar o local de adição da cauda poli-(A) (Howe, 2002). Nessa modificação uma importante proteína ligadora de RNA, PABP (*Poly(A)-binding protein*), apresenta um papel estimulatório, e ainda regula o tamanho da cauda (Mangus *et al.*, 2003). De maneira similar a adição do *cap* 5', o processamento 3' aumenta a resistência dos transcritos a exonucleases e tem papéis funcionais no transporte e tradução dos mRNAs (Howe, 2002).

1.3.4 Proteínas ligadoras de RNA

Atualmente tem sido demonstrado que diversas proteínas ligadoras de RNA (*RNA Binding Proteins* – RBPs), assim como a PABP apresentam papéis centrais nos processos de regulação pós-transcricionais, representando para esses processos o mesmo que os fatores de transcrição representam para a transcrição (Sanchez-Diaz & Penalva, 2006). As RBPs estão envolvidas em todas as etapas que conectam a transcrição dos mRNAs a sua tradução em proteínas, determinando o destino dos pré-mRNAs e RNAs aos quais se ligam alterando padrões de *splicing*, modificando sua estabilidade, seus níveis de tradução e sua localização celular (Siomi & Dreyfuss, 1997).

As RBPs são definidas pela presença de domínios de ligação de RNA característicos (RRM ou RBD, KH e dsRBD). Alguns exemplos intensamente analisados funcionalmente são: PABP (*Poly(A)-Binding Protein*), proteína que se liga

a cauda poli-(A) de mRNAs. É necessária para a síntese da cauda poli-(A) no núcleo, regulando seu tamanho e estimulando a maturação do mRNA (Mangus *et al.*, 2003). Alguns mRNAs também precisam se associar a PABP para serem exportados do núcleo (Mangus *et al.*, 2003); hTRA2 (*Human Transformer 2*), proteína ativadora de *splicing*. Esta RBP pode se ligar a elementos conhecidos como amplificadores exônicos de *splicing* (*Exonic Splicing Enhancer* - ESE) (Kondo *et al.*, 2004); EIF4E (*Eukaryotic Translation Initiation Factor 4E*), proteína envolvida no direcionamento dos ribossomos a estrutura do *cap* 5' dos mRNAs. EIF4E é o componente limitante do aparato de tradução nos eucariotos, sendo por esse motivo de fundamental importância para a regulação espacial e temporal da síntese protéica (Richter & Sonenberg, 2005); e HuD (*HU-Antigen D*), esta RBP se liga e estabiliza uma grande quantidade de mRNA neuronais específicos. Ela também está implicada na exportação do RNA sintetizado no núcleo e na tradução de mRNAs específicos que contêm elementos ricos em AU (*AU-rich elements*) na porção 3' não traduzida (Saito *et al.*, 2004).

1.3.5 Envolvimento de RNAs não-codificantes

Até recentemente apenas os ncRNAs infraestruturais, rRNAs e tRNAs tinham suas funções celulares conhecidas. Mas de uma forma geral outros ncRNAs já tiveram sua funcionalidade determinada, sendo demonstrados como participantes ativos de praticamente todos os níveis de regulação da expressão gênica em eucariotos (Mattick & Makunin, 2006). Especula-se que esses transcritos, particularmente os expressos a partir de introns de genes codificantes para proteína, representem um avanço fundamental na regulação gênica dos organismos superiores, por definirem como e quando certas trajetórias de diferenciação e

desenvolvimento celular seriam tomadas, assim como por responder por diferenças de complexidade entre espécies com programa de expressão protéico muito similares (Mattick, 2004). Porém, os mecanismos específicos que respondem por essa ainda desconhecida rede de regulação não foram totalmente elucidados.

Exemplos na literatura de regulação gênica por transcritos não-codificantes ainda são esparsos, mas têm sido coletados e reunidos em excelentes revisões atuais (Mattick, 2004; Reis *et al.*, 2005; Mattick & Makunin, 2006; Willingham & Gingeras, 2006), o que contribui muito para estimular estudos com ncRNAs, e facilita a geração de hipóteses e experimentos para a determinação da funcionalidade da grande maioria do produto transcricional humano. Alguns desses exemplos, separados por categorias funcionais, serão apresentados nos próximos parágrafos.

Pequenos ncRNAs regulatórios, snoRNAs e microRNAs

Parte do produto transcricional humano apresenta funcionalidade regulatória como pequenos ncRNAs, que pertencem principalmente a duas classes: snoRNAs e miRNAs/siRNAs. Os papéis na regulação da expressão gênica dos snoRNAs ainda são pouco estudados, pois assim como já mencionado, acreditava-se que essas mensagens tinham suas funções restritas a biogênese do ribossomo, através de modificações no rRNA (Mattick & Makunin, 2006). Mas aos poucos o envolvimento desses transcritos na síntese de DNA telomérico (Kiss, 2002), e em especial no processo de *splicing* de mRNAs codificantes de proteína (Kishore & Stamm, 2006), tem sido reconhecido. Essa última função contribui de forma importante na síndrome Prader-Willi (Kishore & Stamm, 2006). Aparentemente a perda de expressão do snoRNA *HBII-52*, que é transcrito a partir da região silenciada do cromossomo 15 classicamente reconhecida como a causa dessa doença congênita,

causa um processamento aberrante do mRNA que codifica o receptor de serotonina *5-HT2CR*, que tem influência na síndrome (Kishore & Stamm, 2006). O *HBII-52* exibe complementaridade de sequência com a região do elemento de silenciamento do exon Vb do mRNA de *5-HT2CR*, causando a inclusão dessa porção da mensagem (Kishore & Stamm, 2006).

Provavelmente os microRNAs são os ncRNAs que tem recebido maior atenção nos últimos anos. Esses transcritos são funcionais como sequências contendo aproximadamente 22 nucleotídeos derivados de moléculas precursoras maiores, tanto RNAs dupla-fita, quanto RNAs que formam estruturas tipo *hairpin* (Bartel, 2004). As funções regulatórias dos microRNAs envolvem a supressão da tradução através de pareamento parcial na região 3' não traduzida (UTR), no geral de 6 a 8 nucleotídeos (Lewis *et al.*, 2005), com mRNAs alvo codificantes de proteína, ou degradação de mRNAs através de pareamento perfeito e direcionamento ao complexo RISC, pelo fenômeno reconhecido como interferência de RNA (Du & Zamore, 2005).

Até pouco tempo, apenas a atuação na inibição da tradução era reconhecida em humanos, na verdade nos vertebrados de uma forma geral. Porém em 2004 foi descrita a associação da supressão da expressão genica de *HOX* por uma degradação do seu mRNA dirigida pelo pareamento perfeito com um microRNA também na porção 3'-UTR (Mansfield *et al.*, 2004; Yekta *et al.*, 2004). Outro exemplo de regulação por esse mecanismo de atuação tem sido reportado para o *locus* que sofre *imprinting*, *Rtl1/Peg11* (Davis *et al.*, 2005). A mensagem transcrita maternalmente antisenso-*Peg11* é processada em diversos microRNAs, que causam degradação mediada por RISC do mRNA paternalmente expresso *Rtl1/Peg11*, após ligação em um segmento na região codificante (Davis *et al.*, 2005).

A expressão dos microRNAs é regulada, e alguns desses transcritos tem sido demonstrados como centrais em vários processos do desenvolvimento, como expressão temporal controlada, proliferação celular, apoptose, metabolismo de gorduras, morfogênese cerebral, diferenciação muscular e divisão de células tronco (Mattick & Makunin, 2006). Assim, alterações aberrantes provocadas por microRNAs têm sido reconhecidas como causais em diversas doenças humanas como o câncer (Mattick & Makunin, 2006). O banco de dados miRBase (<http://microrna.sanger.ac.uk/>) apresenta uma lista com mais de 300 microRNAs identificados experimentalmente em humanos, e também uma lista com os transcritos alvos preditos para esses pequenos ncRNAs (Griffiths-Jones *et al.*, 2006).

Outras classes de ncRNAs, em especial os longos, sem processamento (*splicing* ou excisão similar a que ocorre em alguns pequenos ncRNAs como os microRNAs) e transcritos a partir da região intrônica da fita oposta de *loci* que contêm mRNAs codificantes de proteína, são reconhecidas. Tem sido postulado que a grande maioria desses ncRNAs tem papel na expressão gênica, atuando diretamente nos mRNAs de mesmo *loci* em *cis* (Reis *et al.*, 2005). E apesar de não existirem evidências concretas, ainda não se pode excluir a possibilidade que essas mensagens apenas modulem eventos regulatórios transcricionais ou pós-transcricionais após serem processadas em transcritos menores, e que atuem em *trans*, como os microRNAs. Nos itens seguintes será principalmente discutida a funcionalidade desse tipo ainda pouco estudado de ncRNA, e a figura 6 apresenta uma ilustração dos mecanismos postulados.

Figura 6 – **Mecanismos propostos de regulação da expressão gênica envolvendo Longos RNAs Intrônicos não-codificantes antisense.** (A) Transcrição de pré-mRNA codificante de proteína (alvo) e da mensagem intrônica não codificante antisense (asRNA). (B) Modulação da abundância de isoforma alternativa pelo controle do *splicing*. (C) Controle do perfil de metilação de ilhas CpG de regiões promotoras de genes alvos (**esquerda**) ou modulação da abundancia de RNAs por interferência transcricional (**direita**). Sítios de ligação para fatores de transcrição (TFBS), fatores de transcrição (TF), RNA polimerase II (Pol II), spliceossomo (SC), e DNA metiltransferase (DMt) estão indicados. Figura adaptada de Reis, E. M. *et al.* (2005). As Antisense Gets Intronic. *Omics*, **9**(1):2-12.

Longos RNAs Intrônicos não-codificantes antisense e *splicing* alternativo

Assim como brevemente descrito no tópico 1.3.3, a correta excisão dos introns de um pré-mRNA e a religação dos exons é dirigida pelo spliceossomo, que reconhece sítios nas bordas de introns e exons, assim como certas seqüências regulatórias localizadas principalmente dentro de regiões do mRNA ainda não processado (Black, 2003). A ligação de fatores a elementos regulatórios denominados ISE (*intronic splicing enhancer*) e ISS (*intronic splicing silencer*), está envolvida na modulação do processamento do mRNA, favorecendo ou bloqueando a formação do spliceossomo (Akker *et al.*, 2001; Bruno *et al.*, 2004). A utilização de oligonucleotídeos antisense sintéticos, complementares a essas regiões intrônicas regulatórias é capaz de aumentar tanto a utilização de um exon do gene *FGFR1*, aberrantemente retirado em amostras de glioblastoma humano (Bruno *et al.*, 2004), quanto a exclusão do exon 7 no mRNA do gene *SMN* (Miyaso *et al.*, 2003).

Assim, ficou embasada uma hipótese de atuação de ncRNAs provenientes de regiões intrônicas, naturalmente transcritos com orientação antisense, em mecanismos envolvidos na alteração do padrão de *splicing* de mensagens codificantes. De fato, o trabalho recente de Yan e colaboradores mostrou que após transfecção de células Jurkat e indução do aumento da expressão de *SAF*, um longo (~1,5kb) ncRNA intrônico antisense normalmente transcrito a partir do primeiro intron do gene que codifica para a proteína Fas, ocorrem modificações na freqüência do uso de exons do mRNA *FAS* (Yan *et al.*, 2005). A presença de *SAF* apresenta importantes conseqüências na homeostasia celular, pois as alterações desencadeadas no processamento de *FAS* ocorrem tanto no primeiro exon, que é adjacente ao intron que se sobrepõe ao ncRNA, quanto nos exons da região 3' final

(Yan *et al.*, 2005). Essas últimas mudanças levam a obtenção de mensagens que codificam para uma proteína Fas que não é mais capaz de cumprir seu papel no mecanismo de apoptose através da ativação de caspases, devido a ausência de uma porção carboxi-terminal na sequência protéica primária que determina seu ancoramento a membrana plasmática (Yan *et al.*, 2005). A figura 6B apresenta um esquema geral a respeito do mecanismo de atuação de ncRNAs antisenso nessa etapa da regulação da expressão gênica

Longos ncRNAs intrônicos antisenso e alterações no perfil de metilação de promotores de genes alvos

Modificações na estrutura da cromatina são processos que regulam a expressão gênica através da reorganização de nucleossomos e de alterações químicas nos resíduos de nucleotídeos do DNA (Huang *et al.*, 2003). Esses processos já foram bastante estudados, porém nos últimos anos surgiram evidências do envolvimento direto de moléculas de ncRNA nesses eventos (Akhtar *et al.*, 2000; Tufarelli *et al.*, 2003; Imamura *et al.*, 2004).

Proteínas que realizam transferência de grupos acetil de histonas apresentam um papel fundamental nos mecanismos de rearranjo da cromatina (Van Der Vlag & Otte, 1999; Gebuhr *et al.*, 2000). Nessas e em outras proteínas envolvidas em processos de remodelagem similares, é possível identificar um domínio conservado com propriedades de ligação a RNA, o *chromodomain* (Akhtar *et al.*, 2000; Jones *et al.*, 2000). Esse domínio conservado na histona acetil transferase MOF de *Drosophila sp* interage com o ncRNA *roX*, contribuindo para sua interação ao cromossomo X do macho durante a compensação de dose (Akhtar *et al.*, 2000). Acredita-se que fenômeno similar ocorra em humanos, mas existem poucos estudos

de inativação do cromossomo X e *imprinting* genômico, associados a metilação/desmetilação do DNA, que sugerem um mecanismo dependente de RNA (Lee *et al.*, 1999; Sleutels *et al.*, 2000; Sleutels *et al.*, 2002).

Evidências diretas em humanos de metilação dependente de ncRNAs foram apenas obtidas recentemente. Tufarelli e colaboradores observaram em um paciente afetado com alfa-talassemia um silenciamento não esperado do gene *HBA2* através da hipermetilação de seu promotor, desencadeada pela transcrição aberrante de um longo ncRNA antisenso, *LUC7L*, a partir do mesmo *locus* gênico (Tufarelli *et al.*, 2003). De forma similar, o ncRNA intrônico antisenso *KHPS1a* modula a metilação da ilha CpG do promotor de *SPHK1*, gene codificante de proteína de mesmo *locus* (Imamura *et al.*, 2004). Esse tipo de função está esquematizado na figura 6C (direita).

Interferência transcricional e ncRNAs antisenso

Quando dois RNAs são concomitantemente transcritos a partir de fitas opostas do mesmo *locus* gênico pode ocorrer o fenômeno de interferência transcricional devido a colisão das maquinarias de transcrição (Prescott & Proudfoot, 2002). Durante as observações iniciais desse fenômeno em *Saccharomyces cerevisiae* foi notado que a abundância de ambos RNAs é reduzida pela própria restrição espacial da elongação pela RNAP II (Prescott & Proudfoot, 2002).

Evidências da ocorrência natural desse fenômeno foram obtidas para os *loci* do receptor tipo 2 para fator de crescimento similar a insulina (*Insulin-like growth factor type-2 receptor*) *Igf2r*, e do RNA antisenso *AIR* (Wutz *et al.*, 1997), do fator eucariótico de iniciação 2a, *eIF2a* (Silverman *et al.*, 1992; Noguchi *et al.*, 1994), e do colágeno alfa I (Farrell & Lukens, 1995). Assim, um dos possíveis mecanismos de

atuação dos ncRNAs pode ser regulação através de interferência transcricional, ilustrado na figura 6C (esquerda).

1.4 Genômica comparativa e seqüências não-codificantes

Abordagens utilizando genômica comparativa têm sido largamente utilizadas para inferir conservação e funcionalidade para seqüências de DNA de múltiplas espécies. Esse tipo de análise realizada em larga escala tornou-se possível após o aperfeiçoamento das técnicas de sequenciamento que permitiram obter os genomas completos de mais de 80 organismos microscópicos, e de muitas espécies eucarióticas, incluindo uma levedura (*Saccharomyces cerevisiae*), um nematodo (*Caenorhabditis elegans*), uma mosca (*Drosophila melanogaster*), algumas plantas como *Arabidopsis thaliana*, e uma variedade de arroz (*Oryza sativa japonica*), e de alguns animais tais como o *Fugu ribripes* e humanos (*Homo sapiens*) (Frazer *et al.*, 2003). Nesse arsenal de informações genômicas podem ainda ser incluídos o camundongo (*Mus musculus*), o rato (*Rattus norvegicus*), recém sequenciados, e outros eucariotos como chimpanzés, e peixe zebra cujas seqüências de DNA completa devem ser obtida nos próximos anos, o que ampliou significativamente a possibilidade de realização de estudos comparativos de conservação entre espécies (Frazer *et al.*, 2003).

Baseados na premissa que as regiões mais conservadas que a taxa basal de mutação poderiam estar sobre algum grau de seleção positiva para sua manutenção, uma considerável quantidade de estudos tem usado estratégias comparativas para identificar novas regiões funcionais nos genomas (Makalowski *et al.*, 1996; Hardison *et al.*, 1997; Loots *et al.*, 2000; Desilva *et al.*, 2002; Margulies *et al.*, 2003). A enorme maioria desses estudos comparativos é direcionada para a

identificação de transcritos codificantes para proteína e longas regiões regulatórias conservadas no DNA, utilizando critérios bastante específicos para a determinação de conservação e inferência de funcionalidade. Abordagens de genômica comparativa utilizando o enorme repertório de RNAs não-codificantes (ncRNAs) identificados como transcritos a partir do genoma humano (apresentado no item 1.2) tornou-se um grande desafio ainda não completamente enfrentado.

Aplicando os critérios classicamente utilizados em genômica comparativa, como tamanho mínimo e grau de identidade, diversos trabalhos têm identificado regiões não-codificantes altamente conservadas, e tem postulado funcionalidade para elas. Assim, termos como *Ultra Conserved Regions* (UCR) (Sandelin *et al.*, 2004), *Highly Conserved Regions* (HCR) (Duret *et al.*, 1993), *Multi-species Conserved Sequences* (MCS) (Margulies *et al.*, 2003), *Conserved Non Coding Elements* (CNE) (Woolfe *et al.*, 2005), e seqüências *Conserved Non-Genic* (Dermitzakis *et al.*, 2005) têm sido utilizados para descrever essas regiões genômicas possivelmente funcionais. Estima-se que cerca de 65% dessas seqüências estejam em regiões intergênicas e 35% em regiões intrônicas do genoma humano (Dermitzakis *et al.*, 2005).

Entretanto apesar desses critérios de conservação de identidade terem sido considerados bons para anotação e validação de seqüências funcionais, devido ao viés mencionado para transcritos codificantes, eles podem falhar para identificar novos ncRNAs funcionais. De fato, entre os ncRNAs já é reconhecido que grandes níveis de conservação evolutiva de identidade de seqüência primária não são pré-requisitos para funcionalidade (Numata *et al.*, 2003; Ravasi *et al.*, 2006), já que estes trabalhos indicam um número potencialmente elevado de seqüências não-

codificantes com papéis celulares, expressas em diferentes organismos e possuindo baixa conservação entre si.

1.5 Expressão gênica e câncer

1.5.1 Parálogos e famílias gênicas

A elucidação da seqüência genômica (Lander *et al.*, 2001; Venter *et al.*, 2001), e os esforços para determinação do programa transcricional completo humano (discutido no item 1.1) sugerem que ainda existam mensagens codificantes para proteína a serem descobertas, na maioria transcritos com baixos níveis de expressão, mas que provavelmente são específicos de tecidos ou estágios de desenvolvimento (Ota *et al.*, 2004). E mesmo entre os cerca de 25.000 transcritos codificantes já anotados (She *et al.*, 2004), existe uma considerável fração pouco estudada funcionalmente. Abordagens em escala genômica para entender os processos celulares que envolvem as proteínas codificadas no genoma humano têm sido desenvolvidas, principalmente no contexto de doenças como o câncer, para determinação de novos alvos potenciais para o tratamento com drogas (Harris & Stevens, 2006).

Uma interessante estratégia vem sendo utilizada por diversos pesquisadores nos dias de hoje. *Chemogenomics* é um novo conceito cunhado a luz da nova era nas pesquisas científicas iniciada com o conhecimento do genoma humano, que almeja a identificação e descrição de todos os possíveis alvos genômicos associados às diversas patologias, e suas drogas relacionadas (Caron *et al.*, 2001; Harris & Stevens, 2006). Uma das metodologias utilizadas para alcançar esses ambiciosos objetivos é o estudo de famílias gênicas, definidas pela homologia entre os genes pertencentes, ao nível da estrutura primária da proteína, e que

preferencialmente apresentem propriedades funcionais similares (Caron *et al.*, 2001; Harris & Stevens, 2006). Os genes pertencentes à mesma família gênica podem ser parálogos, quando derivados de um ancestral comum, mesmo que não apresentem necessariamente as mesmas funções celulares (Jensen, 2001). Com isso o estudo de famílias gênicas para identificação de possíveis parálogos de genes envolvidos em doenças humanas, utilizando análise informática de seqüências e posterior caracterização bioquímica, é muito útil para o desenvolvimento de drogas atuantes em múltiplos alvos da mesma família, ou de várias drogas similares efetivas apenas em cada parálogo (Caron *et al.*, 2001; Harris & Stevens, 2006). Um exemplo bem sucedido de resultado obtido com esse tipo de abordagem é exemplificado a partir dos trabalhos realizados com a família das proteínas quinases (Cohen, 1999; Garcia-Echeverria *et al.*, 2000). Muitos inibidores foram desenvolvidos para suprimirem as atividades de várias proteínas dessa família gênica; serinas, treoninas e tirosinas quinases (Cohen, 1999). Embasado nesses estudos, Garcia-Echeverria e colaboradores desenvolveram inibidores específicos para cada tipo de proteína quinase (Garcia-Echeverria *et al.*, 2000).

No desenvolvimento de medicamentos, geralmente uma porcentagem significativa das drogas criadas para agir em um membro da família é ativa em outros membros, cujas funções celulares são freqüentemente diferentes, enquanto outra é apenas efetiva em certos membros de uma família, que possuem funções essencialmente semelhantes (Caron *et al.*, 2001; Harris & Stevens, 2006). Assim a identificação de novos membros parálogos de famílias gênicas envolvidas no surgimento e progressão de doenças pode também elucidar os motivos, e apontar soluções, para a existência de certos pacientes com características clínicas

semelhantes, mas que são refratários aos tratamentos convencionais normalmente efetivos.

Especificamente no surgimento de neoplasias humanas, uma família gênica tem chamado atenção pelo envolvimento de muitos membros em câncer de praticamente todas as origens teciduais, a superfamília Ras de pequenas GTPases monoméricas. De fato, os genes dessa família codificam para proteínas que modulam uma grande variedade de processos celulares, desde o controle do crescimento, rearranjos no citoesqueleto, sobrevivência e apoptose celular, até regulação do tráfego intracelular de vesículas (Bos, 1997; Pasqualato *et al.*, 2002a). Mais de 100 membros da superfamília têm sido descritos em células eucarióticas, e são tradicionalmente classificados em seis famílias (Ras, Rho, Rab, Arf, Ran, e Rad) devido a similaridade de sequência primária com os membros bem caracterizados (Paduch *et al.*, 2001; Oxford & Theodorescu, 2003). Apesar de apresentarem múltiplas funções celulares, todas as pequenas GTPases compartilham um domínio G estruturalmente conservado (Figura 7), que responde pelo mecanismo molecular de ativação (Vetter & Wittinghofer, 2001b). Essas proteínas passam por estados ativos e inativos quando ligadas a GTP ou GDP, respectivamente (Vetter & Wittinghofer, 2001b), e apesar de normalmente possuírem atividade GTPásica intrínseca, suas funções celulares são reguladas por diversos efetores (Figura 8), GAPs (*GTPase-Activating Proteins*), GEFs (*Guanine Nucleotide Exchange Factors*), e GDIs (*Guanine Nucleotide Dissociation Inhibitors*) (Paduch *et al.*, 2001). As funções biológicas das pequenas GTPases também são bastante dependentes de modificações pós-traducionais em sua cadeia de aminoácidos (Paduch *et al.*, 2001). Por exemplo, as famílias Ras, Rho e Rab contêm uma sequência C-terminal que sinaliza e permite modificações que são essenciais tanto para sua localização

celular, quanto para interação com seus efetores (Paduch *et al.*, 2001) Existem muitas possibilidades de alterações na região C-terminal, mas as mais importantes são: ligação covalente de lipídios, metilação e quebra proteolítica (Paduch *et al.*, 2001).

Figura 7 – **Diagrama esquemático do domínio G.** As cinco α -hélices (A1–A5), seis folhas β (B1–B6) e seis voltas de polipeptídeos (G1–G5) que formam o domínio G estão indicadas. Em (A) conformação típica das proteínas da superfamília Ras de pequenas GTPases monoméricas quando ligadas a GTP. Em (B) conformação após a hidrólise a GDP. As maiores alterações conformacionais ocorrem nas regiões do *switch* I e *switch* II. Em (C) é mostrado um esquema do mecanismo de ativação, onde os domínios *switch* I e II estão ligados ao fosfato γ através dos grupos NH da cadeia principal dos resíduos conservados Thr e Gly. A hidrólise do GTP permite que as regiões dos *switches* relaxem, alterando a conformação da enzima. Painéis A e B adaptados de Paduch, M. *et al.* 2001. Structure of small G proteins and their regulators. *Acta Biochim Pol*, **48**(4):829-850. Painel C adaptado de Vetter, I. R. & Wittinghofer, A. 2001. The guanine nucleotide-binding switch in three dimensions. *Science*, **294**(5545):1299-1304.

Como já mencionado, muitas proteínas da superfamília Ras têm sido associadas ao câncer. É documentado que mutações nos genes *RAS* (*H-RAS*, *N-RAS* ou *K-RAS*), protótipos da família, ocorrem em aproximadamente 30% de todos os cânceres humanos (Adjei, 2001). As Rho GTPases também têm sido bastante implicadas em processos que levam a progressão do câncer (Boettner & Van Aelst, 2002). Ainda, alguns membros das famílias Ras, Ral, Rit (Reuther & Der, 2000b), e Rho (Zohn *et al.*, 1998) apresentam comportamentos oncogênicos. Outros como Noey2/Arh1 (Yu *et al.*, 1999a; Luo *et al.*, 2003), Rap1a (Wani *et al.*, 1997b), Rerg (Finlin *et al.*, 2001), e Rig (Ellis *et al.*, 2002a) têm sido descritas como possuindo atividade de supressor tumoral.

Figura 8 – **Modo de ativação das pequenas GTPases monoméricas.** Proteínas GAP estimulam a baixa atividade GTPásica intrínseca para promover a conversão da forma ligada a GTP para a ligada a GDP. Proteínas GDI afetam a dissociação do GDP e mantém as enzimas no estado inativo. Proteínas GEF catalizam a troca do GDP por GTP. Figura adaptada de Paduch, M. *et al.* (2001). Structure of small G proteins and their regulators. *Acta Biochim Pol*, **48**(4):829-850.

Naturalmente, o envolvimento das GTPases relacionadas a Ras no desenvolvimento de tumores coloca essas proteínas como potenciais alvos para tratamentos. A estratégia principal tem como alvo impedir a associação dessas proteínas com membrana citoplasmática. Como mencionado no parágrafo anterior, certas proteínas dessa família estudadas funcionalmente apresentam modificações covalentes em sua estrutura primária, que são importantes para sua funcionalidade. E de fato, são algumas dessas alterações que possibilitam a interação com a face interna da membrana plasmática (Paduch *et al.*, 2001). Assim, estratégias para impedir essa associação têm sido intensamente avaliadas, principalmente através de inibidores de farnesilação que já estão em fases de testes clínicos (Adjei, Croghan *et al.*, 2003; Adjei, Mauer *et al.*, 2003; Rao *et al.*, 2004). Outras abordagens têm como objetivo interromper a via de transdução de sinal dessas proteínas inibindo Raf e MEK, inibir a sinalização autócrina (inibidor do receptor EGF), ou ainda a própria expressão das proteínas relacionadas a Ras, como H-Ras e c-Raf-1 (Cox & Der, 2002). Contudo, a existência de vias bioquímicas alternativas, provavelmente com envolvimento de outros membros similares da família não caracterizados, torna certos pacientes insensíveis aos tratamentos desenvolvidos, sendo uma das principais limitações para o desenvolvimento de metodologias de tratamento mais efetivas.

1.5.2 Marcadores moleculares de diagnóstico e prognóstico em câncer de próstata

O câncer de próstata (PrCa) é uma doença com alta prevalência e é a principal causa de morte por câncer em homens no ocidente (Calvo *et al.*, 2005). O surgimento e o desenvolvimento do PrCa é de fato um processo complexo que inclui

alterações cromossômicas e na expressão gênica, levando a proliferação celular, inibição da apoptose, invasão, angiogênese e metástase (Calvo *et al.*, 2005). Uma revisão detalhada da biologia molecular do PrCa pode ser encontrada no artigo (Abate-Shen & Shen, 2000).

O diagnóstico precoce de câncer de próstata vem sendo realizado principalmente utilizando uma combinação do exame clínico de toque retal e a análise dos níveis de PSA sérico, levando em conta principalmente a idade do paciente (Potter *et al.*, 2001). Para o prognóstico de evolução da doença, tem sido largamente utilizado o método histológico denominado *Gleason Grading System*, onde vários aspectos da arquitetura do tecido, que indicam o nível de diferenciação celular, são analisados e recebem valores de 1 a 5, conforme a diferenciação é perdida (Gleason, 1977). Essa análise é realizada em duas regiões do tecido, o primeiro e o segundo padrão que ocupem as maiores áreas da amostra, e a soma dos dois valores compõem o *Gleason Score* (GS) do tumor do paciente (Gleason, 1977). Amostras de pacientes que apresentam GS de 2 a 4 indicam pouca agressividade e um risco mínimo de morte por câncer de próstata, amostras com valores de GS de 5 e 6 apresentam baixa agressividade com um bom prognóstico, amostras com GS de 8 a 10 são muito agressivas e o prognóstico é ruim, e finalmente valores de GS 7 indicam tumores com agressividade intermediária e prognóstico bastante heterogêneo, de difícil predição (Albertsen *et al.*, 1998). A figura 9 apresenta um esquema do *Gleason Grading System*.

Entretanto, existem severas limitações tanto na metodologia de diagnóstico precoce do PrCa, quanto na utilização do GS como fator de prognóstico. Já é bastante documentado que os níveis sanguíneos de PSA também podem variar devido a fatores não neoplásicos (Oesterling *et al.*, 1997; Carroll, 2001). Como

qualquer procedimento histológico, a avaliação do GS é bastante subjetiva. E por sua vez, a utilização do GS como fator de prognóstico está sujeita a erros amostrais e só pode ser realizada após a remoção cirúrgica radical da próstata. Assim, a busca e identificação de marcadores moleculares pode gerar importantes contribuições para complementar a avaliação clínica dessa doença.

Figura 9 – **Classificação de diferenciação tumoral em próstata por *Gleason Grading System*.** O *Gleason Score* (GS) é dado pela soma dos valores (de 1 a 5) referentes as duas maiores porções do tumor, segundo a característica da arquitetura celular do tecido. Amostras que apresentam GS de 2 a 4 indicam pouca agressividade e um risco mínimo de morte do paciente por câncer de próstata, amostras com valores de GS de 5 e 6 apresentam baixa agressividade com um bom prognóstico, amostras com GS de 8 a 10 são muito agressivas e o prognóstico é ruim, e finalmente valores de GS 7 indicam tumores com agressividade intermediária e prognóstico bastante heterogêneo, de difícil predição. Figura obtida de www.psc.edu/science/Wetzel/images/gleason.png.

Nos últimos anos diversas alterações em transcritos envolvidos em vias de sinalização hormonal, na expressão de oncogenes e supressores tumorais, de fatores de crescimento e seus receptores, de moléculas de adesão e reguladoras de angiogênese têm sido identificadas utilizando a abordagem de análise de expressão

gênica em larga escala por *microarrays* (Calvo *et al.*, 2005). Diferentes trabalhos têm demonstrado a existência de mensagens transcritas cuja expressão diferencial discrimina entre amostras de carcinoma de próstata (PrCa) e amostras com hiperplasia benigna (Dhanasekaran *et al.*, 2001); entre amostras de PrCa e amostras normais de próstata (Luo *et al.*, 2002; Rhodes *et al.*, 2002); e entre amostras localizadas de PrCa e amostras metastáticas (Varambally *et al.*, 2002; Ramaswamy *et al.*, 2003). Além de um potencial uso futuro como marcadores de diagnóstico, as listas de mensagens diferencialmente expressas geradas por essas abordagens podem ajudar na compreensão da biogênese desse processo neoplásico (Calvo *et al.*, 2005).

Similarmente, perfis de expressão gênica também são capazes de classificar subgrupos de pacientes com PrCa cujos tumores apresentam distintos graus de diferenciação histológica dado pelo GS (Singh *et al.*, 2002), assim como de pacientes que apresentam diferentes respostas as terapias atuais (Glinsky *et al.*, 2004). Esse tipo de abordagem pode ter importantes implicações por possibilitar a identificação de mensagens cuja expressão diferencial pode servir como um fator determinante de prognóstico. O prognóstico já poderia ser avaliado durante o diagnóstico, pois o perfil de expressão distingue molecularmente tanto tumores mais agressivos dos menos agressivos, quanto os pacientes que seriam refratários dos pacientes que responderam a terapia utilizada (Calvo *et al.*, 2005).

2 OBJETIVOS E RELEVÂNCIA

No Brasil, no período de 1998 a 2000, o Ministério da Saúde apontava para as neoplasias malignas como a terceira principal causa de morte, suplantada apenas pelas doenças cardiovasculares e por causas externas. Segundo as estimativas realizadas pelo Instituto Nacional de Câncer (INCA) em 2003 (<http://www.inca.org.br/estimativas/2003>), cerca de 400.000 brasileiros desenvolveram algum tipo de câncer, com aproximadamente 130.000 óbitos. Desse total, quase 40.000 seriam novos casos de câncer de próstata, com a morte de aproximadamente 9.000 homens, fazendo com que esse tipo de câncer se apresente com a maior taxa de incidência e a segunda maior taxa de morte por câncer em homens no Brasil, ficando atrás apenas dos tumores malignos de pulmão (Figura 10). Esses dados estatísticos ainda mostravam que, desde 1979, a taxa de mortalidade por câncer de próstata aumentou 141%.

Assim, inserido no contexto pós-genômico amplamente discutido na introdução desse texto, o estudo que foi desenvolvido para a obtenção do título de doutor em bioquímica teve como objetivo inicial a identificação de novos transcritos humanos potencialmente interessantes como marcadores de malignidade para o câncer próstata. O trabalho fez parte do projeto “Cooperação para Análise da Expressão Gênica – sub-área Câncer”, financiado pela Fapesp (Processo nº 99/07390-0), que foi desenvolvido sob coordenação do Prof. Dr. Sergio Verjovski-Almeida para explorar o potencial de informação gerado durante o projeto de sequenciamento em larga escala de ESTs humanas presentes em mais de 20 tipos diferentes de tumores e tecidos normais (HCGP, *Human Cancer Genome Project*). A partir de 2002, esse estudo passou a fazer parte do projeto temático para

“Identificação de marcadores moleculares para diagnóstico e prognóstico em câncer utilizando *microarrays* de DNA” (Processo nº 02/13283-6).

Figura 10 – **Câncer de próstata no Brasil no período de 1998 a 2003.** Taxa de morte em homens brasileiros. Dados computados pelo Instituto Nacional de Câncer (INCA) no período de 1998 a 2000, e estimativas realizadas em 2003 (<http://www.inca.org.br/estimativas/2003>).

Posteriormente uma caracterização mais aprofundada de alguns dos transcritos interessantes identificados pelas diferentes abordagens utilizadas, tanto no contexto de câncer quanto em situações normais de homeostase, foram incluídos entre os objetivos principais, descritos a seguir:

- Identificação e caracterização da expressão de novos transcritos potencialmente interessantes no contexto do câncer de próstata. Entre os novos genes estudados incluem-se parálogos de genes já descritos na literatura como envolvidos em câncer, e mensagens sem anotação funcional, mas diferencialmente expressas em tumores de próstata tanto em relação ao tecido prostático não-tumoral quanto em distintos graus de diferenciação

histológica. Espera-se que esses novos transcritos possam ser marcadores de malignidade (diagnóstico e prognóstico) e/ou alvos para o desenvolvimento, no futuro, de drogas contra o câncer;

- Caracterização da biogênese e do envolvimento em eventos regulatórios transcricionais e pós-transcricionais da nova classe de Longos RNAs não-codificantes, derivados de regiões intrônicas de *loci* gênicos que também expressam mensagens codificantes de proteína, identificados pela abordagem de *microarrays*. Para esse objetivo, células humanas foram estimuladas com o hormônio andrógeno, foram transcricionalmente inibidas pela adição de α -amanitina, e transfectadas com vetores para expressão episomal de proteínas ligadoras de RNA;
- E finalmente, o início de uma caracterização funcional, da especificidade tecidual e da conservação de expressão em outras espécies para os Longos RNAs Intrônicos não-codificantes. Espera-se que o entendimento da funcionalidade desses ncRNAs ajude na compreensão do complexo programa transcricional humano.

3 MATERIAIS E MÉTODOS

3.1 Análises Informáticas

3.1.1 Identificação *in silico* de parálogos de genes relacionados com câncer

Para o estudo de famílias gênicas foram utilizadas cerca de três milhões de etiquetas de seqüências transcritas (ESTs) contidas no GenBank, depositadas por dois projetos independentes de sequenciamento em larga escala de tumores, o CGAP (*Cancer Genome Anatomy Project*), lançado em 1997 pelo *National Cancer Institute*, e o HCGP (*Human Cancer Genome Project*), lançado pela FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo) e pelo Instituto Ludwig de São Paulo em 1999. Foi realizada uma busca computacional por novos membros de famílias gênicas relacionadas à tumorigênese, e para isso foi usada uma lista de 1.127 genes com envolvimento reconhecido em câncer, curada manualmente durante um *workshop* realizado em São Paulo em agosto de 2001. A lista contém genes codificantes de proteínas participantes de processos de angiogênese, crescimento celular, sinalização celular, transporte intracelular, entre outros. Dentre os genes dessa lista 36 membros da superfamília Ras de pequenas GTPases monoméricas estavam representados.

As seqüências protéicas codificadas pelos genes da lista foram utilizadas em uma busca por similaridade contra as ESTs do HCGP-CGAP traduzidas nas seis fases de leitura possíveis, utilizando a ferramenta tBLASTn (NCBI). Candidatos a novos parálogos foram identificados com base na conservação dos resíduos de aminoácido. As ESTs que apresentassem no mínimo 50% de identidade, com uma cobertura de pelo menos 50 resíduos, foram selecionadas, propiciando a distinção

entre proteínas de uma mesma família e proteínas que apenas compartilham pequenos domínios comuns (Futreal *et al.*, 2001). A sequência de nucleotídeos das ESTs que respeitam o critério acima ainda foram comparadas com um banco contendo cerca de 100 mil transcritos humanos já conhecidos utilizando BLASTn (NCBI), para filtrar as ESTs que se alinhassem com 100 % de identidade de bases contra seqüências deste banco, o que indicaria que representavam os próprios genes presentes na lista inicial, e não parálogos. Um esquema completo da metodologia de busca de novos parálogos de genes humanos envolvidos com processos de tumorigênese está apresentado na figura 11.

Figura 11 – **Identificação *in silico* de novos parálogos de genes humanos relacionados com câncer.** Diagrama esquemático dos processos utilizados para identificação de parálogos. Tradução nas 6 fases de leitura possíveis, do conjunto de cerca de 3 milhões de ESTs, comparação com as 1.127 proteínas relacionadas com câncer, (tBLASTn, 50% de identidade, em cerca de 50 aminoácidos), e filtragem das ESTs derivadas dos próprios genes utilizados na busca (BLASTn).

3.1.2 Análises dos produtos protéicos hipotéticos de *RASL11A* e *ARL-9*

Os produtos protéicos hipotéticos, Rasl11a e Arl9, foram utilizados em buscas por proteínas similares presentes no banco de dados de proteínas do GenBank,

utilizando BLASTp (NCBI). Arquivos *multifasta* foram construídos a partir das saídas individuais do BLASTp (NCBI), contendo as novas seqüências e as 50 proteínas humanas mais similares. Para uma análise de ortólogos, arquivos similares foram gerados contendo apenas seqüências provenientes de outros organismos. Esses arquivos foram manualmente inspecionados para redução da redundância e também para possibilitar a adição de alguns membros representativos de famílias menos relacionadas durante a análise das proteínas humanas. As listas finais foram utilizadas na geração de alinhamentos múltiplos de proteína e na reconstrução árvores filogenéticas da superfamília de pequenas GTPases monoméricas.

Todos os alinhamentos e a construção de árvores filogenéticas foram realizados através do programa ClustalX v 1.83 (Thompson *et al.*, 1997), usando os parâmetros padrão. As árvores filogenéticas não enraizadas foram construídas pelo método de *Neighbor Joining*, excluindo as regiões não alinháveis (*gaps*), a confiança dos ramos foi determinada através de permutações (1.000 amostragens), e finalmente, essas foram visualizadas utilizando o programa Tree View v. 1.6.6 (Page, 1996).

3.1.3 Análises Estatísticas para experimentos de expressão gênica com *microarrays*

Marcadores de prognóstico de tumores de próstata

Os valores de intensidade de expressão de alta qualidade (ver item 3.5.5) foram inicialmente analisados utilizando algoritmos de agrupamento para verificação se as amostras diferentes também possuíam perfis globais de expressão distintos. Agrupamentos não supervisionados dos dados foram realizados tanto por *Self-*

Organizing Map (SOM) usando o pacote de análise *GeneCluster2* (Tamayo *et al.*, 1999), ou por *Hierarchical clustering* (Sanchez-Carbayo *et al.*, 2002), através de distância Euclidiana e *Ward linkage* utilizando o *Spotfire Decision Site In Function Genomics* (Spotfire Inc.). Amostragens por permutações foram utilizadas para confirmar a robustez dos agrupamentos observados, onde 10.000 cópias dos dados foram geradas por adição de ruído Gaussiano aos dados originais (Sanchez-Carbayo *et al.*, 2002). Um dendograma consenso foi construído, mostrando a relação entre os nós mais freqüentes nas 10.000 amostragens.

A seleção supervisionada de mensagens cujos perfis de expressão eram significativamente correlacionados ao grau de diferenciação tumoral dado pelo *Gleason Score* (GS) (Gleason, 1977), foi realizada por análise de correlação de Pearson, de maneira similar a descrita anteriormente (Singh *et al.*, 2002). A significância estatística dessa análise foi assegurada pelo cálculo das correlações após 50.000 permutações randômicas dos rótulos de GS, sendo computada a freqüência na qual a correlação medida no grupo original foi também observada nos dados permutados. O intervalo de confiança escolhido foi de $p < 0,001$, onde número esperado de mensagens identificadas ao acaso seria apenas de sete em uma distribuição bicaudal. A aplicação do classificador baixo/alto GS ao perfil de expressão de dezesseis amostras com grau intermediário de diferenciação tumoral (GS 7) foi também realizada por um método supervisionado similar ao utilizado anteriormente (Van 'T Veer *et al.*, 2002).

A robustez do classificador baixo/alto GS foi avaliada por uma abordagem de *leave-one-out* (Singh *et al.*, 2002; Best *et al.*, 2003; Lapointe *et al.*, 2004). Essencialmente, cada paciente foi removido da análise e um novo grupo de mensagens correlacionadas ($p < 0,001$) foi determinado com as amostras restantes.

Transcritos regulados pelo tratamento com andrógeno

As hibridizações para busca de transcritos regulados por andrógeno foram realizadas com uma amostra referência marcada com Cy3 (ver item 3.5.4). Entretanto apenas os valores de intensidade provenientes das amostras de LNCaP marcadas com Cy5 foram utilizadas para determinar as razões entre tratamento (andrógeno) e controle (etanol) durante a análise de expressão diferencial.

A determinação dos transcritos com mudanças de expressão significativa em pelo menos três tempos consecutivos de tratamento com andrógeno foi realizada pela metodologia de *Significance Analysis of Microarray* (SAM) (Tusher *et al.*, 2001). Os parâmetros utilizados nessa abordagem foram: *fold-change* $\geq 1,5$, *K-Nearest Neighbors Imputer* e taxa de falsa descoberta (FDR) $< 0,05$.

Transcritos expressos em tecidos humanos e de camundongos

Foram analisados três diferentes tecidos, como descrito adiante. Os transcritos expressos em pelo menos um tecido segundo os critérios utilizados (ver item 3.5.5), foram submetidos a uma análise de expressão diferencial utilizando uma abordagem complexa. O conjunto foi primeiro analisado pelo método *Significance Analysis of Microarray* (SAM) (Tusher *et al.*, 2001), resposta com múltiplas classes, 1.000 permutações, *K-Nearest Neighbors Imputer*, e taxa de falsa descoberta (FDR) $< 0,1$. Os considerados diferencialmente expressos entre tecidos por SAM, foram também analisados por ANOVA (*Analysis of Variance*), usando-se o limite de significância de $p \leq 0.001$.

Os transcritos considerados expressos diferencialmente pelos dois métodos foram selecionados para prosseguir no estudo.

Transcritos afetados pela inibição da RNA polimerase II

Para determinação dos transcritos diferencialmente expressos durante o tratamento com α -amanitina, molécula inibidora da RNA polimerase II, foi utilizada uma análise estatística combinada. Os valores normalizados das duas réplicas biológicas de LNCaP tratadas ou controle foram analisados pelo método *Significance Analysis of Microarray* (SAM) (Tusher *et al.*, 2001), resposta de duas classes não pareadas, com 1.000 permutações, *K-Nearest Neighbors Imputer*, e taxa de falsa descoberta (FDR) de 0,002 a 0,035; e de *Signal-to-Noise Ratio* (SNR) com 10.000 permutações ($p < 0,05$) assim como em estudo similar (Golub *et al.*, 1999).

O conjunto de transcritos com diferenças significativas de expressão após o tratamento com α -amanitina foi constituído apenas pelas mensagens identificadas simultaneamente por SAM e por SNR.

3.1.4 Identificação *in silico* de elementos de resposta ao andrógeno (ARE)

Cerca de 3kb de seqüência genômica, a 5' dos sítios postulados para o início de transcrição das 39 mensagens intrônicas identificadas pela análise de *microarray* como reguladas por andrógeno, foram utilizados em uma busca pelo consenso de ligação do receptor de andrógeno. Para os transcritos estendidos por RACE-PCR, o novo sítio de início de transcrição foi utilizado.

A seqüência palindrômica AGAACAnnnTGTTCT de elementos de resposta ao andrógeno (*Androgen Response Elements* – ARE) (Shaffer *et al.*, 2004) foi obtida do banco de dados de elementos de DNA regulatórios em eucariotos, o TRANSFAC (Matys *et al.*, 2003) para utilização nessa análise. Os motivos ARE identificados

foram alinhados utilizando o CLUSTAL W (Thompson *et al.*, 1994), usando os parâmetros padrão, e uma representação visual gráfica do consenso foi gerado pelo WEBLOGO (Crooks *et al.*, 2004).

3.1.5 Análises de enriquecimento de categorias de ontologia gênica (GO)

Buscas por categorias enriquecidas de ontologia gênica (GO) entre os transcritos identificados como diferencialmente expressos nas análises de *microarray* foram realizadas utilizando os programas EASE, *Expression Analysis Systematic Explorer* (Dennis *et al.*, 2003) ou eGOn, *Explore Gene Ontology* (<http://www.genetools.microarray.ntnu.no/egon/>). As comparações foram sempre realizadas com relação a uma lista mestra contendo a abundância de transcritos com as mesmas características de mapeamento representados na plataforma de *microarray* em cada categoria de GO. Para mensagens exônicas a comparação foi feita contra uma lista contendo 2.292 transcritos, enquanto a lista mestra de intrônicos possuía 822 transcritos.

Tanto para as mensagens exônicas quanto para intrônicas foi utilizado o número de acesso no GenBank de depósito da EST correspondente à sonda presente no *microarray* para atribuição dos GOs via UniGene (NCBI). Foi utilizada anotação de GO dessa maneira inclusive para os transcritos intrônicos, pois se postula uma atuação desses sobre os próprios transcritos codificantes de proteína de mesmo *locus* gênico (Reis *et al.*, 2005).

3.2 Amostras de tecido e linhagens celulares

3.2.1 Amostras de tecidos humanos

As amostras de tecidos tumorais e normais foram obtidas de pacientes que sofreram remoção cirúrgica total dos órgãos, e são provenientes de convênios com hospitais brasileiros. A utilização dessas nos estudos de expressão gênica descritos aqui foi previamente aprovada pelo comitê de ética das instituições de saúde que as forneceram, e foi formalmente consentida pelos próprios pacientes. Os tecidos foram congelados em nitrogênio líquido até 10 minutos após terem sido removidos dos pacientes, ficando assim armazenados até uma macro-dissecção, ou utilização direta para extração de RNA. Essa inspeção histológica, quando realizada visou definir e separar as porções tumorais e normais de cada peça.

As amostras de tecido prostático foram fornecidas pelo Hospital do Câncer/RJ e o Hospital Sirio-Libanês/SP. No total foram recebidas e armazenadas no banco de tecidos 55 pares de amostras de próstata (o tecido tumoral e seu não-tumoral adjacente) provenientes do Hospital do Câncer/RJ, e 7 amostras tumorais e 5 normais do Hospital Sirio-Libanês/SP. Apenas as amostras que apresentaram um mínimo de 80% de células malignas foram consideradas tumorais para utilização no estudo. As amostras que continham quantidade de tecido insuficiente para a extração e purificação do RNA foram eliminadas. Assim, prosseguiram nos experimentos 20 amostras normais adjacentes ao tumor e 27 amostras tumorais com características histológicas de diferenciação celular diversas: uma com *Gleason Score* (GS) igual a 5, cinco com GS igual a 6, dezesseis com GS igual a 7, uma com GS igual a 9, e quatro com GS igual a 10.

Os Hospital do Câncer/RJ também forneceu as 5 amostras de rim humano não-tumoral adjacente ao tecido tumoral, e 5 amostras de fígado humano normal

foram obtidas do Hospital das Clínicas da USP, de pacientes que receberam transplante.

3.2.2 Amostras de tecidos de camundongos

Amostras de fígado, próstata e rim foram extraídas de cinco camundongos (*Mus musculus*) machos adultos da linhagem BALB/c. Os órgãos dos animais também foram congelados em nitrogênio líquido até 10 minutos após terem sido removidos cirurgicamente, ficando assim armazenados até a extração de RNA.

3.2.3 Linhagens e culturas celulares

As linhagens celulares de carcinoma de próstata LNCaP, DU145 e PC3, a linhagem celular tumoral de tecido renal 293T, e a HepG2, de tumor de fígado humano foram obtidas da ATCC (*American Type Tissue Collection*) e mantidas com os meios sugeridos suplementados com 10% (vol/vol) de soro fetal bovino (FCS), 3mM L-glutamina, 100µg/ml estreptomicina e 100U/ml penicilina. As células LNCaP foram cultivadas em RPMI 1640 enquanto DU145, PC3, 293T e HepG2 em DMEM, para o crescimento em estufa de CO₂ a 37°C. As culturas foram realizadas em placas de petri de 150mm de diâmetro contendo 25ml de meio de cultura completo, atingindo geralmente 60-80% de confluência para replicação, extração de RNA ou realização de manipulações específicas de cada experimento.

Para os experimentos de resposta ao andrógeno as células LNCaP foram cultivadas por 24 horas em meio privado de hormônios, RPMI 1640 com 5% (wt/vol) soro fetal bovino absorvido com carvão (Life Technologies). Após 24h, as células foram plaqueadas em novo meio RPMI 1640 com 10% (wt/vol) soro fetal bovino absorvido com carvão, com adição de 1nM de andrógeno sintético R1881 (NEN -

Life Sciences Products) ou volume equivalente do veículo (etanol). As células foram coletadas após 0, 6, 9, 12, 18, 24, e 48 horas de tratamento para extração de RNA total, e após 4 horas para extração de DNA genômico.

Para os experimentos de imunoprecipitação com proteínas ligadoras de RNA (*RNA Binding Proteins* – RBPs), o crescimento das células 293T foi realizado em placas contendo poli-lisina para garantir uma melhor aderência celular devido as manipulações necessárias. Após atingirem a confluência desejada (60%) as células foram transfectadas com cerca de 20µg de vetor pKALLU com as construções de cada RBP, ou com GFP no controle negativo (ver item 3.9.4), por placa, usando o reagente *GeneJammer* (Stratagene) seguindo estritamente as condições descritas pelo fabricante. Brevemente, em 1ml de DMEM (Invitrogen) sem soro e sem antibióticos, para não prejudicar a formação das vesículas lipossômicas, foi adicionado *GeneJammer* (Stratagene). Após breve incubação à temperatura ambiente por 10 minutos, os vetores foram adicionados e a mistura foi incubada novamente por 10 minutos. Nas placas semiconfluentes com 293T, com 10ml de novo meio de cultura completo, foi gotejado o volume total da solução de lipossomos com o vetor, permanecendo dessa maneira por cerca de 3 horas, quando o volume de meio completo foi elevado até 25ml em todas as placas com células transfectadas.

Para os experimentos de inibição da RNA polimerase II, as células LNCaP foram cultivadas por dois dias (até atingir cerca de 8×10^5 células) quando os meios de cultura iniciais foram substituídos por novos contendo 50µg/mL α -amanitina (Roche) ou apenas veículo. Após 24 horas as células foram coletadas e armazenadas a -80 °C até a extração do RNA. Os procedimentos foram repetidos para obtenção de uma réplica biológica.

3.3 Extração de RNA

As amostras de tecidos foram pulverizadas utilizando mortar e pistilo de porcelana para extração de RNA. Esse procedimento foi realizado em baixa temperatura, utilizando gelo seco e nitrogênio líquido. Todas as culturas de células foram lavadas uma vez com PBS, antes do início do protocolo de extração de RNA.

3.3.1 Extração de RNA poli-(A)+

A extração de RNA mensageiro (mRNA) foi realizada utilizando tanto tecidos pulverizados quanto células. As amostras foram lisadas por adição de tampão de lise na proporção de 40mg de tecido ou 1×10^7 células para cada 1ml, e a população de moléculas de mRNA foi purificada utilizando o *μMACS mRNA Isolation Kit* (Miltenyi Biotec). Esse procedimento foi realizado seguindo estritamente o protocolo do fabricante e baseia-se no pareamento das bases de oligonucleotídeos poli-dT, modificados com uma porção magnética, com as bases da cauda poli-(A) dos mRNAs eucarióticos. A separação é feita pela interação magnética dos complexos mRNA/oligonucleotídeos poli-dT com um ímã. Após diversas lavagem que eliminam proteínas, DNA e outros RNAs, os mRNAs são eluídos pela adição de H₂O a 65°C, que rompe as pontes de hidrogênio formadas entre as bases das caudas poli-A e as bases dos oligonucleotídeos poli-dT. Em média foram obtidos cerca de 5μg de mRNA poli-(A)+ a partir de cada 100mg de tecido processado.

3.3.2 Extração de RNA total

A extração de RNA total foi realizada utilizando tanto tecidos pulverizados quanto células. As amostras foram lisadas e o RNA extraído por adição de Trizol (Invitrogen) na proporção de 50mg de tecido ou 5×10^6 células para cada 1ml. Todo o procedimento seguiu o protocolo padrão do fabricante. O Trizol é uma solução ácida, monofásica de fenol e isotiocianato de guanidina que possibilita a extração de RNA através dos mesmos princípios do método original desenvolvido por Chomczynski e Sacchi (Chomczynski & Sacchi, 1987). Para cada 50mg de tecido ou 5×10^6 células, foram obtidas de 10 a 30µg RNA total.

Para os experimentos de expressão gênica após estímulo com andrógeno, RNA total de células LNCaP, tratadas ou controles, DU145 e PC3, foi extraído utilizando gradiente de CsCl (Chirgwin *et al.*, 1979), após lise com 4M isotiocianato de guanidina; 0,1M β-mercaptoetanol; e 25mM citrato de sódio pH 7,0. RNA total de células da linhagem tumoral de fígado humano HepG2 também foi extraído por gradiente de CsCl. A partir de uma placa de 150mm de diâmetro com de 50 a 80% de confluência, foi obtido de 70-100µg RNA total.

3.3.3 Extração de RNA co-imunoprecipitado com proteínas ligadoras de RNA

Os complexos de proteínas e RNA imunoprecipitados obtidos (ver item 3.8.2) foram tratados com proteinase K (600µl de uma solução livre de RNases contendo: 50mM Tris pH 7,5; 5mM EDTA; 25mM NaCl; 0,5% SDS e 500µg proteinase K) a 50°C por 30 minutos, em agitação constante para degradação proteica. Após uma centrifugação a 1.000×g por 3 min a 4°C, os RNAs co-imunoprecipitados foram

extraídos do sobrenadante com adição de 1 volume de fenol:clorofórmio, centrifugação a 13.000×g por 5 min a 4°C, adição de um volume de clorofórmio, e finalmente, outra centrifugação a 13.000×g por 5 min a 4°C.

As populações de RNA obtidas foram precipitadas usando 0,8 volumes de isopropanol, 80µl de 4M acetato de sódio ou amônia, 3µl de 1M MgCl₂, e 8µl de 20mg/ml glicogênio (Roche), e incubação a -80°C. As amostras de RNA precipitado foram centrifugadas a 13.000×g por 30 min a 4°C, lavadas com 200µl de etanol, secas por 10 min a temperatura ambiente e ressuspendidas em H₂O livre de RNases.

3.3.4 Tratamento com DNase livre de RNase

Contaminações com DNA genômico foram minimizadas nas preparações de RNA total e mRNA. Cerca 0,5µg de RNA poli-(A)+, foi tratado com 1 unidade (U) de *RQ1 RNase-free DNase* (Promega) por 45 minutos a 37°C, em um volume de 15µl na presença de 20U de *RNasin* (Promega), inibidor de RNase. Após inativação enzimática por adição da solução de parada com EDTA (*Stop Solution*) e aquecimento a 65°C por 15 minutos, as amostras de mRNA foram utilizadas diretamente nos experimentos subseqüentes.

As amostras de RNA total foram tratadas com a *RNase-free DNase I* (Qiagen). Cerca de 100µg de RNA total foram tratados e purificados seguindo estritamente o protocolo de digestão de DNA genômico nas colunas do *RNeasy* (Qiagen).

3.3.5 Quantificação e verificação da qualidade dos RNAs extraídos

Para determinação da qualidade e quantidade dos RNAs extraídos, todas as amostras tiveram suas absorbâncias medidas nos comprimentos de ondas 260 e 280nm no espectrofotômetro *ND-1000* (NanoDrop Technologies), e foram separadas por eletroforese utilizando o *2100 BioAnalyzer* (Agilent Technologies).

A quantificação do RNA foi feita utilizando o valor de sua densidade ótica (Abs) em 260nm (40µg/mL RNA possui uma $Abs_{260} = 1$). A razão 260nm/ 280nm foi utilizada para estimar a porcentagem de contaminação de proteína nas amostras de mRNA. Razões entre 1,6-2,0 foram consideradas empiricamente satisfatórias, devido a observação de reações de marcação por transcrição reversa com eficiência similar utilizando RNAs dentro dessa faixa.

As amostras de RNA foram também submetidas a uma eletroforese capilar para determinação simultânea da integridade da população de RNAs e da ausência de grande contaminação por DNA genômico. Essas análises baseiam-se na observação dos perfis dos eletroferogramas, e na comparação com o padrão *RNA 6000 ladder* (Ambion Inc.). A integridade das amostras de mRNA de próstata foi estimada pela maior concentração de moléculas de RNA de maior peso molecular. Para amostras de RNA total a razão da quantidade dos RNAs ribossomais 28S e 18S foi utilizada como parâmetro de qualidade, sendo aceitáveis valores acima de 1,1 para amostras de tecido e acima de 1,5 para amostras derivadas de linhagens celulares.

Para cada tecido prostático no estudo a quantidade de mRNA obtida foi em média de 1,0-2,5µg para cada 50mg processado, sendo que a maioria das amostras possuía no mínimo 100mg de tecido. Como saldo final, para as análises de expressão gênica foram obtidas amostras de mRNA de alta qualidade para 47

tecidos, 20 pares de tumores e suas respectivas regiões normais adjacentes e 7 amostras tumorais não pareadas. As amostras de RNA total provenientes de tecidos humanos ou de camundongo geraram cerca de 10-30µg para cada 50mg processado, enquanto o rendimento de RNA das amostras de linhagem celular foi de cerca de 50µg para cada 1×10^7 células processadas.

3.4 Abordagens de PCR para amplificação de transcritos completos

3.4.1 RACE-PCR

A metodologia para amplificação de extremidades desconhecidas de fragmentos de cDNA, RACE-PCR (*Rapid Amplification of cDNA Ends*), foi utilizada para obtenção das seqüências completas de parte dos novos transcritos humanos identificados. Amplificações por RACE das extremidades 5' e 3' de moléculas de cDNA foram realizadas utilizando bibliotecas comerciais *Marathon Ready cDNA* (BD Biosciences) geradas a partir de RNA poli-(A)+ de uma linhagem de adenocarcinoma de próstata (*Human XG Prostatic Adenocarcinoma*), de testículo (*Human Testis*), de pulmão (*Human Lung*), ou de uma biblioteca feita de RNA poli-(A)+ de colon, clonada direcionamente no vetor lambda gt11 (*Lambda gt11* - BD Biosciences).

Dois pares de oligonucleotídeos iniciadores específicos complementares as seqüências das ESTs candidatas a novos transcritos humanos foram desenhados ancorados em suas pontas, orientados para o centro da mensagem, e de forma a possibilitar a utilização de cada par em reações de PCR seqüenciais (*Nested*), que aumentam a especificidade dos produtos gerados. As reações de RACE-PCR com as bibliotecas *Marathon* foram realizadas em 50µl, com: 2µl de molde de cDNA; 0,2mM de dNTP; 1X *Advantage 2 PCR Buffer*, 1X *Advantage 2 Polimerase Mix* (BD Biosciences); 0,2µM de oligonucleotídeos iniciadores específicos antisense ou sense;

e 0,2 μ M do oligonucleotídeo iniciador universal (*Adaptor Primer*) complementar aos adaptadores existentes nas porções 5' e 3' de cada molécula de cDNA. As reações com as bibliotecas *Lambda gt11* foram similares, exceto pela adição de 0,2 μ M do oligonucleotídeo iniciador complementar a região do vetor próxima a porção 5' ou complementar a região próxima a porção 3' das mensagens clonadas, o que permitiu a determinação da orientação transcricional das mensagens amplificadas. Todas as reações foram incubadas no termociclador *GeneAmp PCR System 9700* (Applied Biosystems), utilizando as seguintes condições de ciclagem: 40 repetições de 95°C por 30 segundos para desfazer as duplas fitas, e de 68°C por 4 minutos para anelamento e extensão, finalizando com 68°C por mais 10 minutos.

Os produtos obtidos foram ligados no *pGEM-T Easy* (Promega) para sequenciamento dideoxi (MegaBace – Armeshan) utilizando oligonucleotídeos iniciadores complementares as regiões promotoras das RNA polimerases T7 e SP6 presentes na seqüência do vetor, flanqueando o inserto. Confirmadas as identidades das porções 5' e 3' obtidas, através de mapeamento no genoma humano, esses foram unidos com uma única reação de amplificação em condições idênticas as descritas acima, porém utilizando ambos fragmentos em quantidades equimolares como molde, e um novo par de oligonucleotídeos iniciadores ancorados nas extremidades do novo transcrito humano. Procedimentos similares de clonagem e sequenciamento foram repetidos com as mensagens completas.

3.4.2 PCR com oligonucleotídeos iniciadores específicos

Para caso particulares onde a metodologia de RACE-PCR não produziu resultados interpretáveis, foram utilizadas abordagens de PCR com oligonucleotídeos iniciadores específicos desenhados utilizando a informação

contida em predições genômicas para novos genes humanos codificantes de proteína, ou em RNAs já depositados no GenBank. Novamente, as bibliotecas comerciais *Marathon Ready cDNA* (BD Biosciences) geradas a partir de RNA poli-(A)+ de uma linhagem de adenocarcinoma de próstata (*Human XG Prostatic Adenocarcinoma*), de testículo (*Human Testis*), e de pulmão (*Human Lung*) foram utilizadas para esse propósito.

As condições de amplificação também foram idênticas às relatadas para RACE-PCR (ver item 3.4.1), exceto na adição simultânea de 0,2 μ M do oligonucleotídeos iniciadores específicos antisense e senso, e na utilização de uma temperatura independente e mais baixa para anelamento (65°C) durante os ciclos. Os produtos obtidos foram também ligados no *pGEM-T Easy* (Promega) para sequenciamento dideoxi (MegaBace – Armeshan) utilizando oligonucleotídeos iniciadores complementares as regiões promotoras das RNA polimerases T7 e SP6 presentes na seqüência do vetor que flanqueia os insertos.

3.5 Estudo de expressão gênica por *microarrays* de cDNA

3.5.1 Plataforma de *microarrays* de cDNA do IQUSP

Parte dos estudos de expressão gênica descritos aqui fizeram parte do projeto CAGE (Cooperação para Análise da Expressão Gênica), sub-área de câncer, financiado pela FAPESP (Processo nº 99/07390-0) (<http://bioinfo.iq.usp.br/cagecancer/>). Com esse financiamento foi construído um *microarray* de cDNA cujas características encontram-se publicadas no banco de dados *Gene Expression Omnibus* (GEO – <http://www.ncbi.nlm.nih.gov/geo/>), sob o número de acesso GPL3985.

Brevemente, a plataforma foi construída após seleção de sondas a serem depositados nas lâminas por análises bioinformáticas com as 1.016.501 seqüências expressas (ESTs) obtidas pela metodologia ORESTES (Dias-Neto *et al.*, 2000) a partir de diversos tipos de câncer durante o Projeto Genoma Humano do Câncer (HCGP - <http://www.ludwig.org.br/ORESTES>). Essa plataforma contém duplicatas de 4.608 sondas de cDNA provenientes dos produtos amplificados das ESTs do HCGP selecionadas, com um total de 9.216 elementos depositados em lâminas de vidro *Type 7 STAR* (Amersham Biosciences Inc.) utilizando o *Generation-III Array Spotter* (Molecular Dynamics). Dentre as sondas depositadas encontram-se 1.301 produtos amplificados de clones cuja EST associada não apresenta similaridade com seqüências completas humanas anotadas. Um conjunto de 2.385 produtos derivados de ESTs similares a seqüências relacionadas com desenvolvimento de câncer, e 922 produtos de genes constitutivos (*housekeeping*) e controles negativos e positivos para os procedimentos de hibridização, também foram depositados. Dentre os controles, 384 cDNAs são provenientes do *Lucidea Microarray ScoreCard* (Amersham Bioscience), onde além de controles positivos e negativos (cDNA de planta), existe um outro conjunto de fragmentos de DNA intergênico de levedura que não hibridizam com cDNAs humanos.

Do conjunto de transcritos humanos sem similaridade com transcritos conhecidos encontram-se 822 transcritos derivados de ESTs que mapeiam em regiões intrônicas de mensagens anotadas, 241 derivados de seqüências que mapeiam em regiões sem mensagens conhecidas (intergênicas). Detectamos também 238 transcritos contaminantes de outros organismos, que durante a seleção não foram detectados pelos filtros informáticos para seqüências, e agora foram eliminados das análises.

Nas lâminas as sondas estão organizadas em 12 *sub-arrays*, cada um contendo 12 linhas e 32 colunas (384 sondas/*sub-array*). Em cada *sub-array*, as linhas de 1 a 5 contêm sondas para mensagens humanas sem similaridade com transcritos conhecidos; as linhas de 6 a 10 contêm as sondas para mensagens humanas conhecidas; e finalmente, as linhas 11 e 12 de cada *sub-array* contêm sondas para os 32 controles do *ScoreCard* e 32 sondas para conhecidos marcadores de próstata e genes *housekeeping*.

3.5.2 Obtenção de alvos fluorescentes

Marcação direta para geração de alvos fluorescentes

Os alvos fluorescentes foram obtidos por marcação direta usando o *CyScribe First-Strand cDNA Labelling Kit* (Amershan Biosciences), através de transcrição reversa na presença de dCTP modificado com Cy3, fluoróforo que emite uma fluorescência verde (550nm de absorção e 570nm de emissão) ou Cy5, que emite uma fluorescência vermelha (650nm de absorção e 670nm de emissão).

Brevemente, 15µg de RNA total combinado apenas com oligo-dT, ou 1µg de RNA poli-(A)+ com oligo-dT e nonâmeros randômicos (9 bases) foram utilizados para síntese de cDNA, seguindo estritamente o protocolo do fabricante, usando-se 100 unidades de transcriptase reversa (*CyScript reverse transcriptase*) a 42°C, durante 90 minutos. Os moldes de RNA foram removidos por hidrólise alcalina com 2µl de 2,5M NaOH a 37°C por 15 minutos, e em seguida as reações foram neutralizadas com 10µl de 2M MOPS. Cinco volumes de tampão de purificação (5,3M Guanidina-HCl; 150mM KOAc) foram adicionados às amostras para promover a eliminação dos nucleotídeos modificados não incorporados usando *96-well*

Multiscreen Filter Plates (Millipore). Quatro lavagens com etanol 80% foram realizadas por centrifugação das placas a 3.000 x g por 5 minutos, e uma última centrifugação por 5 minutos foi feita para eliminação de etanol residual. Os alvos marcados foram eluídos das placas com 10mM Tris pH 8,5 através de centrifugação a 3.000×g por 5 minutos, desidratados totalmente em uma *SpeedVac* (Eppendorf) e mantidos a -20°C protegidos da luz até o uso.

Em todas as reações de marcação realizadas foram incluídos 2,5µl da mistura de mRNAs controle contidos no *Lucidea Microarray ScoreCard* (Amersham Bioscience). Esses mRNAs são complementares ao grupo de 32 seqüências de levedura depositadas no *microarray* de cDNA (detalhes no item 3.5.1).

Marcação indireta para geração de alvos fluorescentes

Os alvos fluorescentes foram obtidos por marcação indireta usando *CyScribe Post-Labeling Kit Protocol* (Amersham Biosciences), e *SuperScript™ Indirect cDNA Labeling System* (Invitrogen) através de transcrição reversa utilizando aminoallyl-dUTP, e acoplamento a esses resíduos de uracila modificados, respectivamente, com Cy5 ou *AlexaFluor 647* (Molecular Probes). Ambas moléculas emitem uma fluorescência vermelha (650nm de absorção e 670nm de emissão).

As reações com o *CyScribe Post-Labeling Kit Protocol* (Amersham Biosciences) utilizaram 10-15µg de RNA total combinado apenas com oligo-dT para síntese de cDNA seguindo estritamente o protocolo do fabricante, usando-se 100 unidades da transcriptase reversa (*CyScript reverse transcriptase*) a 42°C, durante 90 minutos. Os moldes de RNA foram removidos por hidrólise alcalina com 15µl de 1N NaOH a 70°C por 10 minutos, em seguida as reações foram neutralizadas com 15µl de 1N HCl, e 20µl de 3M Acetato de Sódio pH 5,2 foi adicionado a mistura. Os

nucleotídeos não incorporados foram eliminados das amostras através do protocolo *SNAP Column* (Invitrogen). As amostras de cDNA purificado foram incubadas a -20°C por cerca de 16 horas após adição de 10µl de 3M Acetato de Sódio pH5,2; 2µl de 20mg/ml de glicogênio; e 300µl de 100% etanol gelado. Os produtos precipitados foram centrifugados a 14.000×g por 20 minutos a 4°C, lavados com 250µl de etanol 75% gelado (14.000×g por 2 minutos a 4°C), e secos por 10 minutos a temperatura ambiente. Quinze microlitros de H₂O foram utilizados para ressuspender as amostras onde foram também adicionados 15µl de Cy5 (*Cy5TM Dye Vials* – Amershan Biosciences) reconstituído com 0,1M bicarbonato de sódio pH 9. Após 2 horas à temperatura ambiente e protegidas da luz, as reações de acoplamento foram finalizadas por adição de 15µl 4M hidroxilamina e incubação a 15min à temperatura ambiente. Os alvos marcados e purificados foram obtidos após repetição do protocolo das *SNAP Column* (Invitrogen), sendo esses mantidos a -20°C e protegidos da luz até o uso.

As reações utilizando o *SuperScriptTM Indirect cDNA Labeling System* (Invitrogen) foram feitas com oligo-dT mais hexâmeros randômicos (6 bases) e toda população de RNA imunoprecipitada em conjunto com as proteínas ligadoras de RNA (ver item 3.8.2) para síntese de cDNA, seguindo estritamente o protocolo do fabricante, usando-se 800 unidades da transcriptase reversa (*SuperScriptTM III RT*) a 46°C, durante 3 horas. As etapas de hidrólise alcalina, purificação do cDNA com aminoallyl, e purificação do cDNA acoplado com *AlexaFluor 647* (Molecular Probes), foram idênticas as descritas acima para as reações feitas com *CyScribe Post-Labeling Kit Protocol* (Amershan Biosciences). Entretanto, as reações de acoplamento foram feitas com adição de 5µl *2X Coupling Buffer* as amostras de cDNA purificadas, adição de 5µl *AlexaFluor 647* (Molecular Probes) ressuspendido

com 2µl de DMSO e 3µl H₂O, e incubação por 2 horas a temperatura ambiente e protegidas da luz.

Apenas nas reações realizadas com o *CyScribe Post-Labeling Kit Protocol* (Amersham Biosciences) foram incluídos 2,5µl da mistura de mRNAs controle contidos no *Lucidea Microarray ScoreCard* (Amersham Bioscience). Esses mRNAs são complementares ao grupo de 32 seqüências de levedura depositadas no *microarray* de cDNA (detalhes no item 3.5.1).

3.5.3 Hibridizações em lâminas de *microarrays* de cDNA

Cada lâmina foi hibridizada com no mínimo 12pmol de cDNA marcado com Cy3 e/ou Cy5. Os alvos marcados foram ressuspensos em 250µl de tampão de hidrização, composto por 50% formamida, 25% de H₂O, e 25% de *Microarray Hybridization Buffer Version 2* (Amersham Biosciences); desnaturados por 2 minutos a 92°C; e centrifugados a 14.000×g por 5 minutos.

Todos os passos de processamento das lâminas (bloqueio, hibridização, lavagens) foram realizados no *Automated Slide Processor* (Amersham Biosciences). As lâminas foram incubadas por 16 horas a 42°C e posteriormente lavadas a temperatura ambiente com 1xSSC/0,2%SDS por 5 minutos; com 0,1xSSC/0,2%SDS por 5 minutos; e com 0,1XSSC por 3 minutos. Finalmente, lâminas prontas para obtenção das imagens por excitação com laser foram obtidas após um jato de isopropanol, e um jato de ar a 42°C para secá-las.

3.5.4 Desenho dos experimentos com *microarrays* de cDNA

Marcadores de prognóstico de tumores de próstata

Vinte pares de amostras tumorais e suas respectivas regiões normais adjacentes marcadas respectivamente com Cy3-dCTP e Cy5-dCTP, foram hibridizadas em 20 lâminas. Um novo conjunto de hibridizações com as amostras tumorais e normais marcadas com fluoróforos invertidos (*dye-swap*) foi realizado para facilitar a correção de diferenças de incorporação e de emissão características de cada molécula fluorescente utilizada. As 7 amostras tumorais não pareadas foram marcadas duas vezes, uma com Cy3 outra com Cy5, e hibridizadas em uma mesma lâmina.

Todas a hibridizações, assim como os procedimentos de lavagem foram realizados automaticamente no *Automated Slide Processor* (Amershan Biosciences), como descrito (ver item 3.5.3).

Transcritos regulados pelo tratamento com andrógeno

Cada amostra proveniente de células LNCaP tratadas com andrógeno ou controle apenas com etanol foram marcadas com Cy5-dCTP, e hibridizadas duas vezes contras as lâminas de *microarray* de cDNA. Uma referência universal composta por quantidades equimolares de RNA proveniente de células LNCaP, DU145 e PC3, foi marcada com Cy3-dCTP, e hibridizada em conjunto com cada amostra experimental, totalizando 26 lâminas.

Todas a hibridizações, assim como os procedimentos de lavagem foram realizados automaticamente no *Automated Slide Processor* (Amershan Biosciences), como descrito (ver item 3.5.3).

Transcritos expressos em tecido humanos e de camundongo

Os RNAs extraídos de cada tecido (rim, próstata e fígado) humano e de camundongos provenientes de 5 espécimes diferentes foram unidos em amostras únicas para a síntese de cDNA e posterior marcação com Cy5 (ver item 3.5.2).

Os alvos marcados de cada tecido (humano e camundongo) foram divididos em três alíquotas iguais para serem hibridizados em três diferentes lâminas de *microarray* de cDNA. Ao todo, 18 experimentos foram realizados, sendo 9 com material proveniente de tecido humano e 9 de tecido de camundongo.

Todas a hibridizações, assim como os procedimentos de lavagem foram realizados automaticamente no *Automated Slide Processor* (Amershan Biosciences), como descrito (ver item 3.5.3).

Transcritos co-imunoprecipitados com proteínas ligadoras de RNA

As imunoprecipitações com cada proteína ligadora de RNA (RBP) foram realizadas em duplicata com quantidades similares de RBP ou da proteína controle GFP, normalizadas pela quantidade total de proteínas presentes nos extratos de 293T (ver item 3.9.5), e todo o RNA obtido foi transformado em cDNA marcado com *Alexa Fluor 647* (Molecular Probes). Cada amostra foi hibridizada sem referência em uma lâmina de *microarray* de cDNA, gerando um conjunto de 10 experimentos.

Todas a hibridizações, assim como os procedimentos de lavagem foram realizados automaticamente no *Automated Slide Processor* (Amershan Biosciences), assim como descrito (ver item 3.5.3).

3.5.5 Obtenção de imagens e valores de intensidade comparáveis

Após os procedimentos de hibridização e lavagens, as lâminas de *microarray* foram excitadas com laser nos devidos comprimentos de onda utilizando o *Gen III Confocal Scanner* (Molecular Dynamics) ou *GenePix 4000B* (Axon Instruments), para obtenção das imagens relativas a cada experimento. Os seguintes parâmetros foram utilizados como padrão nesse procedimento: resolução de 10µm por pixel; comprimento de onda de excitação de 532nm (para alvos marcados com Cy3) ou 633nm (para alvos marcados com Cy5 ou *AlexaFluor 647*); e tensão PMT de 350 volts (para alvos marcados com Cy3) e 600 volts (para alvos marcados com Cy5 ou *AlexaFluor 647*).

Os valores de intensidade de cada um dos canais (Cy3, e Cy5 ou *AlexaFluor 647*) para cada sonda depositada foram obtidos a partir das imagens geradas, utilizando o programa *ArrayVision* (Imaging Research Inc.), com a abordagem de *Artifact-Removed* (ARM). O valor de ARM representa a média de todos os valores de pixels restantes provenientes daquela sonda após a retirada dos valores que excedem quatro desvios absolutos da mediana (MADs). Esse tratamento minimiza artefatos provenientes, por exemplo, de partículas de poeira. Os valores brutos de intensidade foram obtidos após subtração dos valores de hibridização inespecífica na lâmina (*background*) nos arredores da sonda (distância de 2 pixels com 3 pixels de largura) dos valores de ARM calculados.

Quando possível, os arquivos de saída com os valores brutos de intensidade foram importados para o programa do *Lucidea Microarray ScoreCard* (Amersham Bioscience), onde os valores de intensidade dos genes controles contidos no *microarray* foram usados na avaliação da qualidade das hibridizações.

Marcadores de prognóstico de tumores de próstata

Os valores brutos de intensidade obtidos de experimentos com boa qualidade para determinação de marcadores de prognóstico de tumores de próstata foram normalizados pelo algoritmo *LOWESS* (*locally weighted linear regression algorithm*), disponível no pacote R (<http://www.r-project.org/>), para redução dos erros sistemáticos nos valores de intensidade gerados pela diferença de incorporação e emissão de fluorescência de Cy3 e Cy5 (Quackenbush, 2002). Esses valores normalizados passaram depois por um tratamento estatístico para eliminação de dados artefatuais, ou pouco reproduzíveis, no programa *ArrayStat* (Imaging Research Inc.). Esse programa constrói um histograma de frequências a partir da média de intensidade de cada sonda, além de gerar um *scatterplot* do desvio padrão dos dados em relação à mediana, revelando que o erro dos dados varia de acordo com a faixa de intensidade de sinal dos genes. Isso cria um modelo estatístico em cada experimento para eliminação de valores com desvios acima do desvio-padrão observado em cada faixa de intensidade, através das curvas ajustadas de erro (Nadon *et al.*, 2001). Com esses tratamentos finalizados, foram obtidos dados de intensidade de expressão de alta qualidade para as análises de expressão diferencial.

Transcritos regulados pelo tratamento com andrógeno

Os valores brutos provenientes de cada sonda presente no *microarray* foram comparados ao valor médio de intensidade mais três desvios-padrão dos controles negativos do *Lucidea Microarray Scorecard* (G1p, um gene de planta) presentes em cada *sub-array*, para determinação do limite mínimo de detecção de cada hibridização. Os transcritos representados pelas sondas que possuíam intensidade

acima deste limite mínimo foram considerados acessos naquele lado da lâmina. Cada um dos lados foi considerado individualmente, e a média dos valores de intensidade, retirando-se 20% das sondas com maior e 20% das sondas com menor intensidade de sinal, de cada réplica foi utilizada como fator de normalização, para tornar os experimentos comparáveis entre si.

Transcritos expressos em tecidos humanos e de camundongos

Os valores brutos provenientes de cada sonda gerados pelos experimentos de hibridização com amostras humanas e de camundongo foram normalizados pela intensidade dos genes *housekeeping* presente nas diferentes lâminas (ver item 3.5.1). A média dos valores de intensidade, retirando-se 20% dessas sondas com maior e 20% das sondas com menor intensidade de sinal, de cada réplica foi utilizada como fator de normalização para tornar os experimentos comparáveis.

Os transcritos considerados como expressos em cada tecido passaram por um critério bastante conservador, visto que a hibridização de mensagens de camundongo em uma plataforma com sondas humanas poderia ser bastante criticada. Os diferentes transcritos foram considerados acesos em um dos lados da lâmina se sua intensidade estivesse igual ou acima da intensidade do controle negativo do *Lucidea Microarray Scorecard* (G1p, um gene de planta) presente no mesmo *sub-array*, acrescido de três desvios-padrão da média dos valores em toda a lâmina. Cada transcrito só foi considerado aceso em uma lâmina se os valores de intensidade das sondas de ambos os lados satisfizessem o critério anterior. E finalmente, um transcrito só foi considerado como expresso no tecido, se tivesse sido considerado aceso em, no mínimo, duas das três lâminas hibridizadas.

Transcritos co-imunoprecipitados com proteínas ligadoras de RNA

Os valores brutos provenientes de cada sonda gerados pelo experimento de imunoprecipitação com proteínas ligadoras de RNA (RBPs) foram comparados com a intensidade média dos 12 controles negativos do *Lucidea Microarray Scorecard* (G1p, um gene de planta) presente em cada *sub-array* da lâminas. Foram considerados como acessos as mensagens cujas intensidades excederam o valor da média dos controles negativos mais 3 desvios-padrão, em no mínimo 3 dos 4 valores gerados (dois de cada uma das lâminas, réplicas técnicas).

Contudo, só foram considerados como especificamente imunoprecipitados os transcritos que estavam acessos nas amostras imunoprecipitadas com as RBPs (PABP, EIF4E, hTRA2 e HUD), cujos valores de intensidade ainda eram no mínimo 3 desvios-padrão acima de seu valor médio observado nas amostras imunoprecipitadas com GFP. Esse último filtro minimizou a presença de resultados indicativos de RNAs imunoprecipitados inespecificamente ligados a proteína GFP, que não tem domínios de ligação a RNAs conhecidos, ou a região comum de proteína A presente em todos produtos expressos pelo sistema do vetor pKALLU (ver item 3.9.4).

3.6 Hibridizações em *Northern Blots*

A expressão em diferentes tecidos, o tamanho, e a orientação transcricional de algumas das novas mensagens identificadas foi investigada através da hibridização de sondas radioativas em *Northern Blots*. Para esses procedimentos foram utilizadas tanto a membrana de nitrocelulose comercial *Human Multiple Tissue Northern (MTN) Blot II* (Clontech Laboratories Inc.), onde encontram-se transferidos 2µg de RNA poli-(A)+ de alta qualidade, isolados de 9 tecidos humanos diferentes:

baço, timo, próstata, testículo, ovário, intestino delgado, colon e leucócitos sanguíneos periféricos, quanto membranas construídas no laboratório.

Cerca de 4µg de RNA poli-(A)+ isolados de 3 linhagens celulares diferentes de tumores de próstata, PC3, DU145 e LNCaP (ver item 3.3.1 para detalhes da extração) foram utilizadas na construção de membranas de *Northern Blot*. Os RNAs foram separados por eletroforese em gel desnaturante com 1,5% de agarose; 20mM ácido bórico, pH 8,3; 0,25mM de EDTA; e 6,5% de formaldeído (Merck). Antes de aplicadas no gel, as amostras foram desnaturadas a 60°C por 10 minutos em uma solução contendo formamida, formaldeído e tampão de corrida (20% glicerol; 20mM ácido bórico, pH 8,3; 0,25mM EDTA; 50% formamida; 16% formaldeído; e azul de bromofenol). Após 5 minutos no gelo, as amostras foram aplicadas no gel e submetidas à eletroforese a 70 volts em tampão de corrida durante cerca de 4 horas.

O gel foi lavado em 4x SSC (0,6 M NaCl, 60 mM citrato de sódio, pH 6,45) e 10x SSC (1,5 M NaCl, 0,15 M citrato de sódio, pH 6,45) durante 20 minutos para eliminação do excesso de formaldeído. As amostras foram transferidas por capilaridade a uma membrana de nylon (Hybond-N) com 10x SSC durante 16 horas, e os RNA foram fixados na membrana por radiação ultravioleta utilizando o *UV Crosslinker* (Biorad).

3.6.1 Detecção de mensagens utilizando sondas de fita dupla

A marcação das sondas com [$\alpha^{32}\text{P}$]-dCTP seguiu o protocolo do *Rediprime II* (Amershan Pharmacia), usando como molde fragmentos gerados a partir de RACE-PCR (descrito no item 3.2.6), e um produto com cerca de 500bp proveniente da região codificante do gene da β -actina (*ACTB*). Os procedimentos de hibridização foram realizados de acordo com as condições especificadas no manual fornecido

pelo fabricante das membranas *Multiple Tissue Northern (MTN) Blot II* (Clontech Laboratories Inc.). A detecção foi realizada no *Storm Phosphoimager* (Molecular Dynamics) após dois dias de exposição das membranas no cassete *Storm Phosphor Screen* (Molecular Dynamics).

3.6.2 Detecção de mensagens utilizando sondas de fita simples

Duas abordagens distintas foram utilizadas para gerar sondas radioativas que fossem específicas para definir a orientação transcricional de mensagens intrônicas em hibridizações com membranas de Northern Blot. Uma foi baseada no uso de oligonucleotídeos marcados com ^{32}P em suas regiões 5'. A marcação dessas foi realizada através de reações de fosforilação com $[\gamma^{32}\text{P}]\text{-ATP}$ e uma polinucleotídeo quinase (T4 PNK – Invitrogen). A hibridização dos oligonucleotídeos marcados foi realizada em uma membrana comercial de nylon contendo $2\mu\text{g}$ de RNA poli-(A)+ isolado de vários tecidos humanos (*Multiple Tissue Northern blot II* - Clontech Inc.), seguindo as especificações recomendadas pelo fabricante.

Na outra abordagem a síntese das sondas radioativas foi realizada através de transcrição *in vitro* usando a T7 RNA polimerase na presença de $[\alpha^{32}\text{P}]\text{-dCTP}$ (Riboprobe - Promega). Produtos gerados pelos experimentos de RACE-PCR (descrito no item 3.2.6) foram utilizados como molde após sua linearização orientada por reações de PCR com um oligonucleotídeo de sequência igual à região anterior ao sítio promotor da RNA polimerase (T7 ou SP6) e outro complementar a região do inserto com a orientação desejada. A figura 12 ilustra esse procedimento. A geração das sondas de RNA seguiu as especificações do fabricante. As hibridizações ocorreram em membranas de nylon construídas com $4\mu\text{g}$ de poli-(A)+ RNA isoladas

de 3 linhagens celulares diferentes de tumores de próstata, PC3, DU145 e LNCaP (ver item 3.6).

Em ambas abordagens, a detecção foi realizada no *Storm Phosphor Screen* (Molecular Dynamics) após dois dias de exposição da membrana marcada no cassete apropriado.

Figura 12 – **Transcrição *in vitro* para geração de sondas fita-específicas.** (A) Molde linearizado por PCR usando os oligonucleotídeos **GSP_F** e **Lud_R** é utilizado para gerar sondas complementares à um RNA com orientação antisense usando-se SP6 RNA polimerase. (B) Molde linearizado por PCR usando os oligonucleotídeos **Lud_F** e **GSP_F** é utilizado para gerar sondas complementares à um RNA com orientação antisense usando-se a T7 RNA polimerase. As denominações senso e antisense se referem ao mRNA codificante de proteína. Em azul consenso do promotor da T7 RNA polimerase, em vermelho da SP6, e em verde as regiões do transcrito alvo. As setas cortadas indicam os oligonucleotídeos de linearização.

3.7 Análise de expressão gênica por PCR

3.7.1 Determinação da orientação de mensagens utilizando cDNA transcrito reversamente com oligonucleotídeos iniciadores específicos

Transcrições reversas com orientação específica (*Orientation-Specific Reverse Transcription*) seguidas de amplificação por PCR (OSRT-PCR) foram também utilizadas para determinação da natureza transcricional de mensagens intrônicas. Utilizando 100ng de mRNA ou 2-4 μ g de RNA total, 200U de transcriptase reversa *SuperscriptII* (Invitrogen) mais 0,9 μ M de oligonucleotídeos iniciadores complementares a fita de RNA antisenso ou a fita senso de cada região intrônica, as reações de transcrição reversa foram mantidas a 50°C por 60 minutos (Shendure & Church, 2002; Cawley *et al.*, 2004). Controles para a ausência de *RNA self-priming* e de contaminação com DNA genômico, foram realizados com reações sem adição de oligonucleotídeos iniciadores e de transcriptase reversa, respectivamente.

Após a transcrição reversa, quantidades de cDNA equivalentes a 10ng de RNA reversamente transcrito (utilizando uma estimativa de 2% de RNA reversamente transcrito na população de RNA total, quando necessário) foram utilizadas como molde para amplificações das mensagens intrônicas senso e antisenso, por 37-40 ciclos no *GenAmp 5700 Sequence Detector* (Applied Biosystems) utilizando o *SYBR[®] Green PCR Master Mix Kit* (Applied Biosystems), como meio de reação. As amplificações ocorreram com 0,8 μ M de oligonucleotídeos desenhados internamente aos usados durante a transcrição reversa, de modo a garantir uma maior especificidade nas reações. Finalmente, os produtos obtidos foram separados por eletroforese em gel com 3% de agarose. A figura 13 apresenta um diagrama esquemático do processo.

Figura 13 – **Transcrições reversas com orientação específica seguidas de amplificação por PCR.** A figura apresenta um esquema da metodologia utilizando o locus de *RASSF1* como exemplo (**superior**). Um cDNA com orientação definida é sintetizado por reações de transcrição reversa com oligonucleotídeos (RT-F e RT-R) complementares a RNAs antisense, ou senso (pré-mRNA no exemplo). Gel de 3% de agarose para visualização dos produtos amplificados por PCR (**inferior**) utilizando oligonucleotídeos internos para a sequência de sobreposição das mensagens (**direita**) ou para a sequência que só existe com orientação senso (**esquerda**), realizado aqui como controle extra. Controles para ausência de contaminação com DNA genômico (- RT) e de auto-anelamento do RNA (- oligo) durante a transcrição reversa são fundamentais para interpretação dos resultados. Figura adaptada de Reis, E. M. *et al.* (2005). *As Antisense Gets Intronic. Omics*, **9**(1):2-12.

3.7.2 Avaliação da abundância relativa de mensagens utilizando PCR quantitativo

Algumas diferenças de expressão observadas nas análises por *microarrays* de cDNA foram confirmadas através de PCR quantitativo, utilizando a técnica de Real Time PCR com o aparelho no *GenAmp 5700 Sequence Detector* (Applied Biosystems) e o *SYBR[®] Green PCR Master Mix Kit* (Applied Biosystems). Precedente às validações, foram necessárias otimizações das condições para garantir eficiências comparáveis de amplificação. A concentração ideal dos pares de oligonucleotídeos iniciadores foi determinada realizando reações idênticas de amplificação, porém com quantidades crescentes desses, de 200nM-800nM, através da identificação da situação na qual não ocorre a formação de dímeros e onde a amplificação é mais robusta. Outro fator importante atestado foi se as eficiências de amplificação dos transcritos alvos e do transcrito referência, em função da concentração de RNA utilizado, são aproximadamente iguais.

Avaliação da abundância de mensagens utilizando PCR quantitativo em tempo real

Amostras contendo 5µg de RNA total ou 500ng de mRNA, extraídas das condições a serem comparadas, foram reversamente transcritas em cDNA utilizando o protocolo padrão da *SuperScript[™] II First-Strand Synthesis System for RT-PCR* (Invitrogen), ou a metodologia para obter transcrições reversa com orientação específica (*Orientation-Specific Reverse Transcription*) descrita anteriormente (ver item 3.7.1). A quantificação foi baseada na observação da abundância dos produtos no ciclo de PCR onde as reações estão em fase exponencial de amplificação. O

acúmulo dos produtos de PCR nessa fase foi detectado em tempo real, através da ligação do *SYBR[®] Green*, que é uma molécula cuja emissão de fluorescência aumenta várias vezes quando ligada a DNA dupla-fita. Através da definição arbitrária da região exponencial de amplificação, foram determinados os *Cycle Threshold* (C_T) de cada transcrito. O valor relativo da expressão foi dado pela diferença (ΔC_T) entre o C_T do gene alvo e o C_T do gene de referência, tubulina (*TUBA6*) (um gene estrutural que normalmente não apresenta variação de expressão entre as amostras estudadas). Os resultados finais foram expressos como *fold-change*, usando a formula $2^{-\Delta C_T}$ (Pfaffl, 2001). A significância estatística das diferenças de expressão observadas (*p-values*) foi determinada usando *Student's t-test*, assumindo uma distribuição bicaudal e um modelo de variância heterocedástico.

3.8 Experimentos de Imunoprecipitação

3.8.1 Ensaio de imunoprecipitação da cromatina

Ensaio de imunoprecipitação da cromatina (*Chromatin Immunoprecipitation* - ChIP) foram desenhados para detectar a ligação do complexo do receptor de andrógeno à regiões promotoras dos transcritos intrônicos. Nessas regiões foi predita por análise informática a existência de um ARE, elemento de resposta a andrógeno (ver item 3.1.4). As condições e os reagentes para realização dos experimentos de ChIP foram do *Chromatin Immunoprecipitation Assay Kit* (Upstate Biotechnology Inc.).

Brevemente, células LNCaP tratadas com andrógeno ou as células controle foram fixadas com 1 % de formaldeído por 10 minutos a temperatura ambiente para ligação covalente das proteínas ao DNA, e foram lisadas com tampão contendo SDS. Esses extratos celulares foram sonificados para quebra mecânica do DNA em

tamanhos de 200-1000bp usando o *Sonic Dismembrator Model 500* (Fischer Scientific). As imunoprecipitações foram realizadas com 3µg do anticorpo policlonal contra receptor de andrógeno (anti-AR) denominado *H-280* (Santa-Cruz) ou 1µg de IgG de camundongo (controle negativo - Upstate) usando 100µl de lisados contendo aproximadamente 2×10^6 células. Após reversão das ligações entre DNA e proteínas, a cromatina imunoprecipitada foi tratada com RNase A e proteinase K, e os fragmentos de DNA foram purificados utilizando colunas de purificação de produtos de PCR *GFX* (Amersham Biosciences).

A detecção das regiões promotoras identificadas *in silico* foi realizada por reações de PCR contendo os oligonucleotídeos iniciadores para as regiões do ARE predito para o transcrito intrônico antisense do locus *MYO5A* (TTTGCTGGATAAGGATTCCCA, AGTTCCTTAACTTTAGCTTAGAAAAAGGA); para o AREIII do *PSA* (controle positivo) (GCTCAGCCTTTGTCTCTGATGA, TGCAAGATGATATCTCTCTCAGATCC); e para uma região arbitrária do transcrito de *TUBA6* (controle negativo) (GACTACAGGTGTGCGCCATCAT, TGCCGTGTTCCAGGCAGTAG).

3.8.2 Imunoprecipitação de proteínas ligadoras de RNA (RBPs)

Volumes apropriados de cada lisado de 293T transfectado com pKALLU contendo as RBPs, ou contendo GFP (controle negativo), normalizado pela quantidade total de proteína (ver item 3.9.5) foram centrifugados a 13.000×g por 10 minutos a 4°C e os sobrenadantes foram transferidos para tubos novos contendo 10 volumes de NT2 buffer (50mM Tris; pH 7,4; 150 mM NaCl; 1 mM MgCl₂; 0,05% NP40) suplementado com 25mM EDTA, 2mM DTT, 100U/ml RNaseOUT, e 0,4 mM *Vanadyl Ribonucleoside Complex*, VRC. Para cada lisado foi adicionado cerca de

200µl de *Sepharose 6B beads* (Sigma), e a mistura foi incubada por 30 minutos a 4°C, com agitação constante. Essa primeira incubação visou minimizar a quantidade de proteínas ou RNAs que podem ser inespecificamente precipitados através de interação direta com os *Sepharose beads*. Ainda para minimizar interações inespecíficas, as alíquotas de *IgG Sepharose 6 fast flow beads* (GE Health Care) que foram usados para imunoprecipitar as RBPs através da interação da IgG com a proteína A, foram também incubados por 30 minutos com agitação constante a 4°C, com o mesmo tampão descrito acima, porém além dos suplementos já mencionados foi adicionado 5% BSA.

As amostras contendo os *Sepharose beads* com os extratos das células que expressaram as RBP, e os *IgG Sepharose 6 fast flow beads* em solução com BSA foram então centrifugadas a 1.000×g por 5 minutos a 4°C. Os *IgG Sepharose 6 fast flow beads*, após retirada do sobrenadante por sucção, foram lavados 4 vezes com o NT2 suplementado. O sobrenadante com as RBPs foi transferido para os tubos contendo cerca de 200µl de *IgG Sepharose 6 fast flow beads* lavados. Essa nova mistura foi incubada por no mínimo 3 horas, a 4°C com agitação constante.

Finalizada essa última incubação, os *IgG beads* com as RBPs e seus RNAs alvo ligados foram centrifugados novamente a 1.000×g por 5 minutos a 4°C, e foram lavados por 4 vezes com cerca de 5 volumes (referente aos *beads*) do tampão NT2. Após esse procedimento, os RNA co-imunoprecipitados com as RBPs puderam ser extraídos (ver item 3.3.3) para os experimentos de expressão gênica por *microarrays* de cDNA.

3.9 Sistemas de expressão episomal

3.9.1 Construção do sistema de expressão de Rasl11a/b em procarioto

A estratégia escolhida para obtenção de proteínas das novas GTPases monoméricas da subfamília Rasl11 foi a expressão em bactérias através de estímulo de promotores ativados por IPTG (*isopropyl-beta-D-thiogalactopyranoside*). No vetor de expressão *pET21* (Novagen) foram clonadas as regiões codificantes dos genes de interesse *RASL11A* e *RASL11B*. A metodologia de clonagem foi baseada nas seqüências únicas de corte por enzimas de restrição presentes no sítio de clonagem do referido vetor, que não possuísem correspondência nas regiões codificantes de *RASL11A/B*. Assim, foram desenhados pares de oligonucleotídeos iniciadores complementares ao começo e final das fases abertas de leitura dos genes. Esses oligos foram modificados para apresentarem tanto as seqüências de reconhecimento pela enzima de restrição *NdeI* para a região 5' inicial, quanto de *XhoI*, mas com substituição de uma base para cancelamento do códon de parada de tradução da região 3' terminal. Essa última característica foi acrescentada aos insertos para que a região codificante ficasse em fase com uma repetição de nucleotídeos que codificam para histidina. A presença dessa porção final leva a expressão da proteína dos membros da subfamília Rasl11 em fusão com uma cauda de seis histidinas. A presença dessa cauda em proteínas recombinantes permite sua purificação através de cromatografia de troca iônica usando colunas de níquel, e sua detecção em *Western Blots* ou *Dot Blots* com anticorpos gerados contra um peptídeo contendo múltiplas histidinas.

Os insertos amplificados da forma descrita acima, foram clonados primeiramente em *pGEM-T Easy* (Promega), tanto para a confirmação da identidade das seqüências obtidas, por sequenciamento dideoxi (Megabace – Amershan

Pharmacia), quanto para facilitar o reconhecimento dos sítios de digestão pelas enzimas a serem utilizadas durante a transferência para o vetor de expressão *pET21*.

Confirmadas as identidades, os insertos foram transferidos para o vetor de expressão *pET21*. Quantidades referentes a 200µg de produtos de mini-preparações de DNA plasmidial (*Plasmid Mini Kit* Qiagen) de *pET21* vazio (sem inserto) e 100µg de pGEM-T-Ras11a/b foram utilizados em reações simultâneas de digestão com as enzimas *NdeI* e *XhoI* (New England Biolabs) seguindo as recomendações do fabricante quanto ao número de unidades utilizadas, tampão de reação, temperatura e tempo de incubação. A seguir, 50µg de vetor digerido foram adicionados a uma quantidade dos insertos (em mol) relativa a três vezes a quantidade do vetor digerido, para uma reação de ligação com *T4 DNA Ligase* (Promega) por 12-16 horas a 6°C. As ligações após dialisadas para a diminuição da concentração de sal, foram utilizadas para transformar por eletroporação células de *E. coli* da linhagem DH10B, destinadas à propagação das construções pET21-Ras11a/b. Nessa etapa os insertos foram mais uma vez seqüenciados para assegurar a expressão de proteínas com total ausência de mutações.

As colônias transformantes foram crescidas em meio LB líquido com ampicilina para novas preparações de DNA plasmidial (*Plasmid Mini Kit* - Qiagen), a serem utilizadas nos experimentos de indução da expressão em células de *E. coli* da linhagem ORIGAMI. A presença de um gene para resistência a ampicilina no vetor *pET21*, assim como no *pGEM-T*, e o acréscimo de ampicilina no meio foram utilizados para selecionar as bactérias transformantes em todas as etapas.

3.9.2 Indução da expressão de Rasl11b em bactéria

A construção pET21-Rasl11b, e em paralelo o vetor pET21 vazio (sem inserto), foi utilizada para transformação de células de *E. coli* da linhagem para expressão ORIGAMI. A cada experimento de indução de expressão uma fração da preparação de DNA plasmidial das construções (descrito no item 3.9.1) foi utilizada para nova transformação, pois após um breve período de armazenamento em glicerol, os clones de bactérias ORIGAMI transformadas perdem a capacidade de expressão heteróloga.

Uma colônia de bactérias transformantes com pET21-Rasl11b foi pré-inoculada em 5ml de LB contendo ampicilina por aproximadamente 1 hora a 37°C, até atingirem densidades ópticas em 600nm (Abs_{600}) de 0,5-1,0. Estando as bactérias crescendo em fase exponencial, esses 5ml foram transferidos para inóculos contendo 200ml de LB com ampicilina, para crescimento por cerca de 3 horas a 37°C, até novamente atingirem Abs_{600} de 0,5-1,0. Em seguida a expressão das proteínas foi induzida pelo IPTG adicionado ao meio em concentração final de 1mM, e incubando as células a 37°C por até 6 horas. Através de centrifugação por 15 minutos a 5.000×g (4°C) as bactérias foram isoladas do meio de cultura para congelamento a -80°C até serem lisadas para verificação da expressão. As Abs_{600} finais foram medidas para normalização da quantidade de tampão de lise utilizado, visando garantir quantidades de proteína total por ml de solução relativamente semelhantes.

Os aglomerados de bactérias induzidas foram ressuspensos em um tampão de lise contendo 20mM de Tris-Cl pH 8,0 e 1mM PMSF (*Phenylmethylsulfonyl fluoride*, um potente inibidor de proteases) no volume apropriado, e lisados em uma *French Press* através de três passagens a pressão de 2.000 psi. Uma fração

referente ao lisado total foi armazenada e o restante foi centrifugado a 20.000×g por 30 minutos (4°C) para separação das frações solúvel (sobrenadante) e insolúvel (precipitado). Para visualização das proteínas expressas, quantidades normalizadas pelas Abs₆₀₀ finais, foram separadas por SDS-PAGE com 12% de poliacrilamida, ao lado de um padrão de peso molecular de proteínas *Prestained Protein Marker* (Fermentas), corado posteriormente por *Coomassie Blue*.

3.9.3 Detecção de Rasl11b expresso em bactéria com anticorpo anti-histidinas.

Para certificação da natureza da proteína expressa, foi utilizada a cauda de seis histidinas fusionada à região carboxi-terminal de Rasl11b para realização de um *Dot Blot*. Para isso, após as etapas descritas acima (ver item 3.1.6), uma diluição seriada do lisado total, das frações solúvel e insolúvel, referentes à construção pET21-Rasl11b, pET21 vazio (controle negativo), e 6xHis-Ca⁺²-ATPase (controle positivo), foram aplicadas em uma membrana de nitrocelulose utilizando o *Filtration Manifold System* (Gibco).

As membranas foram bloqueadas em 10ml de Blotto (5% de leite em pó desnatado; 0,03% de azida sódica) durante 3-6 horas. Em seguida, foram incubadas com 10µl de anticorpo primário anti-histidinas (*Anti-His Antibody* - Amershan Biosciences) em nova solução de Blotto por 16 horas, sendo em seguida lavada com PBS Tween 0,1%, uma vez por 15 minutos e duas vezes por 5 minutos. Usando 10ml de Blotto sem azida, a membrana foi então incubada com 10µl de anticorpo secundário (*Anti-mouse IgG, Horseradish Peroxidase* - Amershan Biosciences), por 1 hora, sendo as lavagens descritas repetidas em seguida. A revelação foi realizada

por quimioluminescência (*ECL Western Blotting Detection Reagents* - Amershan Biosciences), seguindo o protocolo do fabricante.

3.9.4 Sistema de expressão de proteínas ligadoras de RNA em células mamífero

Para investigar se algumas das proteínas ligadoras de RNA (*RNA Binding Proteins* – RBPs) envolvidas em mecanismos de regulação pós-transcricional ligam-se aos longos RNAs intrônicos, foi utilizada uma abordagem de imunoprecipitação, similar a já utilizada em estudos anteriores (Penalva & Keene, 2004; Penalva *et al.*, 2004), para posterior análise por *microarrays* de cDNA.

As RBPs de interesse foram expressas em um sistema composto pelo vetor pKALLU, baseado no vetor comercial *pcDNA3* (Invitrogen), previamente construído (Penalva *et al.*, 2004), e pela linhagem celular tumoral de tecido renal 293T, obtida da ATCC (*American Type Tissue Collection*). A escolha dessa linhagem foi baseada em sua velocidade de crescimento e em sua eficiência de transfecção, cerca de 90%. O vetor utilizado apresenta como principais características uma região codificante para proteína A, proveniente do *pEZZ* (Amershan Biosciences) para ser expressa em fusão com a RBP de interesse; um promotor de CMV para expressão constitutiva; e genes de resistência a antibióticos.

O estudo foi realizado com cinco construções em pKALLU: um controle negativo com uma proteína reconhecidamente não ligadora de RNA, GFP (Green Fluorescent Protein); e quatro RBPs, PABP (*Poly(A)-Binding Protein*), hTRA2 (*Human Transformer 2*), EIF4E (*Eukaryotic Translation Initiation Factor 4E*), e HuD (*HU-Antigen D*).

3.9.5 Detecção das RBPs expressas em 293T com anticorpo contra proteína A.

Quarenta e oito horas após a transfecção com pKALLU contendo as RBPs descritas acima ou GFP (ver item 3.2.3), as células foram coletadas e tratadas de modo a ser possível a extração de RNAs íntegros associados às proteínas ligadoras de RNA em estudo. As células foram lavadas duas vezes com PBS gelado, coletadas usando um *scraper* e centrifugadas a 2.000×g por 5 minutos a 4°C. O precipitado celular foi ressuspensionado e homogeneizado em 2 volumes de tampão de lise (100mM KCl, 25mM EDTA; 5mM MgCl₂; 10mM HEPES, pH 7,0; 0,5% Nonidet® P-40 (NP40); 10% glicerol; 2mM DTT; 0,4 mM *Vanadyl Ribonucleoside Complex*, VRC, da New England Biolabs; 1 tablete/50 ml de *Complete Proteinase Inhibitor* da Roche; e 100U/ml *RNaseOUT* da Invitrogen). Os lisados celulares foram centrifugados a 13.000×g por 10 minutos a 4°C e o sobrenadante foi transferido para um tubo novo. O precipitado celular recebeu novamente o mesmo tratamento descrito, mas dessa vez usando apenas 1 volume de tampão de lise. Ambos sobrenadantes foram centrifugados e armazenados a –80°C em alíquotas de 1ml para a realização das imunoprecipitações (ver item 3.8.2), e uma alíquota menor com cerca de 100µl para quantificação protéica através da metodologia de Bradford (Harlow & Lane, 1998) e para a verificação da expressão das RBP específicas.

A verificação da expressão específica das RBPs foi realizada por *Western Blot*. Um volume de lisado celular referente à cerca de 50µg de proteína total de cada proteína expressa foi separado por SDS-PAGE, 10% de acrilamida. As proteínas foram transferidas para uma membrana de nitrocelulose, e essa membrana foi bloqueada com 5% de leite desnatado em PBS com 0,1% Tween® 20 por cerca de 16 horas. Essa membrana foi incubada com uma diluição em PBS de

1:1000 de anti-proteína A por 2 horas, seguida de uma incubação com o anticorpo secundário anti-IgG de coelho (1:1000) conjugado com peroxidase (*horseadish*) por 1 hora. A detecção foi feita por quimioluminescência (*ECL Western Blotting Detection Reagents* - Amershan Biosciences), seguindo o protocolo do fabricante.

3.9.6 Sistema de expressão de transcritos intrônicos em células de mamífero

Os longos RNAs intrônicos selecionados para posterior caracterização funcional foram clonados no sistema de expressão em células de mamífero *pBudCE4.1* (Invitrogen). Esse vetor foi escolhido para o início desse estudo devido à existência de duas unidades de expressão distintas, uma sob o controle do promotor CMV, e outra do promotor EF-1 α .

A estratégia constou em clonar em uma das unidades o gene que codifica para a proteína fluorescente verde (GFP), que facilitará a realização de ensaios transientes através de seleção das células transfectantes pelo *sorter* do citômetro de fluxo; e na outra os transcritos intrônicos a serem estudados. Devido a baixa eficiência de transfecção da linhagem celular de próstata tumoral, LNCaP, escolhida como modelo para esses estudos funcionais, a co-expressão de GFP é muito útil para rápida seleção das células. Entretanto, como recurso adicional poderá também ser realizada seleção clonal das células onde ocorrer integração do vetor ao seu genoma, utilizando o antibiótico *Zeocin* (Invitrogen), letal tanto para células eucarióticas quanto procarióticas.

Através de procedimentos de PCR, sítios para enzimas de restrição de corte único, *HindIII* e *BamHI*, foram adicionados às regiões 5' e 3', respectivamente, de um fragmento de cDNA contendo toda a porção codificante de proteína de GFP. Da

mesma maneira, sítios de restrição para as enzimas *NotI* e *XhoI* foram adicionados às pontas 5' e 3', respectivamente, de transcritos intrônicos correlacionados com o grau de diferenciação tumoral em próstata cuja sequência completa foi obtida por RACE (Reis *et al.*, 2004). Uma clonagem intermediária dos produtos obtidos em *pGEM-T* (Promega) foi realizada, tanto para confirmação da natureza e integridade das mensagens obtidas através de sequenciamento dideoxi, quanto para facilitar o reconhecimento das enzimas de restrição ao sítio criado.

Os clones confirmados foram crescidos em LB com ampicilina para extração de DNA plasmidial com o *Plasmid Mini Kit* (Qiagen). Através de reações de digestão sequencial (GFP), e de dupla digestão (transcritos intrônicos), dos plasmídeos extraídos, seguidas de procedimentos de purificação com o *Gel Extraction Kit* (Qiagen), foram obtidos os insertos finais para ligação ao vetor de expressão. Em primeiro lugar foi gerada a construção pBud-GFP, através da inserção do fragmento de GFP modificado ao vetor digerido com *HindIII* e *BamHI*, na região localizada bem em frente ao promotor CMV. Além de ser o vetor base para as clonagens dos transcritos a serem estudados, o pBud-GFP será utilizado como um dos controles negativos para determinação se os efeitos desencadeados nas células são específicos para as mensagens intrônicas. Seguindo procedimentos similares foram finalmente geradas as construções, utilizando o pBud-GFP como base, com os transcritos intrônicos sob o controle do promotor EF-1 α .

Todos os sistemas construídos foram amplificados por propagação em bactérias, em quantidades suficientes para uma preparação média de DNA plasmidial com o *Plasmid Midi Kit* (Qiagen). Os produtos finais a serem utilizados para modulação da expressão de mensagens intrônicas estão representados esquematicamente na figura 14, que também resume a estratégia de clonagem.

Figura 14 – **Sistema de expressão de Longos RNAs Intrônicos não-codificantes antisense em células de mamífero.** Esquema do sistema construído para superexpressão em células de mamífero dos ncRNAs intrônicos *asRASSF1* e *asPPHLN1* utilizando o vetor comercial pBudCE4.1 (Invitrogen) como base. Foram geradas as construções pBud-GFP, pBud-GFP-*asRASSF1* e pBud-GFP-*asPPHLN1* a partir de uma estratégia de clonagem utilizando enzimas de restrição (destacadas na caixas vermelhas). Os promotores CMV e EF-1 a, e o gene de resistência ao antibiótico Zeocin (Invitrogen) estão indicados.

3.10 Experimentos para obtenção de anticorpos

3.10.1 Preparo de antígenos de Rasl11b para imunização de coelhos.

Os lisados totais de bactérias transformadas com pET21 Rasl11b, obtidos da mesma maneira relatada anteriormente, foram fracionados por uma centrifugação em menor velocidade, 10.000×g, e por um tempo inferior, 15 minutos, visando minimizar o efeito de “co-purificação” de proteínas bacterianas. Em seguida os corpos de inclusão foram lavados 2 vezes com solução contendo uréia (50mM Tris-Cl pH 8,0 com 1M Uréia) e finalmente ressuspensos em tampão para corpos de inclusão (50mM Tris-Cl pH 8,0 com 100mM NaCl e 1mM EDTA) para quantificação

protéica seguindo a metodologia de Bradford, a partir de uma curva padrão de albumina sérica (Harlow & Lane, 1998). Quantidades suficientes de proteína para as imunizações foram ainda centrifugadas ($20.000\times g$ por 30 minutos) para a obtenção da preparação de proteína-adjuvante. Essa preparação corresponde a uma mistura 1:1 da solução do tampão para corpos de inclusão com adjuvante de Freund completo (Difco), na dose inicial, e incompleto no reforço.

3.10.2 Imunização de coelhos com corpos de inclusão de Rasl11b.

Foram utilizadas duas fêmeas jovens da linhagem *New Zeland*, cuja saúde foi atestada após um período de observação (15 dias) anterior ao início do experimento. Cada fêmea recebeu doses subcutâneas em vários pontos do dorso (4-8 pontos) da preparação de proteína-adjuvante (descrito anteriormente) totalizando cerca de $100\mu g$ de proteína por espécime. Quatro semanas após, foi dado um reforço contendo as mesmas quantidades de proteína. Foram coletadas amostras de sangue, até 20ml, proveniente da artéria central da orelha do animal para a obtenção de: soro pré-imune (antes da primeira inoculação), o soro imune da primeira dose (logo antes do reforço) e o soro imune final, para serem testados.

A obtenção dos soros foi conseguida através de coagulação do sangue coletado a $37^{\circ}C$ por 1 hora com um palito de madeira no centro do tubo. Após completa coagulação ao redor do palito, o tubo foi armazenado a $4^{\circ}C$ por 12 horas para contração do coágulo. O líquido ao redor do coágulo foi reservado, e o coágulo foi centrifugado a $1.500\times g$ a $4^{\circ}C$, por 10 minutos, para obtenção do soro residual. Todo soro extraído foi então centrifugado a $10.000\times g$ a $4^{\circ}C$, por 10 minutos para retirada das células e material insolúvel. Finalmente, os soros foram aliquotados, identificados e armazenados a $-80^{\circ}C$ até seu uso.

3.11 Análise de expressão gênica por *microarrays* de oligonucleotídeos

3.11.1 Plataforma de *microarrays* de oligonucleotídeos

Alguns dos estudos de expressão gênica descritos aqui fizeram parte do projeto temático para “Identificação de marcadores moleculares para diagnóstico e prognóstico em câncer utilizando *microarrays* de DNA”, financiado pela FAPESP (Processo nº 02/13283-6). Com parte desse financiamento foi desenhado pelo grupo um *microarray* de oligonucleotídeos, cujas características se encontram publicadas sob o número de acesso GPL4051, no banco de dados *Gene Expression Omnibus* (GEO – <http://www.ncbi.nlm.nih.gov/geo/>).

Brevemente, sondas de oligonucleotídeos foram desenhadas para 7.715 RNAs totalmente intrônicos e para 5.289 mensagens parcialmente intrônicas (com uma região de sobreposição com exons de *RefSeq*). Esses transcritos foram escolhidos randomicamente a partir de listas de *contigs* de ESTs sem *splicing* geradas pelo projeto de mapeamento em larga escala que estudou o conteúdo transcricional intrônico humano (Nakaya *et al.*, 2006).

Para cada transcrito intrônico foram obtidas sondas de 60 bases que satisfizessem uma série de critérios (Hughes *et al.*, 2001). As sondas deveriam ter menos de oito bases derivadas de regiões repetitivas do genoma, e menos de sete bases de segmentos com baixa complexidade (homopoliméricos); as sondas deveriam ter um conteúdo de GC de 35–55%; e uma temperatura de desnaturação calculada (T_m = *melting temperature*) de 68–76°C. Para minimizar a possibilidade de hibridizações cruzadas, cada sonda de 60 bases foi submetida a uma busca com BLASTn (NCBI) contra um banco de dados específico contendo todas as regiões genômicas humanas onde existe evidência de transcrição por mRNAs ou ESTs. O

melhor *hit* de cada sonda contra este banco é exato, ou seja, tem por definição 100% de identidade ao longo das 60 bases, o que corresponde a um BLAST *bit-score* de 119. Um *bit-score* de 42.1 corresponde a um alinhamento máximo de 21 bases sem *gap*. As sondas cujo segundo melhor *hit* contra esse banco de dados tivessem um *bit-score* mais baixo que 42.1 foram consideradas com baixa probabilidade de reação cruzada, ou seja, seriam sondas específicas, e prosseguiram para o próximo passo de seleção.

Os oligonucleotídeos considerados específicos foram mapeados em seus respectivos transcritos e a sonda mais próxima a porção 3' de cada alvo foi selecionada, sempre assumindo uma orientação antisenso para mensagem intrônica em relação ao transcrito exônico codificante de proteína correspondente. Para cada alvo, uma segunda sonda com orientação oposta foi obtida utilizando o reverso complementar da primeira sequência de 60 bases selecionada, sendo que para os transcritos parcialmente intrônicos essa segunda sonda corresponde ao exon do *RefSeq* onde a mensagem intrônica se sobrepõe.

Para possibilitar a detecção em paralelo dos níveis transcripcionais das mensagens codificantes de proteína de mesmo *loci* dos RNAs intrônicos, foram incluídos 14.393 elementos correspondentes a sondas desenhadas pela própria *Agilent* para a plataforma comercial *Whole Human Genome Oligo Microarray*. Foram incluídas também 2.256 sondas para controles positivos e negativos presentes nessa plataforma comercial. As lâminas desta plataforma *Intronic 37k microarray*, desenhada pelo grupo como descrito acima, foram construídas por encomenda pela *Agilent Technologies*.

3.11.2 Marcação de alvos fluorescentes por amplificação de RNA

As amostras de RNA provenientes de duas réplicas biológicas de células LNCaP tratadas com α -amanitina e de células controle foram marcadas através de amplificação do RNA seguindo as recomendações do *Agilent Low RNA Input Fluorescent Linear Amplification Kit* (Agilent Technologies). Brevemente, cRNAs marcados com Cy5 e Cy3 foram obtidos usando 500ng de RNA total como molde para amplificação de RNAs poli-(A)+ pela T7 RNA polimerase. Cada amostra foi marcada um vez com cada fluoróforo, sendo incluídos os mRNAs sintetizados *in vitro* *Agilent RNA Spike-In Kit* (Agilent Technologies) para serem utilizados como controle.

A marcação utilizando cRNA amplificado por transcrição *in vitro* com a T7 RNA polimerase apresenta uma grande vantagem em relação a abordagem de marcação do cDNA com transcrição reversa. A orientação das sondas geradas é sempre preservada, enquanto estruturas secundárias nos mRNA podem gerar artefatos durante a transcrição reversa que causam interpretações incorretas da orientação das moléculas.

3.11.3 Hibridização dos experimentos com lâminas de *microarrays* de oligonucleotídeos

Hibridizações com 750ng cada de cRNA marcado com Cy3 e Cy5 foram realizadas seguindo as recomendações do *Agilent In situ Hybridization Kit-plus* (Agilent Technologies). Foram utilizadas 4 lâminas no total, sendo que em cada par foi hibridizada uma das réplicas biológicas com cada uma das amostras marcadas com cada fluoróforo (*dye-swap*).

As lâminas foram lavadas e processadas de acordo com o protocolo *Agilent 60-mer Oligo Microarray* e as imagens foram obtidas por excitação com laser utilizando o *GenePix 4000B* (Molecular Devices). Os seguintes parâmetros foram utilizados como padrão para esse procedimento: resolução de 10 μ m por pixel; comprimento de onda de excitação de 532nm (para alvos marcados com Cy3) ou 633nm (para alvos marcados com Cy5); e tensão PMT de 350 volts (para alvos marcados com Cy3) e 600 volts (para alvos marcados com Cy5).

3.11.4 Obtenção de imagens e valores de intensidade comparáveis

Seguindo a rotina de tratamento de dados estabelecida de nosso laboratório durante as análises com as lâminas de cDNA, os valores de intensidade foram extraídos das imagens obtidas com as lâminas de oligonucleotídeos utilizando o *ArrayVision 8.0* (Imaging Research Inc.), e os valores de intensidade de réplicas marcadas com Cy5 e Cy3 foram corrigidos usando algoritmo *LOWESS* (*locally weighted linear regression algorithm*), disponível no pacote R (<http://www.r-project.org/>). O processo de normalização para deixar os diferentes experimentos comparáveis foi utilizar como fator de normalização a média das 300 sondas presentes no *microarray* complementares aos mRNAs controle que foram adicionados durante a marcação, removidos os 20% dos valores de intensidade extremos, mais altos e mais baixos.

4 RESULTADOS

4.1 Identificação e caracterização de membros representativos de novas famílias de pequenas GTPases monoméricas.

Uma das estratégias utilizadas para contribuir na exploração do conhecimento potencial gerado durante o HCGP (*Human Cancer Genome Project*) baseou-se em uma busca informática por mensagens codificantes para novos parálogos, isto é, produtos novos porém similares a proteínas já conhecidas, e que além disso tivessem envolvimento demonstrado em processos de tumorigênese (detalhes no item 3.1.1). Com essa abordagem um conjunto de candidatos foi selecionado, e desses, dois novos transcritos humanos similares a mensagens codificantes para proteínas membros da superfamília de pequenas GTPases foram identificados e parcialmente caracterizados. Os resultados descritos a seguir mostram os passos desse estudo que apontou para evidências mais concretas do envolvimento de um dos transcritos em câncer, e ainda estimulou o início de uma caracterização mais aprofundada do seu produto protéico.

4.1.1 *RASL11A* e *ARL-9*, dois novos genes codificantes para pequenas GTPases monoméricas.

Por comparação do potencial codificante do conjunto de cerca de três milhões de ESTs geradas a partir de RNA proveniente de tumores humanos presentes no GenBank com a lista de 1.127 proteínas humanas com função já relacionada com câncer, foram encontradas 16 ESTs candidatas a parálogos de genes humanos. O conjunto de informações referentes aos transcritos codificantes com relacionados com o câncer, assim como das ESTs candidatas a parálogos estão na Tabela 1.

Tabela 1 – Novos parálogos: dados de alinhamento das seqüências parálogas preditas e de seu gene relacionado com câncer (CRG) correspondente.

CRG	CRG Locus	CRG mapeamento	CRG tamanho	CRG estrutura	CRG domínios	Parálogo tamanho	Parálogo estrutura	Parálogo mapeamento	Parálogo domínios	% Similaridade Local (amino ácidos)	% Identidade Local (amino ácidos)	% Cobertura CRG (amino ácidos)	% Cobertura Parálogos (amino ácidos)	Gaps	Evidência experimental
Talin 1	TLN1	9p13	2,541 aa	57 exons	smant000096 smant000307 smant000307	2,777 aa	59 exons	15q22.2	smant00295 smant00307 pfam01608	85	76	95	100	20	EST, mRNA
Growth arrest-specific 2	GAS2	11p14.3-15.2	313 aa	6-7 exons	LOAD_gas2arco pfam0187 smant00743 smant00033	182 aa	8 exons	12q23.1	pfam0187 smant00243 LOAD_gas2arco smant00033	71	62	50	86	2	EST
Lactucan/calmodulin-dependent protein kinase I	CAMK1	3p25.2	370 aa	11 exons	smant000096 smant00219 LUSU_1ain smant00091 pfam01429	365 aa	12 exons	4q31.23	smant000096 smant00219	62	47	74	80	7	EST
Methyl-CpG binding domain protein 2	MBD2	18q21	411 aa	3 exons	smant00091 pfam01429	194 aa	1 exon	19p13.2	smant00091	59	42	44	94	2	EST, mRNA
Cytokine-inducible kinase	CNK	1p34.1	607 aa	15 exons	smant00220 pfam00069 smant00219 pfam00659	336 aa	14 exons	19p13.3	smant00220 pfam00069	56	42	61	99	13	EST, mRNA (FLJ30246)
Dual specificity phosphatase 3 (vaccinia virus phosphatase \41-related)	DUSP3	17q21	185 aa	3 exons	smant00195	1024 aa	3 exons	1q24.1	smant00195	55	41	97	19	3	EST
Ras-related protein 2A	RAP2A	13q34	183 aa	2 exons	smant00173 smant00176 smant00174 smant00176	242 aa	4 exons	13q12.2	smant00173 smant00176 smant00174 pfam01926	55	37	89	70	3	EST
Cytochrome P450, subfamily 1B (phenobarbital-inducible), polypeptide 6	CYP2B6	19q13.2	491 aa	9 exons	pfam00067	388 aa	5 exons	11p15.2	pfam00067	54	38	76	97	6	EST
GTP binding protein	ARL3	10q23.3	182 aa	5 exons	pfam00026 smant00177 smant00094 smant00102 smant00186 smant00194	123 aa	2 exons	4q12	pfam00026 smant00177	51	33	60	93	2	EST
Phosphatase and tensin homolog	PTEN	10q23.3	403 aa	9 exons	smant00102 smant00186 smant00194	254 aa	6 exons	7p13	smant00404 pfam0102	51	38	52	83	6	EST
Butyrylcholinesterase	BCHE	3q26.1-q26.2	602 aa	3 exons	pfam0135	656 aa	5 exons	17p13.1	pfam0135	50	37	74	69	12	EST, mRNA (KIAA1386, partial cds)
Cyclin-dependent kinase 2	CDK2	12q13	298 aa	7 exons	smant00020 pfam00069 smant00219	341 aa	2 exons	1p36.11	smant00020 pfam00069 smant00219	50	35	96	97	17	EST, mRNA
Forkhead protein FREAC-1	FOXF1	16q24	354 aa	2 exons	smant000339	570 aa	6 exons	2p11.1	smant000339 pfam0260 pfam00691	47	38	92	65	19	EST
PIG10	ENC1	5q12-q13.3	516 aa	1 exon	pfam00651 smant00612 smant00612	649 aa	9 exons	2q37.3	smant00612 LOAD_ketch pfam01344	47	31	93	83	17	EST
DAP-kinase 1	DAPK1	9q34.1	1401 aa	24 exons	smant00220 pfam00023 smant00005	171 aa	5 exons	2q11.1	pfam00023	46	31	12	84	7	EST
XPG-complementing protein	ERCC5	13q22	1,180 aa	15 exons	pfam00752 pfam00867	687 aa	7 exons	2p24.22	smant00485 pfam00867	37	24	50	94	24	EST

Análise de similaridade/identidade Local: BestFit (pacote GCG). Matrix de comparação: Blossum 35, Parâmetros: -GAPweight=10 -LENGhtweight=1 -ENDWeight

Dentre as seqüências identificadas pela abordagem, duas aparentemente codificavam para proteínas similares a membros da superfamília Ras de pequenas GTPases monoméricas, cuja representatividade na lista inicial de genes relacionados com câncer era de 36 membros. Devido a grande participação de muitos membros dessa família em processos relacionados à tumorigênese (discutido no item 1.5.1), o estudo se concentrou na clonagem e caracterização parcial desses dois novos transcritos humanos.

O primeiro gene candidato parece codificar para um produto que possui 55% de similaridade com a proteína Rap2a, e foi posteriormente denominado *RASL11A* (*RAS-like, family 11, member A*). O segundo, denominado *ARL-9* (*ADP-ribosylation factor-like 9*) parece codificar para um produto que possui 51% de similaridade com a proteína Arl-3. Os transcritos de *RASL11A* (Figura 15A) e *ARL-9* (Figura 15B) amplificados a partir de cDNA de próstata, mapeiam separados, nas regiões cromossômicas 13q12.2 e 4q12, respectivamente, e cada um possui quatro exons distintos. Como representado na Figura 15A, o transcrito completo de *RASL11A*, obtido por RACE-PCR, possui 1.543bp com uma fase aberta de leitura (*Open Reading Frame* – ORF) que codifica para uma proteína com 242 a.a., enquanto o fragmento de *ARL-9*, obtido por PCR com oligonucleotídeos específicos, possui 836bp com uma ORF que gera uma proteína com 187 a.a. (Figura 15B).

4.1.2 *RASL11A* e *ARL-9* constituem novas subfamílias de pequenas GTPases.

Utilizando as seqüências protéicas hipotéticas codificadas por essas novas GTPases monoméricas foi possível identificar no GenBank (descrito no item 3.1.2) outras proteínas similares a essas e alinhá-las para: i) identificar com quais

proteínas, e/ou subfamílias de pequenas GTPases, os novos genes seriam mais relacionados; ii) localizar os resíduos conservados presentes nos domínios característicos das proteínas da superfamília; e iii) inferir sobre novos motivos conservados nas proteínas mais relacionadas aos novos genes.

Figura 15 – *RASL11A* e *ARL-9* mapeiam nas regiões cromossômicas 13q12.2 e 4q12. Exons de *RASL11A* (A) e *ARL-9* (B) são mostrados como caixas preenchidas. Regiões codificantes para proteína são preenchidas em cinza e as não traduzidas (UTR) em preto. O tamanho das proteínas preditas está no topo de cada painel. Tamanhos dos exons e introns estão em pares de base (bp).

Com uma abordagem por comparação das seqüências primárias, foi determinado que entre as 10 proteínas mais similares a Rasl11a constam claramente membros representativos das famílias Ras/Rho/Rab (Figura 16A), enquanto entre as 10 mais similares a Arl-9 aparecem proteínas pertencentes às famílias Arf/Arl (Figura 16B). Com essa análise ainda foi possível identificar outros possíveis parálogos, também não descritos, para cada uma das novas mensagens. Com 74% de similaridade de seqüência primária com o produto protéico putativo de

RASL11A, foi identificado *RASL11B* (NP_076429), anteriormente apenas anotado como codificante para uma proteína hipotética humana (Figura 16A). De forma semelhante, foi identificado *ARL-10* (NP_775935), cujo produto protéico possui 62% de similaridade com Arl-9 (Figura 16B).

Figura 16 – **Produtos protéicos hipotéticos de *RASL11A* e *ARL-9* possuem parálogos e são similares a membros de subfamílias distintas de pequenas GTPases.** As seqüências de Rasl11a (**A**) e Arl-9 (**B**) foram alinhadas com seqüências do banco de dados de proteína do NCBI, usando BLASTp. Os 10 melhores *hits* estão indicados no eixo X de cada painel. Todas as proteínas mais similares a Rasl11a pertencem a subfamílias Ras/Rho/Rab, incluindo seu parálogo Rasl11b, enquanto as de Arl-9 e Arl-10 pertencem a subfamílias Arf/Arl. Os parâmetros de alinhamento obtidos pelo BLASTp são *bit-score* (azul claro), porcentagem de identidade (azul escuro) e porcentagem de substituições conservativas ou similaridade (vermelho).

As proteínas da subfamília Rasl11, filogeneticamente mais próximas da família Ras, compartilham a maioria dos resíduos conservados do domínio G das pequenas GTPases (Figura 18). Entretanto essas proteínas exibem uma substituição por Asn na posição correspondente ao resíduo Thr35 (relativo a H-Ras), situada no sub-domínio denominado *Switch* 1 do domínio G (Figura 18) e conservado em todas as GTPases já caracterizadas funcionalmente (Vetter & Wittinghofer, 2001a), que está esquematizado na figura 7. Este achado sugere que o mecanismo de ativação possa ser estruturalmente diferente nessa nova subfamília de pequenas GTPases, assim como discutido posteriormente (ver item 5.1). Além disso, essas proteínas provavelmente não apresentam as modificações lipídicas em aminoácidos características da família, pois não possuem o motivo CaaX (Reuther & Der, 2000a) na região carboxi-terminal. Porém, apresentam agrupamentos de resíduos básicos que devem garantir sua ligação à membrana plasmática (Lee *et al.*, 1996), normalmente essencial para suas funções biológicas (Figura 18, resíduos em verde).

Arl-9 e Arl-10 apresentam além de resíduos do domínio G a assinatura canônica característica do *Interswitch Toggle* (wDvGGqxxxRxxW) da família Arf de pequenas GTPases (Pasqualato *et al.*, 2002b), com apenas algumas alterações conservativas (Figura 18) já verificadas em outras proteínas funcionais (Kuge *et al.*, 1994; Eckley & Schroer, 2003). Importante ressaltar que alguns desses resíduos estão apenas conservados em Arl-9 se usarmos como códon de iniciação de tradução CUG. O uso de desse tipo de iniciação é documentado em proto-oncogenes, fatores de transcrição e fatores de proliferação, e normalmente aparece em um contexto genômico de sequência (Boeck & Kolakofsky, 1994) similar ao observado para o *ARL-9*.

Figura 18 – Alinhamento das novas pequenas GTPases com membros selecionados da superfamília. Membros selecionados das subfamílias alinhados com as novas proteínas Ras111a/b e Arl-9/10. As cinco regiões conservadas do domínio G estão entre as caixas. Resíduos normalmente invariantes estão em vermelho. O resíduo Asn que está na posição da Thr no switch 1 está em negrito. Resíduos do *interswitch toggle* da família Arf estão em azul. Sinais de prenilação estão sombreados em cinza e os resíduos básicos carboxi-terminais estão em verde.

4.1.3 Produtos protéicos codificados por *RASL11A* e *ARL9* são muito conservados entre os eucariotos.

Foram encontradas também diversas proteínas hipotéticas, entre os demais eucariotos, similares as proteínas da subfamília Rasl11 (Figura 19A e B) e Arl-9/Arl-10 (Figura 20A e B), mostrando uma grande conservação evolutiva dessas novas proteínas. Cada membro da subfamília humana Rasl11 (Figura 19A), assim como da subfamília Arl-9/Arl-10 (Figura 20A), situa-se em ramos separados, sendo que em cada ramo encontram-se ortólogos muito relacionados.

Ortólogos próximos para Rasl11a foram encontrados apenas em *Mus musculus* e em *Rattus norvegicus*, entretanto para Rasl11b foram encontrados também em *Danio rerio* e *Xenopus leavis* (Figura 19). Ainda foram encontrados ortólogos em *Drosophila sp* e *Anopheles gambiae*, ligeiramente menos relacionados à subfamília Rasl11 (Figura 19). Importante notar pelo múltiplo alinhamento que as proteínas Rasl11 são muito conservadas dentro dos metazoários e que todos os ortólogos encontrados apresentam a substituição do resíduo canônico Thr do *switch* 1, por um resíduo de Asn (Figura 19B).

Resultados bastante similares foram obtidos para Arl-9/Arl-10. Enquanto foram encontrados ortólogos próximos para Arl-9 apenas em *Mus musculus* e em *Rattus norvegicus*, para o Arl-10 foi encontrado também em *Xenopus leavis* (Figura 20A). O alinhamento múltiplo dessas proteínas mostra elevada conservação de aminoácidos, inclusive nas alterações observadas no consenso do *Interswitch Toggle* (Figura 20B).

Figura 19 – **As proteínas Rasl11 são muito conservadas entre os metazoários.** (A) Árvores filogenéticas da subfamília Rasl11 incluindo os ortólogos encontrados no GenBank nessa análise. Os números representam a confiança nos ramos (em 1.000 amostragens). (B) Alinhamentos dos ortólogos de Rasl11 mostrando as cinco regiões conservadas do domínio G entre as caixas e os resíduos invariantes da superfamília estão em vermelho. O resíduo Asn que está na posição da Thr no *switch* 1 está em negrito. Os resíduos básicos carboxi-terminais estão em verde.

Figura 20 – As proteínas da subfamília Arl-9/Arl-10 são muito conservadas entre os metazoários. (A) Árvore filogenética da subfamília Arl-9/10 incluindo os ortólogos encontrados no GenBank nessa análise. Os números representam a confiança nos ramos (em 1000 amostragens). (B) Alinhamento dos ortólogos de Arl-9/10 mostrando as regiões conservadas do domínio G estão entre as caixas e os resíduos invariantes da superfamília em vermelho. Resíduos do interswitch toogle da família Arf estão em azul.

4.1.4 Níveis do transcrito *RASL11A* estão diminuídos em tumores de próstata.

A proteína funcionalmente já estudada mais semelhante a Rasl11A, Rerg com 58% de similaridade e 37% de identidade (Figura 16A), foi recentemente descrita como possuindo atividade supressora tumoral em carcinomas de mama (Finlin *et al.*, 2001). Essa observação, o interesse particular do grupo em estudar

esse processo tumoral, e também a possibilidade de obter amostras de mRNA derivado de tecido prostático não-tumoral e neoplásico, motivaram a verificação do padrão de expressão de *RASL11A*. Utilizando PCR quantitativo em tempo real (ver item 3.7.2), verificamos que a abundância do transcrito *RASL11A* está significativamente ($p=0,00065$) diminuída em amostras de adenocarcinoma de próstata (dos pacientes 4, 9, 525, 110) em relação ao tecido prostático não-tumoral adjacente a tumores (dos pacientes 6, 105, 106, 86, 88) (Figura 21A), sugerindo um possível papel de supressor tumoral também para este novo transcrito.

Figura 21 – ***RASL11A* tem expressão diminuída em tumores de próstata.** RT-PCR quantitativo em tempo real a partir de mRNA isolado de tumores (barras pretas) ou tecido não-tumoral adjacente ao tumor (barras cinzas) de próstata. Os níveis dos transcritos de *RASL11A* (**A**) e *ARL-9* (**B**) foram normalizados pelo mRNA de *beta-tubulina*. Valores de p para a expressão diferencial foram determinados para *RASL11A* ($p=0,00065$) e *ARL-9* ($p=0,52$).

A expressão em diferentes tecidos e o tamanho desse transcrito foram também investigados através da hibridização de sondas radioativas com membranas comerciais de nitrocelulose contendo mRNA de diversos tecidos (ver item 3.6.1). Através de Northern Blot foi observada a presença constitutiva de *RASL11A* em vários tecidos humanos. Duas formas alternativas para o transcrito de *RASL11A* (1.6 e 1.2 kb) foram claramente detectadas em bexiga, próstata, testículo e colón (Figura 22). A expressão de *RASL11A* parece ser bastante fraca em ovário e ausente em timo e leucócitos (Figura 22).

Figura 22 – **Duas isoformas de *RASL11A* são transcritas nos tecido humanos.** (A) Transcrição de *RASL11A* em diversos tecidos detectada por Northern blot. (B) A presença de quantidades comparáveis de mRNA em cada poço foi confirmada pela subsequente hibridização da membrana com sonda de *beta-actina*.

4.1.5 Sistema de expressão de proteínas Rasl11 em procariotos.

Para possibilitar o início de uma caracterização mais aprofundada dessa nova subfamília de pequenas GTPases identificada nesse trabalho (Louro *et al.*, 2004)⁴ com possível envolvimento em câncer de próstata foi construído um sistema de expressão heteróloga das proteínas Rasl11 em bactérias (descrito no item 3.9.1).

⁴ Louro *et al.* 2004 (Artigo publicado). ANEXO 4 - incluso no CD-ROM suplementar a esta Tese.

Dessa forma, bactérias transformadas com o vetor pET21 vazio (sem inserto) e com a construção pET21-Ras11b foram induzidas a expressar as proteínas de interesse com adição de IPTG (descrito no item 3.9.2). Esses experimentos se iniciaram com a proteína Ras11b devido a problemas de contaminação durante a construção do pET21-Ras11a. Quando essa construção foi seqüenciada, identificou-se na verdade um outro clone da construção com Ras11b.

Na figura 23 pode-se visualizar um SDS-PAGE mostrando o acúmulo de proteína de cerca de 27kDa no lisado total das bactérias contendo pET21-Ras11b, comparado com o vetor vazio, durante os tempos de indução por IPTG. Observou-se que até 6 horas de indução a quantidade de proteína expressa no tamanho compatível com o esperado para as novas GTPases monoméricas da sub-família Ras11 (~27kDa) ainda está aumentando, proporcionalmente ao tempo de indução utilizado.

Figura 23 – **Indução da expressão de Ras11b por IPTG em bactéria.** SDS-PAGE com 12% de acrilamida mostrando a expressão heteróloga de Ras11b induzida por IPTG em diferentes tempos: 1,5h; 3h; 4,5h. Marcador de massa molecular de proteínas *Prestained Protein Marker* (Fermentas). V = vetor pET21b sem inserto (controle negativo) e R = pET21b com inserto da região codificante de Ras11b.

Devido à existência da cauda de seis histidinas fusionada às proteínas em estudo, foi possível assegurar a natureza do produto expresso através de *Dot Blot* com anticorpos anti-histidina (descrito no item 3.9.3). A figura 24 apresenta a confirmação da expressão de uma proteína com cauda de histidinas presente apenas nas bactérias transformadas com pET21-Ras11b.

Figura 24– **Confirmação da expressão da proteína de fusão His₆-Ras11b por Dot Blot.** Diluições seriadas de cada fração do extrato de bactérias aplicadas a uma membrana de nitrocelulose e incubadas com anticorpos contra poli-histidinas. His₆-Ca⁺²ATPase adicionada como controle positivo e uma construção de pET21 sem inserto como controle negativo. Dois clones de His₆-Ras11b estão mostrados.

4.1.6 Expressão de Ras11b recombinante em corpos de inclusão para imunização de coelhos.

Foi desenhada uma estratégia de imunização para geração de anticorpos policlonais a partir de uma preparação de corpos de inclusão obtidos da expressão heteróloga de Ras11b. A utilização dessa metodologia foi escolhida devido à

observação inicial, através de SDS-PAGE, que as condições testadas para indução da expressão de Rasl11b levavam a um acúmulo de proteínas na fração insolúvel (Figura 25 – coluna 3). Provavelmente esses produtos não estão corretamente dobrados e foram armazenadas em corpos de inclusão devido a uma indução de expressão muito robusta e acelerada.

Esse resultado foi extremamente indesejado para o prosseguimento da caracterização funcional das GTPases monoméricas através de ensaios bioquímicos, pois além de purificação por cromatografia de troca iônica em coluna de níquel necessitaria de uma etapa adicional de renaturação estrutural das proteínas. Todavia, a proporção relativa da proteína induzida nesses corpos de inclusão é bastante elevada, sendo relatada na literatura como de no mínimo 70% (Harlow & Lane, 1998). Assim, essa característica foi utilizada para a obtenção dos antígenos utilizados em uma estratégia de imunização. É possível notar que quantidades baixas de proteínas contaminantes de bactéria que foram “co-purificadas” a partir de interações hidrofóbicas com os corpos de inclusão (Figura 25 - coluna 3), foram praticamente eliminadas após a lavagem do precipitado dos corpos de inclusão (ver item 3.10.1) (Figura 25 - coluna 6).

Esses antígenos foram então utilizados para imunização de coelhos, assim como descrito (item 3.10.2), entretanto durante esse doutorado não chegou a ser realizado um teste sistemático de reatividade dos soros.

Figura 25 – **Corpos de inclusão com His₆-RasI11b para imunização em coelhos.** SDS PAGE com 12% de acrilamida, mostrando a eliminação das proteínas contaminantes de bactéria “co-purificadas” com os corpos de inclusão, após o tratamento descrito no texto. Marcador de massa molecular de proteínas *Prestained Protein Marker* (Fermentas). **Poço 1** – Lisado total. **Poço 2** – Fração solúvel. **Poço 3** – Fração insolúvel com corpos de inclusão. **Poço 4** – Lavagem com uréia. **Poço 5** – Corpos de inclusão após a lavagem. **Poço 6** – Corpos de inclusão usados para imunização.

4.2 Transcritos intrônicos marcadores de prognóstico em câncer de próstata.

Como parte das estratégias para aprofundar a exploração das informações geradas durante o projeto HCGP, o grupo de pesquisa coordenado pelo Prof. Dr. Sergio Verjovski-Almeida construiu uma plataforma de *microarray* enriquecida em mensagens novas, provenientes principalmente de regiões intrônicas de genes codificantes para proteína (detalhes da plataforma no item 3.5.1). Essa abordagem resultou na identificação de um conjunto de novas mensagens humanas, longas e sem *splicing*, sem potencial codificante para proteínas, transcritas a partir de introns de mRNAs já anotados, e mais importante com perfil de expressão correlacionado com o grau de diferenciação tumoral de próstata. O conjunto de resultados que levou a esse importante achado está descrito abaixo.

4.2.1 Transcritos intrônicos são expressos em próstata

Inicialmente, para verificar qual fração dos novos transcritos é expressa em tecidos prostáticos foram realizadas hibridizações contra a plataforma intrônica de *microarrays* de cDNA utilizando RNA poli-(A)+ derivado de 20 amostras normais adjacentes ao tumor, e 27 amostras tumorais provenientes de 27 pacientes. Como pode ser verificado na tabela 2, foram detectadas frações comparáveis de transcritos expressos em próstata que mapeiam em regiões intrônicas de mensagens manualmente curadas (*RefSeq* - *Reference Sequence*), em regiões não anotadas (intergênicas), assim como em exons conhecidos de genes humanos bem anotados (genes *RefSeq*).

Mapeamento do clone ORESTES	Número de regiões genômicas únicas	Número de clones ORESTES		
		Total na coleção	Total no <i>microarray</i>	Expresso em próstata
Introns de <i>RefSeq</i>	37.899	150.236	619	346 (56%)
Regiões não anotadas	45.400	136.628	1.019	539 (53%)
Exons conhecidos de <i>RefSeq</i>	29.444	390.263	2.048	1.369 (67%)

Tabela 2 – Atividade transcricional em próstata de fragmentos de mensagens que mapeiam em diferentes regiões do genoma humano, e representação dos clones. “Total na coleção” se refere ao número de clones na coleção gerada pelo *Human Cancer Genome Project* pela técnica ORESTES (Dias-Neto *et al.*, 2000). “Expresso em próstata” se refere aos fragmentos de mensagens que cujas intensidades passaram pelo filtro de expressão (3 desvios padrão acima da média das intensidades dos controles negativos da lâmina) em pelo menos 5 amostras de tecidos (total de 47). A porcentagem de fragmentos expressos está indicada entre parênteses. A anotação de mapeamento de cada fragmento foi realizada em 2004.

A observação de que grande parte dos transcritos totalmente intrônicos sem *splicing* representados na plataforma são expressos em próstata (Tabela 2) é um forte argumento contra abordagens informáticas conservadoras, que normalmente excluem de análises de produtos de transcrição esse tipo de mensagem intrônica, argumentando com a possibilidade de serem contaminantes de DNA genômico (Sorek & Safer, 2003). Além disso, esse resultado aponta para uma falha nas análises de expressão gênica para estudos de doenças, que utilizam plataformas exônicas comerciais, contendo apenas sondas para genes com *splicing* e codificantes para proteínas, previamente caracterizados. Com essa abordagem mais ortodoxa, usando somente sondas de genes conhecidos, não é possível detectar alterações nessas novas mensagens intrônicas, potencialmente regulatórias como discutido posteriormente, nas condições patológicas em estudo.

4.2.2 Níveis de transcritos intrônicos não codificantes são correlacionados ao grau de diferenciação tumoral em próstata

Com o objetivo de verificar se os novos transcritos expressos em próstata, principalmente os derivados de regiões intrônicas, possuem alguma significância biológica foram realizadas análises utilizando apenas os valores de expressão obtidos nas hibridizações com amostras tumorais de próstata. Já sem nenhuma análise estatística prévia, o perfil de expressão das 27 amostras apontava para a existência de uma assinatura gênica diferenciada para amostras com *Gleason Score* (GS) (Gleason, 1977) altos (GS 9 e 10). Um *Self-Organizing Map* e uma análise de *Hierarchical clustering* não supervisionada mostraram que todas amostras com alto-GS (9 e 10) agrupam-se juntas, apontando para um perfil de expressão bastante distinto nessas amostras (Figura 26A e B). De fato, alguns trabalhos já haviam

identificados genes com perfis de expressão altamente correlacionados ao GS (Singh *et al.*, 2002; Best *et al.*, 2003; Lapointe *et al.*, 2004).

Figura 26 – **Classificação não-supervisionada de 27 amostras de tumor de próstata.** (A) *Self organizing map* (Tamayo *et al.*, 1999). (B) *Hierarchical clustering* (Sanchez-Carbayo *et al.*, 2002). Círculos coloridos abaixo de cada painel indicam o *Gleason score* (GS) das amostras tumorais, servindo aqui apenas como referência já que a “clusterização” não foi supervisionada. Todas as amostras alto GS (9 e 10) estão agrupadas em conjunto em (A) e em (B), indicando assinaturas de expressão distintas no conjunto total de seqüências do *microarray* (linhas em B). As 27 amostras tumorais são apresentadas nas colunas em (B). O padrão de expressão de cada mensagem está representado pelo número de desvios-padrão acima (vermelho) ou abaixo (azul) do valor médio da expressão em todas as amostras. Os números na árvore indicam a robustez da classificação em 10.000 amostragens.

Em seguida, utilizando a metodologia de análise de correlação de Pearson com permutações (detalhes no item 3.1.3) foi realizada uma busca por transcritos significativamente correlacionados com o GS, apenas entre as amostras de valores baixos (5 e 6) e de valores altos (9 e 10). Com essa abordagem foi identificado um conjunto classificador com 56 mensagens cujos níveis de expressão eram significativamente ($p < 0,001$) correlacionados com a distinção entre baixo/alto GS (Figura 27A). Dentro desse, foram identificados um subgrupo 27 seqüências representativas de exons conhecidos de RefSeq, 23 transcritos intrônicos de RefSeq e 6 mensagens provenientes de regiões genômicas sem anotação.

Figura 27 - **Transcritos correlacionados com o grau de diferenciação (GS) em tumores de próstata.** Os painéis mostram a matriz de expressão de 56 transcritos correlacionados com a distinção de classe baixo/alto GS ($p \leq 0,001$) (A) e sua aplicação nos perfis de expressão dos pacientes de GS intermediário (GS 7) (B). Os transcritos (colunas) estão ordenados pela sua correlação de Pearson e os pacientes (linhas) pela sua correlação (r) com o perfil médio das amostras de baixo GS, mostrada nos painéis da direita. Fragmentos intrônicos têm o nome do *locus* gênico em azul. Níveis de expressão em cada paciente estão representados pelo número de desvios-padrão acima (vermelho) ou abaixo (azul) da média do transcrito em todas as amostras.

A aplicação do classificador baixo/alto GS ao perfil de expressão de dezesseis amostras com grau intermediário de diferenciação tumoral (GS 7), revelou dois grupos com padrão de expressão distintos dentre pacientes com características

histológicas bastante similares (Figura 27B). Enquanto a maioria das amostras GS 7 exibiram uma correlação positiva com as amostras de baixo GS, quatro amostras (pacientes 17, 19, 47 e 3) apresentaram perfil de expressão mais correlacionado com amostras de alto GS. Esse resultado provavelmente está relacionado ao fato de que pacientes com tumores de próstata de GS 7 apresentam características clínicas bastante heterogêneas (Epstein *et al.*, 1996).

Do conjunto de transcritos correlacionados, a fração de mensagens mais robustas como classificadores de baixo/alto GS foi determinada por *leave-one-out* (detalhes no item 3.1.3). Transcritos que apareceram em pelo menos 60% dos classificadores gerados são mostrados na figura 28. Surpreendentemente, sete das doze mensagens mais robustas são transcritas a partir de regiões intrônicas de *RefSeq*. As cinco demais representam exons conhecidos de *RefSeq*, dos quais três genes, *EIF4E*, *TMEPA1* e *RNF139*, têm sido associados com a progressão de diferentes tipos de câncer (Gemmell *et al.*, 1998; Brunschwig *et al.*, 2003; Ramaswamy *et al.*, 2003).

4.2.3 Caracterização de transcritos intrônicos relacionados com câncer

Para caracterizar o tamanho e a orientação transcricional dos sete transcritos intrônicos identificados no classificador baixo/alto GS mostrados na figura 28, foram utilizadas abordagens de RACE-PCR (detalhes no item 3.4.1), Northern Blot (detalhes no item 3.6.2), e transcrições reversa com orientação específica (OSRT) seguida de PCR (detalhes no item 3.7.1).

Figura 28 - **Transcritos intrônicos mais robustos na distinção de classe baixo/alto GS.** (A e B) 12 mensagens classificadoras de grau de Gleason. O painel (A) mostra os transcritos correlacionados ($p \leq 0,001$) que aparecem em pelo menos 60% dos classificadores (fração indicada no eixo x), identificados por *leave-one-out*. Barras verde-escuras correspondem a exons conhecidos de *RefSeq*, e barras verde-claras são transcritos intrônicos. O painel (B) mostra o nível de expressão de cada transcrito (linha) para cada amostra de tumor (coluna). Níveis de expressão em cada paciente estão representados pelo número de desvios-padrão acima (vermelho) ou abaixo (azul) da média do transcrito em todas as amostras. (C) Mensagens intrônicas foram detectadas tanto com orientação senso quanto antisenso. Os controles para garantir a ausência de contaminação por DNA genômico (- RT) e de auto-anelamento do RNA (- oligo) estão indicados.

O resultado de RACE-PCR para a sequência intrônica de *TLE3* mostrou que essa mensagem é na verdade um novo exon do transcrito de mesmo *locus* (Figura 29A), que codifica para uma proteína membro da via de sinalização Notch, envolvida em diferenciação epitelial (Liu *et al.*, 1996). O resultado do Northern blot utilizando sonda específica para esse novo exon mostrou a presença de uma banda com cerca de 5,5 kb, a qual é totalmente compatível com uma mensagem completa de *TLE3* com esse novo fragmento incluso (Figura 29B). A adição desse exon introduz um códon de terminação de tradução precoce que trunca o domínio WD-40 da porção carboxi-terminal, sendo observada a expressão dessas formas alternativas na maioria dos metazoários cuja função acredita-se estar envolvida na regulação negativa das proteínas Tle3 (Chen & Courey, 2000).

Figura 29 – **Novo exon no transcrito de *TLE3***. O mapeamento genômico dos fragmentos de cDNA dupla-fita presentes no *microarray* (em azul) e as coordenadas cromossômicas estão indicadas. (A) Forma alternativa de *splicing* do gene *TLE3*, confirmadas por RACE-PCR (bloco vermelho). (B) Forma alternativa de *splicing* do gene *TLE3*, um transcrito com 5,53 kbp detectado por *Northern blot* em três linhagens derivadas de tumores de próstata, utilizando uma sonda complementar a fita senso do novo exon de *TLE3*.

Utilizando RACE-PCR com uma biblioteca comercial de pulmão foi determinado que as mensagens intrônicas de *RASSF1*, *AP1GBP1* e *PPHLN1* eram longas (0,6 a 1,1 kb), sem *splicing* e contidas totalmente nas regiões intrônicas dos RNAs codificantes de proteína de mesmo *loci* (Figuras 30A, B e C - caixas amarelas). Aparentemente essas mensagens estendidas não geram proteínas, pois análise utilizando *ESTScan* (Iseli *et al.*, 1999) não apontou nenhum potencial codificante. Para as mensagens intrônicas de *RASSF1* e *AP1GBP1*, um RACE-PCR utilizando a biblioteca de cDNA de cólon clonada com orientação determinou a presença de transcrição antisenso (Figuras 30A e B - setas vermelhas). Ainda, a orientação antisenso das mensagens intrônicas de *RASSF1* e *PPHLN1*, assim como seu tamanho, foi confirmado por Northern Blot com sondas de fita simples (Figuras 30D e E). Produtos de RACE-PCR para os 3 restantes transcritos intrônicos, com as bibliotecas utilizadas, não estenderam os respectivos fragmentos de cDNA depositados na plataforma de *microarray*.

Figura 30 - Longos RNAs Intrônicos antisense correlacionados com GS caracterizados por RACE-PCR e Northern Blot. O mapeamento genômico dos fragmentos de cDNA dupla-fita presentes no *microarray* (em azul) e as coordenadas cromossômicas estão indicadas. (A–C) Mapeamento dos longos transcritos intrônicos de *asRASSF1*, *asPPHLN1* e *asAP1GBP1* estendidos por RACE-PCR. Flechas vermelhas mostram transcritos obtidos por RACE-PCR utilizando bibliotecas de cDNA humano com orientação definida (côlon). Blocos amarelos mostram transcritos obtidos por RACE-PCR com biblioteca de pulmão. (D–E) Longos transcritos intrônicos de *asRASSF1* e *asPPHLN1* detectados por Northern Blot utilizando sondas complementares à fita antisense das respectivas regiões intrônicas.

A abordagem de OSRT-PCR (ver item 3.7.1) mostrou, e também confirmou, que das sete mensagens intrônicas analisadas, seis eram transcritas com orientação antisenso em relação ao intron correspondente (Figura 28C). Surpreendentemente, foram também detectadas mensagens intrônicas senso sobrepostas às antisenso nas amostras de mRNA testadas (Figura 28C).

4.2.4 Validação por PCR quantitativo em tempo real das mensagens intrônicas de *RASSF1*

Utilizando *microarrays* de cDNA fita dupla não é possível determinar separadamente os níveis de expressão provenientes de mensagens senso e antisenso sobrepostas. Assim os resultados de expressão diferencial observados para essas mensagens intrônicas poderiam refletir tanto alterações em uma, quanto em ambas mensagens. Para discriminar entre essas possibilidades foi utilizada uma abordagem de PCR quantitativo em tempo real, após a realização de transcrições reversas com orientação específica (OSRT-qPCR) (para detalhes ver item 3.7.2) para confirmar a expressão diferencial dos transcritos intrônicos de *RASSF1*, o mais correlacionado com o grau de diferenciação tumoral.

A figura 31A mostra que ambos transcritos intrônicos, senso e antisenso, estão significativamente diminuídos em amostras de alto GS, em comparação com amostras de baixo GS ($p < 0,024$ e $p < 0,014$; respectivamente). Como mostra a figura 31B, os níveis dos transcritos exônicos codificantes de proteína de *RASSF1* também são significativamente diminuídos em amostras de alto GS ($p < 0,009$). Na figura 31C vê-se que os níveis de *RASSF1* codificante de proteína são menores ($p < 0,002$) em amostras de tumores de próstata GS 7 tumors cujos perfis de expressão

são negativamente correlacionados com o padrão de baixo GS (pacientes 17, 19, 47 e 3), o que confirma os dados do experimento de *microarray* mostrado na figura 27B. Já havia sido demonstrado na literatura que a hipermetilação da região promotora de *RASSF1* correlacionava-se com o GS de tumores de próstata (Liu *et al.*, 2002). Nossos dados acrescentam o fato de que a expressão do transcrito codificante da proteína *RASSF1* também está correlacionada com o GS de tumores de próstata.

Figura 31 - **Expressão diminuída de *asRASSF1* está em tumores com alto GS.** OSRT-qPCR em tempo real foi usada para medir os níveis de mRNA. Frações relativas foram calculadas em relação à abundância média de mRNA entre todas as amostras testadas. **(A)** Os níveis de mensagens intrônicas antisense (barras hachuriadas) e sentido (barras cheias) transcritas a partir do *locus* de *RASSF1* foram medidos em 3 amostras de tumor com grau de baixo GS (barras cinzas) e alto GS (barras pretas). **(B)** Transcritos codificantes para a proteína Rassf1, medidos nas mesmas amostras. **(C)** Transcritos codificantes para a proteína Rassf1 medidos em amostras de tumor com grau de GS 7 que apresentaram correlação negativa (barras pretas) ou positiva (barras cinzas) com o perfil de expressão médio das amostras de baixo GS obtido nos experimentos de *microarray*. Diferenças de expressão estatisticamente significativas com $p < 0,05$.

4.3 Transcritos intrônicos regulados pelo estímulo de andrógeno.

O trabalho que identificou os Longos RNAs Intrônicos não-codificantes (ncRNAs), e os relacionou com a progressão do câncer de próstata (Reis *et al.*, 2004) descrito anteriormente (ver item 4.2), além de ilustrar a enorme complexidade do programa transcricional humano, permitiu a elaboração de hipóteses sobre as possíveis funções biológicas destes transcritos intrônicos relacionadas ao controle da expressão gênica em eucariotos (Reis *et al.*, 2005). Assim, foram priorizados os estudos relacionados à caracterização dessas mensagens, visando principalmente agregar evidências sobre essas funções regulatórias dos ncRNAs intrônicos. Para alcançar esse objetivo foram desenhados experimentos para avaliar a existência de um mecanismo comum de controle da transcrição dos RNAs intrônicos, que fosse compartilhado com o de mRNAs codificantes para proteína. Essa modulação da expressão em condições fisiológicas normais envolveria, por exemplo, um estímulo hormonal, que acarreta sensíveis alterações nos perfis transcricionais.

4.3.1 Identificação de transcritos intrônicos regulados por andrógeno

Utilizando a plataforma de *microarray* de cDNA enriquecida em transcritos intrônicos (detalhes da plataforma no item 3.5.1) foi realizado um estudo de expressão gênica para identificar mensagens cujos níveis de RNA fossem significativamente alterados em células da linhagem tumoral de próstata em resposta a um tratamento com andrógeno sintético R1881 em diferentes tempos de incubação (detalhes no item 3.2.3). Precedente às hibridizações, a eficiência desses tratamentos foi avaliada por PCR quantitativo em tempo real para determinação dos níveis de expressão de *KLK3* (*PSA*, *Prostate Specific Antigen*) (Nelson *et al.*, 2002) e

TMEPAI (Xu *et al.*, 2000), dois transcritos que já haviam sido demonstrados como regulados por andrógeno no epitélio da próstata (Figura 32).

Figura 32 – **Indução de genes regulados por andrógeno em células de próstata.** Níveis de expressão de *KLK3* (*PSA*, *Prostate Specific Antigen*) e *TMEPAI* determinados por PCR quantitativo em tempo real para confirmar que o tratamento foi efetivo para induzir a resposta de genes regulados em células LNCaP suplementadas com andrógeno sintético de 6 a 48h. *Fold-change* Relativo indica os níveis de RNA de *KLK3* (**superior**) e *TMEPAI* (**inferior**) para cada ponto com relação ao tempo 0, e são mostrados para células tratadas com andrógeno (barras pretas) e para as células controle sem tratamento (barras brancas).

Com a validação do estímulo do andrógeno nas células LNCaP, um total de 26 hibridizações, correspondentes aos tempos de tratamento e controles, foi realizada. Os valores de intensidade normalizados provenientes desses experimentos foram analisados com a abordagem de SAM (Tusher *et al.*, 2001) para a identificação dos transcritos com alterações de abundância significativas após

exposição ao andrógeno de 6 a 48 horas (detalhes no item 3.1.3). Como resultado, foram detectadas 168 mensagens exônicas codificantes para proteína (Figura 33A) e 39 transcritos intrônicos (Figura 33B) com padrão de expressão temporal estatisticamente diferente em células tratadas com andrógeno em comparação com os controles sem tratamento. A lista completa dos transcritos exônicos codificantes para proteína regulados por andrógeno é apresentada na Tabela Suplementar 1 (ANEXO 5).

Os RNAs intrônicos regulados pelo andrógeno puderam ser agrupados em três diferentes conjuntos baseados em distintos padrões de expressão temporal após o tratamento (Figura 33C). O Grupo I contém RNAs cujos níveis foram diminuídos de 2 a 3 vezes até 24 horas de exposição ao andrógeno. Os Grupos II e III incluem RNAs cujos níveis foram aumentados com 24 horas de tratamento: enquanto o Grupo II apresenta RNAs aumentados de 2 a 3 vezes, o grupo III apresenta mensagens com aumentos de 3 a 6 vezes. Na maioria dos casos, os níveis de expressão dos transcritos intrônicos retorna a um valor próximo do inicial (tempo 0) após 48 hs de introdução do andrógeno na cultura (Figura 33C).

Todos os 39 RNAs intrônicos identificados são mensagens sem *splicing*, e aparentemente não são utilizados para sintetizar proteínas, pois não apresentam potencial codificante significativo em análises por *ESTScan* (Iseli *et al.*, 1999). Assim, é pouco provável que esses transcritos intrônicos representem novos exons das mensagens codificantes de proteína correspondentes. A lista completa dos RNAs intrônicos regulados por andrógeno é apresentada na Tabela Suplementar 2 (ANEXO 6).

Figura 33 – Perfil temporal de expressão de transcritos que respondem ao tratamento com **andrógeno**. Análise por *Hierarchical clustering* (Sanchez-Carbayo *et al.*, 2002) de transcritos codificantes para proteínas (exônicos) (**A**) e intrônicos (**B**) significativamente ativados ou reprimidos após a exposição das células ao andrógeno é mostrada. Transcritos intrônicos que respondem ao tratamento com andrógeno e que apresentam perfis temporais de expressão semelhantes foram agrupados usando *K-means* (**C**). Grupos de mensagens intrônicas com padrões similares de expressão são indicados por barras verticais coloridas (I–III) no painel **B**. *Fold-change* médio de cada grupo (Grupos I–III) é indicado em **C** por uma linha colorida.

4.3.2 RNAs intrônicos regulados por andrógeno mapeiam em genes envolvidos em transdução de sinais

A ferramenta de comparação de ontologias gênicas (GO) denominada eGOn (<http://www.genetools.microarray.ntnu.no/egon/>) foi utilizada para investigar o enriquecimento de categorias específicas de GO entre os transcritos regulados por andrógeno (detalhes no item 3.1.5).

Para os transcritos intrônicos, foi postulado que representam uma nova classe de RNAs que podem atuar em *cis* como fatores regulatórios (Reis *et al.*, 2005). Dessa forma, cada mensagem intrônica foi anotada com as categorias de GO do mRNA codificante de proteína transcrito a partir do mesmo *locus* gênico. A análise dos *loci* que contém transcritos intrônicos regulados por andrógeno mostrou enriquecimento em categorias de GO normalmente envolvidas em respostas celulares a hormônios: *cell organization and biogenesis* (GO:0016043), *protein binding* (GO:0005515) and *signal transducer activity* (GO:0004871).

A análise com as mensagens exônicas reguladas por andrógeno mostrou várias categorias de GO significativamente enriquecidas ($p < 0,05$), entre elas algumas diretamente relacionadas ao estímulo por andrógeno, como *regulation of metabolism* (GO:0019222), *specific DNA-binding* (GO:0043566) e *intracellular receptor-mediated signaling pathway* (GO:0030522).

4.3.3 Caracterização dos RNAs intrônicos regulados por andrógeno

Um sub-conjunto de RNAs intrônicos (17 mensagens) regulados por andrógeno foi selecionado para caracterização. Transcrições reversas com orientação específica (OSRT) seguida de PCR (detalhes no item 3.7.1) foram utilizadas para determinação da orientação transcricional desse grupo de

mensagens (Figura 34) em relação aos mRNAs codificantes de proteína de mesmo *loci*. Para a grande maioria (13 de 17) o PCR forneceu resultados. Para 10 *loci* dos 13, um transcrito intrônico antisenso foi detectado (Figura 34B), sendo que em seis desses (*DNAJC3*, *SAP18*, *KDEL2*, *FLJ10154*, *MYO5A* e *ATF2*) foi também detectada uma mensagem senso sobreposta; nos quatro *loci* restantes (*GAS6*, *RAB25*, *ITGA6* e *ADD3*) apenas o RNA antisenso foi detectado. Ainda, para três outras mensagens intrônicas entre as 13 (*loci* *ACTN4*, *STARD13* e *ERG*) foram detectadas apenas transcritos com orientação senso (Figura 34B).

Figura 34 – **Orientação transcricional de mensagens intrônicas que respondem ao tratamento por andrógeno.** Transcrição reversa orientação-específica seguida de PCR (OSRT-PCR) foi realizada em um conjunto selecionado de 13 mensagens intrônicas reguladas por andrógeno identificadas pelas análises de *microarray* (**A**) para determinação da orientação transcricional da mensagem intrônica relativa à mensagem codificante para proteína de mesmo *locus*. Nomes a esquerda se referem aos mRNAs RefSeq codificantes para proteína que mapeiam nos mesmos *loci* e definem as regiões intrônicas onde atividade transcricional foi identificada. (**B**) Produtos de PCR separados em gel de agarose 3% detectaram transcrição antisenso (*GAS6*, *RAB25*, *ITGA6* e *ADD3*), senso (*ACTN4*, *STARD3* e *ERG*) ou antisenso/senso simultânea (*DNAJC3*, *SAP18*, *KDEL2*, *FLJ10154*, *MYO5A* e *ATF2*). Os controles negativos, para garantir ausência de contaminação com DNA genômico (-RT) e de auto-anelamento do RNA (- oligo) durante a transcrição, estão indicados.

Como todos transcritos intrônicos identificados não apresentaram potencial codificante pela análise por *ESTScan*, eles realmente devem representar RNAs não-codificantes (ncRNAs) originados de unidades transcricionais independentes. Provavelmente isso também é verdade para as três mensagens com orientação senso, pois apesar de expressos a partir da mesma fita, eles aparentemente não são exons novos das mensagens codificantes para proteína dos *loci*. Transcrição intrônica senso em unidades independentes têm sido observada em outros trabalhos (Kapranov *et al.*, 2002; Bertone *et al.*, 2004).

Para determinação do tamanho completo de algumas mensagens intrônicas, foram realizados experimentos de RACE-PCR com uma biblioteca feita a partir de RNA de adenocarcinoma de próstata (detalhes no 3.4.1). Entre os RNAs intrônicos selecionados para determinação do transcrito completo estavam mensagens expressas com orientação antisenso (*GAS6* e *ITGA6*), senso (*STARD3*) e mensagens antisenso/senso sobrepostas (*MYO5A* e *DNAJC3*).

Os resultados dessa abordagem revelaram que os RNAs antisenso dos *loci* *DNAJC3*, *MYO5A* e *ITGA6* são longos (0,5 a 0,8kb), e mapeiam sem *splicing* totalmente dentro de regiões intrônicas das mensagens codificantes de proteína correspondentes (Figura 35A-C). Um transcrito intrônico do *locus* *STARD3*, também sem *splicing*, com aproximadamente 1,4kb foi estendido a partir do fragmento com orientação senso determinada por OSRT-PCR (Figura 35D).

O RNA intrônico antisenso que mapeia no *locus* *GAS6* foi estendido utilizando uma abordagem de PCR com oligonucleotídeos iniciadores específicos (detalhes no item 3.4.2). A presença de uma única mensagem composta pelo mRNA com cerca de 3,2kb obtido de um biblioteca de cDNA de cérebro, cuja sequência encontra-se disponível no GenBank (AK123478), e o fragmento intrônico de *GAS6* depositado no

microarray foi confirmada pela obtenção de um produto com cerca de 4,2kb em cDNA de adenocarcinoma de próstata (Figura 35E – caixa amarela).

Figura 35 – Caracterização dos transcritos intrônicos não codificantes que respondem ao tratamento com andrógeno. Transcritos intrônicos selecionados foram estendidos por RACE-PCR (A–D, caixas vermelhas) ou PCR com oligonucleotídeos específicos (E, caixa amarela). Os longos transcritos intrônicos antisense sem *splicing* estendidos estão mapeados na sequência genômica (A – E, linhas cinza) com orientação relativa ao transcrito codificante para proteína expresso a partir da fita oposta (A – E, caixas verdes). Fragmentos de cDNA dupla-fita depositados nos *microarrays* estão representados como caixas azuis (A – E). Um transcrito intrônico antisense sem *splicing* que mapeia no *locus* de GAS6 é mostrado (E, flecha preta). Coordenadas cromossômicas e o tamanho dos fragmentos intrônicos estendidos estão indicadas. *Northern blot* de múltiplos tecidos usando sondas complementares ao transcrito antisense do *locus* de GAS6 confirma a existência de uma mensagem com aproximadamente 5,0kb.

Mais uma vez, a análise *in silico* utilizando ESTScan (Iseli *et al.*, 1999) apontou que não existem fases abertas de leitura (ORFs – *Open Reading Frames*) nas mensagens intrônicas estendidas dos *loci* DNAJC3, ITGA6, MYO5A e GAS6. E apesar de um peptídeo com 74 resíduos ter sido predito para o STARD13 intrônico, buscas utilizando essa sequência protéica com BLASTp (Altschul *et al.*, 1997) não resultaram em identificação de proteínas similares no GenBank, nem de domínios

conservados presentes no CDD (Marchler-Bauer *et al.*, 2005). Dessa maneira, é definitivamente improvável que os RNAs intrônicos regulados por andrógeno, especialmente os estendidos por PCR, sejam utilizados para codificar proteínas.

Ainda para o transcrito intrônico de *GAS6*, o tamanho e a orientação transcricional foram confirmados, e a expressão em outros tecidos foi determinada por *Northern blot* utilizando sondas de fita simples geradas por transcrição *in vitro* (detalhes no item 3.6.2). Esse resultado mostra a expressão de uma mensagem longa, sem *splicing* e transcrita antisenso a partir do *locus* de *GAS6* com aproximadamente 5,0kb em diferentes tecidos humanos (Figura 35F). A observação que RNAs intrônicos regulados por andrógeno puderam ser detectados em tecidos que normalmente não sofrem esse tipo de controle hormonal (Figura 35F), sugere que a regulação dos níveis dessas mensagens pode também ocorrer por vias independentes do andrógeno.

4.3.4 Efeitos do andrógeno nos níveis dos RNAs intrônicos e dos transcritos codificantes de proteína correspondentes

Transcrições reversas com orientação específica (OSRT) seguidas de PCR quantitativo em tempo real (detalhes no item 3.7.2), usando RNA de células tratadas e células controle, foram utilizadas para confirmar as alterações na abundância das mensagens intrônicas após a exposição ao andrógeno. Foram testadas duas mensagens intrônicas detectadas apenas com orientação antisenso (*loci GAS6* e *ITGA6*) e duas onde transcritos antisenso e senso sobrepostos foram observados (*loci DNAJC3* e *KDELR2*) (Figuras 36A – D, Painéis “Intrônico”). Para cada *locus*, a abundância de mensagens exônicas codificantes para proteína correspondentes foi também quantificada (Figuras 36A – D, Painéis “Exônico”).

Foi confirmada a diminuição significativa ($p < 0,005$) dos níveis do transcrito intrônico antisenso de *GAS6* após o tratamento com andrógeno (Figura 36A, Painel “Intrônico”). Porém as mensagens codificantes de proteína do *locus GAS6*, que contêm os exons constitutivos 4 e 5 estão significativamente aumentadas ($p < 0,025$) nas mesmas condições experimentais (Figura 36A, Painel “Exônico”). Para o *locus DNAJC3*, foi observado que as abundâncias dos RNAs intrônicos senso e antisenso são moduladas em direções opostas com o tratamento com andrógeno (Figura 36B, Painel “Intrônico”). O aumento observado para o transcrito intrônico antisenso de *DNAJC3* corrobora o perfil observado nos experimentos de *microarray* (Figura 36B, Painel “Intrônico”), sendo observado comportamento similar para o transcrito codificante de proteína de *DNAJC3* (Figura 36B, Painel “Exônico”).

A alteração na abundância dos RNAs intrônicos de *KDELR2* e *ITGA6* também foi confirmada (Figura 36C e D, Painéis “Intrônico”). Para *KDELR2*, tanto os níveis do RNA antisenso, quanto do senso intrônico foram aumentados pelo andrógeno ($p < 0,0035$ e $p < 0,02$, respectivamente) (Figura 36C, Painel “Intrônico”). De maneira similar, o RNA intrônico antisenso do *locus ITGA6* foi significativamente aumentado ($p < 0,0012$) pelo tratamento (Figura 36D, Painel “Intrônico”). Entretanto, para esses últimos *loci* não foram detectadas alterações significativas na abundância dos respectivos transcritos codificantes de proteína, quando oligonucleotídeos iniciadores desenhados para medir exons constitutivos foram utilizados (Figura 36C e D, Painéis “Exônico”).

Figura 36 – Transcrição intrônica antisense pode regular a abundância e o processamento pós-transcricional de mRNAs codificantes de proteína do mesmo *locus*. (A - D) Níveis de expressão de transcritos intrônicos senso e antisense (Painéis “Intrônico”) ou de exons específicos de mensagens codificantes de proteína (Painéis “Exônico”) dos *loci* de *GAS6* (A), *DNAJC3* (B), *KDEL2* (C) ou *ITGA6* (D) medidos a partir de células tratadas com (barras pretas) ou células controle (barras brancas). mRNAs curados completos que mapeiam em cada *loci* são mostrados como caixas e setas pretas. mRNAs não curados e ESTs são mostrados como caixas e setas cinza. Transcritos intrônicos sem *splicing*, regulados por andrógeno, são mostrados como setas hachuradas. *Fold-change* Relativo (eixo Y) indica aumento ou diminuição dos níveis do transcrito após o tratamento com andrógeno em relação aos seus controles. Diferenças de expressão com significância estatística ($p < 0,05$) estão indicadas com asterisco.

Em seguida, experimentos similares aos descritos acima foram desenhados, porém visando testar o possível envolvimento dos longos transcritos intrônicos antisense não-codificantes regulados por andrógeno no mecanismo de *splicing*, assim como já demonstrado anteriormente para um transcrito antisense super-expresso artificialmente (Yan *et al.*, 2005). Foi investigada a existência de correlação entre as abundâncias das mensagens intrônicas e os níveis de alguns mRNAs codificantes de proteína expressos a partir do mesmo *locus*, usando oligonucleotídeos iniciadores desenhados para exons regulados (ausentes em pelo menos uma das isoformas documentadas). Para o *locus KDELR2* foi observado que os exons 2/3 foram significativamente ($p < 0,015$) aumentados pelo tratamento com andrógeno (Figura 36C, Painel “Exônico”). De fato, o exon 2 está ausente no transcrito de *KDELR2* representado pelo mRNA BP269541, sendo possível que o aumento observado nos níveis do RNA intrônico antisense do *locus KDELR2* após o tratamento induza a retenção do exon 2. Não foi observada nenhuma mudança significativa na abundância dos transcritos que contêm os exons 1/3 ou 1/4 (Figura 36C, Painel “Exônico”). Mas surpreendentemente, foi observada diminuição na abundância do exon 1 de *KDELR2* ($p < 0,026$), quando a quantificação foi feita com oligonucleotídeos iniciadores desenhados apenas para esse exon (Figura 36C, Painel “Exônico”). Esse último resultado só pode ser explicado se ocorrer acúmulo após o tratamento com andrógeno de um transcrito ainda não caracterizado de *KDELR2* com a região 3' do exon 1 mais curta.

Para o *locus ITGA6*, a abundância das isoformas codificantes de proteína que contêm um exon terminal mais longo (exon 22D) foi diminuída ($p < 0,04$) pelo tratamento (Figura 36D, Painel “Exônico”), sendo inversamente correlacionada com o aumento observado ($p < 0,0012$) para o RNA intrônico antisense nas mesmas

condições (Figura 36C, Painei “Intrônico”). Entretanto, não é possível determinar se o resultado observado para o transcrito intrônico definitivamente causou a diminuição das isoformas de mRNA de *ITGA6* que contêm o exon 22D, pois foi também detectada diminuição de outra mensagen intrônica, naturalmente transcrita a partir do *locus ITGA6* (BX537483) que ainda possui uma sobreposição com a região 3’ do exon terminal da mensagen codificante de proteína (Figura 36D).

4.3.5 Aumento da ligação de AR a um motivo ARE a 5’ do *asMYO5A*

A regulação da atividade transcricional pelo andrógeno envolve sua ligação direta com o receptor de andrógeno citoplasmático e a translocação do complexo formado para o núcleo, onde interage especificamente com regiões da cromatina, que possuem o motivo de ligação de andrógeno (ARE - *Androgen response elements*), localizados próximos das seqüências alvo (Nelson *et al.*, 2002). Visando agregar evidências de um mecanismo de controle transcricional direto atuando nos RNAs intrônicos regulados pelo andrógeno, foi utilizada uma abordagem bionfórmica baseada na literatura (Nelson *et al.*, 2002), de busca de possíveis motivos AREs localizados na região intrônica, a 5’ da região de mapeamento das mensagens intrônicas antisenso (detalhes no item 3.1.4). Com essa abordagem foram identificados motivos ARE putativos para 24 dos 39 transcritos intrônicos identificados como regulados por andrógeno (Figura 37).

Para demonstrar experimentalmente a ligação do complexo andrógeno/receptor de andrógeno a um dos motivos ARE identificados para os RNAs intrônicos, foi utilizado o ensaio de imunoprecipitação da cromatina (*Chromatin Immunoprecipitation* – ChIP) com anticorpos reativos contra receptor de andrógeno (detalhes no item 3.8.1). Com essa metodologia foi possível constatar

que o aumento de abundância verificado por OSRT-qPCR da mensagem intrônica antisense do *locus MYO5A (asMYO5A)* após o tratamento com andrógeno (Figura 38A), correlaciona-se diretamente com o aumento na detecção da ligação do receptor de andrógeno ao DNA (Figura 38B) em sua possível região promotora com motivo ARE.

GenBank Accession (EST)	Nome do Locus	Seqüência ARE	Posição	Localização Genômica	% Match
CK327134	N A	A G G A C A G A A T G T T T T	-2,056	2p23.2	83
BF333281	NIBP	A G A A G G T G C T G T T C A	-621	8q24.3	75
BF879099	GAS6	G G A A A A T C A T G T T C A	-1,528	13q34	75
AW860958	DST	A G A A C A T C T T T T T G C	-799	6p12.1	75
BF881469	ALMS1	G G A C C A G T T T T T T C T	-1,091	2p13.1	75
CK327189	P2RY14	A G A A C A G C A G G C T C T	-1,837	3q25.1	83
BF368584	DNAJC3	A G A A T A C C A T G T T G C	-1,608	13q32.1	75
BF768459	ACTN4	G G A A C A T G A A T T T C T	-785	19q13.2	75
BF768642	N A	T G A A C T G G G T T T T C T	+170	Xq27.1	75
BF882783	ACTN4	G G A A C C A C T G G T T C T	-1,839	19q13.2	75
CK327094	ATF2	T G A A C A G A C A C T T C T	-920	2q31.1	75
BE087892	STARD13	A G A T C A T G T G G T T C T	-2,239	13q13.1	83
CK327184	ITGA6	A G A A C A G G T T T T T G A	-2,788	2q31.1	75
BF926454	ADD3	T G A C C A T G C T G G T C T	-2,285	10q25.1	75
BF848956	MYO5A	A G T A C A T A A T T T A C T	-812	15q21.2	75
BF805741	PMF1	A G A A A T T C T T G C T C T	-2,104	1q22	75
AW805635	ERG	A G A A C T G C T T T G T C T	+246	21q22.2	75
BF350758	KDELR2	A G A A C A G C T G C T T C T	+219	7p22.1	83
BE156190	N A	A G A C C A G T T T T A T C T	-181	4q32.3	75
BE062809	UBE2V1	A G A A C A A F A T G T T A A	-2,786	20q13.13	83
AW819863	PPP2R2A	A G A G C A G G A A G C T C T	-881	17q23.3	75
BF332494	SAP18	A G A C C A G C C T G G T C A	-414	13q12.11	75
BF364131	N A	T G A C C A G G C T G G T C T	-2,481	16q23.2	75
CK327190	ZNF644	A G A A T C T A C T G T T C C	-283	1p22.2	75

Figura 37 – Identificação *in silico* de motivos ARE em regiões a 5' de RNAs intrônicos regulados por andrógeno. Seqüências de DNA genômico (3 kb) localizado a 5' dos sítios postulados de início de transcrição de cada um dos 39 RNAs intrônicos regulados por andrógeno identificados foram utilizadas em uma busca pelo consenso AGAACAnnnTGTTCT, característico de elementos de resposta a andrógeno (ARE). GenBank Accession esta indicado para os 24 transcritos onde os motivos ARE foram encontrados. Alinhamentos dos 24 AREs identificados foi realizado utilizando CLUSTALW, e os resíduos conservados estão mostrados em preto (**superior**). Uma representação gráfica da seqüência consenso dos 24 AREs, com a freqüência de cada base sendo proporcional ao tamanho do caractere foi gerada utilizando WEBLOGO (Crooks *et al.*, 2004).

Figura 38 – **Aumento na abundância do transcrito intrônico antisense *MYO5A* está correlacionada com a ligação do complexo andrógeno/receptor de andrógeno ao ARE postulado.** (A) Níveis de expressão do transcrito intrônico antisense *MYO5A* foram medidos em células tratadas com andrógeno (barras pretas) ou células controle (barras brancas). *Fold-change* Relativo (eixo Y) indica o aumento ou diminuição dos níveis do transcrito após o tratamento com andrógeno em relação aos seus controles. Diferença de expressão apresentou significância estatística ($p < 0,05$) assim como indicado pelo asterisco. (B) Extratos celulares de culturas tratadas com andrógeno (+) ou controles (–) foram utilizadas em ensaios de ChIP na presença de anticorpos contra a proteína do receptor do andrógeno (anti-AR). Reações de PCR foram utilizadas para detecção do DNA imunoprecipitado contendo o ARE para transcrito intrônico antisense *MYO5A* (*asMYO5A* ARE). O ARE do *Prostate Specific Antigen* ARE (*PSA* ARE), e uma região do gene da tubulina alfa 6 não relacionada com o motivo ARE (*TUBA6*) foram utilizados como controles positivos e negativos respectivamente. Um anticorpo não-relacionado (IgG camundongo) foi utilizado como controle para imunoprecipitação inespecífica. A quantidade inicial de DNA utilizada relativa a cada ensaio de ChIP é mostrada (input).

4.4 Transcritos intrônicos sintênicos com baixa conservação expressos diferencialmente em tecidos humanos e de camundongos.

A observação de RNAs intrônicos com envolvimento em câncer (Reis *et al.*, 2004), assim como de mensagens dessa classe sendo reguladas por estímulos hormonais (Louro *et al.*, 2006)⁵, aumentou ainda mais o interesse no estudo dessas moléculas. Entre as estratégias utilizadas para inferência de funcionalidade para novas seqüências, as mais utilizadas são a avaliação da especificidade de expressão em diferentes tecidos, e as análises de genômica comparativa em

⁵ Louro *et al.* 2006 (Artigo submetido). ANEXO 7 - incluso no CD-ROM suplementar a esta Tese.

diferentes organismos. Assim, utilizamos a análise comparativa entre homem e camundongo como uma abordagem alternativa. Foram obtidos os resultados descritos abaixo, que identificam um conjunto de mensagens intrônicas com perfil de expressão espacialmente conservado em humanos e camundongos, o que sugere funções tecido-específicas desempenhas por esses RNAs; contudo, essas funções ainda necessitam ser determinadas.

4.4.1 Detecção de transcritos intrônicos em tecidos humanos e de camundongos.

A plataforma de *microarray* de cDNA enriquecida em transcritos intrônicos humanos (detalhes da plataforma no item 3.5.1) foi utilizada na identificação de transcritos intrônicos não-codificantes funcionalmente conservados entre humanos e camundongos. A observação que mensagens não-codificantes para proteína, mesmo sem altos níveis de conservação de identidade entre diferentes espécies, podem desempenhar importantes funções celulares (Numata *et al.*, 2003; Ravasi *et al.*, 2006) embasaram o desenho experimental desse estudo (descrito no item 3.5.4). Amostras de RNA de três tecidos diferentes (fígado, rim e próstata) de espécimes humanos e camundongos foram hibridizados independentemente contra a plataforma de *microarray* de cDNA humana.

De todas as sondas intrônicas depositadas na plataforma, foram identificadas 63 (7,7 %) e 21 (2,6 %) expressas em pelo menos um dos tecidos de humanos ou camundongos, respectivamente. A grande maioria (90%) dos transcritos detectados em camundongos foram também detectados como expressos em humanos. De maneira similar, das 2.292 sondas exônicas foram observadas 212 (9.2 %) e 65 (2.8 %) expressas em pelo menos um tecido humano ou de camundongo, com uma

sobreposição de 94 % dos transcritos detectados em camundongo com os detectados em humano.

4.4.2 Expressão diferencial de transcritos intrônicos em tecidos humanos e de camundongos.

Os valores de intensidade normalizados, provenientes dos experimentos realizados independentemente com amostras de tecido de humanos e camundongos foram analisados estatisticamente para identificação de mensagens diferencialmente expressas (detalhes no item 3.1.3). Dos 275 transcritos detectados nos tecidos humanos 227 (82%) foram identificados como diferencialmente expressos por ANOVA, compreendendo 51 transcritos intrônicos e 176 exônicos. Proporções similares foram encontradas nas análises com os tecidos de camundongo, onde 87% (75/86) dos transcritos expressos foram identificados por ANOVA como diferencialmente expressos. Entre esses, um total de 17 provenientes de sondas intrônicas e 58 de exônicas. Importante notar que 100 % das mensagens intrônicas, assim como 93% das exônicas diferencialmente expressas em tecidos de camundongo também foram detectadas como diferencialmente expressas nos tecidos humanos.

Para deixar a abordagem mais estrita foi também realizada uma análise utilizando SAM (Tusher *et al.*, 2001). Com as mensagens intrônicas expressas, a análise por SAM revelou 17 diferencialmente expressas compartilhadas por humanos e camundongos, sendo 14 mais expressas em fígado, duas mais expressas em rim e apenas uma considerada mais expressa em próstata humana e em fígado de camundongo. Similarmente, com as mensagens exônicas, 54 transcritos compartilhados por humanos e camundongos foram detectados como

diferencialmente expressos, sendo 42 mais expressos em fígado, oito mais expressos em rim, um em próstata, e ainda três mais expressos em rim humano e em fígado de camundongo. Mais uma vez, foi possível notar a alta proporção de mensagens (~94%) com padrão de expressão similar nas duas espécies analisadas.

Uma análise de *hierarchical clustering* combinada foi realizada com as listas de transcritos com diferença de expressão, obtidas de maneira independente para cada espécie. Assim, foi possível identificar uma assinatura de expressão tecido-específica independente da espécie analisada, como pode ser observado pelo agrupamento dos perfis de expressão pelos tecidos e não por espécie (Figura 39). A Tabela Suplementar 3 (ANEXO 8) apresenta os transcritos identificados nessa análise.

4.4.3 Validação da existência e orientação transcricional de mensagens intrônicas sintênicas diferencialmente expressas em tecidos humanos e de camundongos.

Para confirmar a existência e determinar a orientação transcricional das mensagens intrônicas cuja expressão preferencial em tecidos parece ser conservada entre as espécies estudadas, foram selecionados 10 transcritos intrônicos diferencialmente expressos em ambas espécies. A sequência conhecida de cDNA humano foi alinhada contra o genoma humano utilizando BLAT (*UCSC Genome Browser*), para obtenção da região sintênica intrônica no genoma de camundongo. Esse conjunto de 10 regiões intrônicas sintênicas foi caracterizado por transcrição reversa com orientação específica seguida de PCR (OSRT-PCR) com RNA extraído do fígado de cinco camundongos. Como resultados preliminares, foram identificados dois transcritos antisense, quatro senso, e quatro pares de transcritos antisense e senso expressos nas regiões intrônicas sintênicas dos 10 *loci* testados.

Com a mesma abordagem, foi determinada a orientação transcricional de 5 dessas mensagens intrônicas humanas, assim como apresentado na figura 40, ao lado do resultado confirmado para as mensagens sintênicas de camundongo. A determinação da orientação transcricional foi realizada com RNA extraído de células da linhagem HepG2, de tumor de fígado humano (Figura 40B).

Figura 39 – **Assinatura de expressão intrônica específica de tecidos e comum entre humanos e camundongos.** *Hierarchical clustering* (Sanchez-Carbayo *et al.*, 2002) dos transcritos intrônicos independentemente identificados como diferencialmente expressos em tecidos humanos e de camundongos. Rótulos coloridos abaixo do painel indicam a espécie (H – humano, M – camundongo), o tecido (próstata, rim ou fígado) e a réplicas (1, 2 ou 3; L – lado esquerdo da lâmina, R – lado direito da lâmina) das hibridizações realizadas. O padrão de expressão de cada mensagem está representado pelo número de desvios-padrão acima (vermelho) ou abaixo (verde) do valor médio da expressão da mensagem em todas as amostras.

Figura 40 – **Determinação da orientação transcricional das mensagens intrônicas expressas diferencialmente em tecidos humanos e de camundongos.** (A e B) Transcrição reversa orientação-específica seguida de PCR (OSRT-PCR) foi realizada em um conjunto selecionado de 5 mensagens intrônicas diferencialmente expressas em tecidos humanos e de camundongos, identificadas pelas análises de *microarray*, para determinação da orientação transcricional da mensagem intrônica relativa a mensagens codificantes para proteína de mesmo *locus*. Nomes a esquerda se referem aos mRNAs RefSeq codificantes para proteína que mapeiam nos mesmos *loci* e definem as regiões intrônicas onde atividade transcricional foi identificada. (A) Produtos de PCR gerados a partir de amostras de fígado de camundongo e separados em gel de agarose 3% detectaram transcrição antisenso (*mMGC*), senso (*Wwtr1*, *Cent2*, *Etf1* e *Rab11*) ou antisenso/senso simultânea (*Sap18*). (B) Produtos de PCR gerados a partir de células da linhagem HepG2 de fígado humano e separados em gel de agarose 3% mostram transcrição antisenso (*SAP18*), senso (*CENT2* e *RAB11*) ou antisenso/senso simultânea (*WWTR1*, *ETF1* e *MGC4707*). Os controles negativos, que avaliam a ausência de contaminação com DNA genômico (–RT) e de auto-anelamento do RNA (–oligo) durante a transcrição, estão indicados.

Surpreendentemente, para as mensagens provenientes dos *loci* humanos foi observada correspondência total na orientação com a observada para os transcritos de camundongo apenas para as mensagens intrônicas senso de *CENT2* e *RAB11* (Figura 40A e B). Apesar de também terem sido detectados transcritos antisenso para os *loci* de *SAP18* e *MGC4707*, um senso correspondente foi apenas observado em humano para o *locus MGC*, e em camundongo apenas para o *Sap18* (Figura 40A e B). E ainda, os pares de mensagens antisenso/senso para *WWTR1* e *ETF1*

observados são exclusivos de fígado humano, pois em camundongo *Wwtr1* e *Etf1* apresentaram apenas mensagens intrônicas senso (Figura 40A e B).

4.5 Interação de transcritos intrônicos com proteínas ligadoras de RNA.

Tendo como hipótese que essa nova classe de mensagens não-codificantes, os Longos RNAs Intrônicos, identificada e extensivamente estudada nesse projeto seja biologicamente relevante e não apenas um tipo de ruído transcricional, foi investigada a existência de associação entre transcritos intrônicos e proteínas ligadoras de RNA (RBPs – *RNA Binding Proteins*). Essas proteínas estão amplamente descritas na literatura como essenciais na regulação pós-transcricional de RNAs, através de processamento do RNA mensageiro imaturo (*splicing*), transporte e estabilidade do mRNA, assim como de sua tradução (Siomi & Dreyfuss, 1997). Assim, os resultados descritos a seguir apesar de ainda necessitarem de experimentos confirmatórios, já apresentam evidências que sustentam o envolvimento dessas mensagens em atividades regulatórias no programa transcricional humano.

4.5.1 Expressão das proteínas ligadoras de RNA em células 293T

Após as transfecções com as construções para expressão episomal de RBPs (ver item 3.2.3), as células 293 T foram coletadas e lisadas para permitir a extração de RNA associado às proteínas ligadoras de RNA em estudo (ver item 3.3.3). Uma alíquota com cerca de 100µl dos lisados foi utilizada para quantificação protéica através da metodologia de Bradford (Harlow & Lane, 1998), assim como para a verificação da expressão das RBPs específicas por *Western Blot*.

Os lisados celulares de 293T produzidos apresentaram tipicamente concentrações de proteína total bastante similares (8-10mg/ml), independente da RBP expressa, não sendo necessário qualquer tipo de normalização precedente as imunoprecipitações. As quantidades similares de proteína podem ser confirmadas na figura 41. Todo o conteúdo das amostras de RNA derivadas das imunoprecipitações (ver item 3.3.3) foi utilizado na geração de alvos fluorescentes marcados (ver detalhes no item 3.5.2), assim como toda a quantidade de alvo fluorescente obtida foi utilizada nas hibridizações (ver item 3.5.4).

Figura 41 – **Expressão transiente das proteínas ligadoras de RNA (RBPs).** Western Blot para detecção da expressão das RBP em 293T. Um volume de lisado celular referente à cerca de 50µg de proteína total de cada proteína expressa (RBPs e o controle GFP) separado por SDS-PAGE, 10% de acrilamida foi transferido para uma membrana de nitrocelulose, para detecção da proteína de fusão com anti-proteína A. **Poço 1:** HuD (40kD); **Poço 2:** GFP (27kD); **Poço 3:** EIF4E (24kD); **Poço 4:** PABP (72kD); **Poço 5:** h-TRA2 (40kD). Massas moleculares indicadas referentes às proteínas com o tag de proteína A.

4.5.2 Identificação, utilizando *microarrays* de cDNA, de RNAs longos intrônicos não-codificantes associados a proteínas ligadoras de RNA

Alvos fluorescentes imunoprecipitados e marcados foram hibridizados com *microarrays* intrônicos (ver item 3.5.4). Após as análises para determinação das

mensagens especificamente co-imunoprecipitadas (ver item 3.5.5) foi possível detectar 296, 201, 175 e 122 transcritos que foram imunoprecipitados por estarem presentes nas células superexpressando PABP, HUD, hTRA2 e EIF4E, respectivamente.

De modo a tentar caracterizar o conjunto de transcritos intrônicos candidatos a ligação com RBP, foi feita uma análise de enriquecimento de genes pertencentes a uma determinada categoria de ontologia gênica (GO), utilizando o programa eGOn v1.0 (Explore Gene Ontology - <http://www.genetools.no/>). Algumas categorias apresentaram-se como significativamente enriquecidas em comparação com todos os transcritos intrônicos presentes em nossa plataforma de *microarrays* (Tabela Suplementar 4 – ANEXO 9).

4.6 Análise de expressão de RNAs intrônicos em células LNCaP tratadas com α -amanitina

Para ampliar o estudo da biogênese dos transcritos intrônicos iniciado com a identificação de RNAs regulados por andrógeno (Louro *et al.*, 2006), foram realizados experimentos de análise de expressão gênica com células de próstata tumoral LNCaP tratadas com α -amanitina. Essa molécula é um octapeptídeo bicíclico extraído do cogumelo *Amanita phalloides* que inibe fortemente a atividade da RNA polimerase II (RNAP II) impedindo sua translocação sobre o DNA durante a elongação da transcrição (Bushnell *et al.*, 2002). Como tem sido apenas postulado o envolvimento da RNAP II na transcrição de muitos RNAs não codificantes, esse estudo visou testar se essa é a polimerase responsável pela síntese da maioria dos longos RNAs intrônicos que vêm sendo identificados pelas análises descritas.

Um conjunto de experimentos piloto (não apresentados aqui) utilizou a plataforma de *microarrays* de cDNA, tendo sido hibridizadas duas amostras tratadas e duas controle; uma parcela considerável dos transcritos intrônicos cuja expressão foi detectada em LNCaP, apresentaram sua expressão significativamente aumentada na células tratadas com o inibidor da RNAP II. Esse inesperado aumento na abundância dos transcritos intrônicos está em concordância com resultado observado na literatura para algumas mensagens codificadoras de proteínas humanas (Kravchenko *et al.*, 2005), e apesar de sugerir várias hipóteses, ainda necessita de experimentação para uma melhor compreensão do fenômeno. Entretanto, deixa claro que a RNAP II não está envolvida na transcrição das mensagens intrônicas onde esse fenômeno foi observado.

4.6.1 Uma grande fração de RNAs intrônicos aparentemente não é transcrita pela RNAP II

Para ampliar e validar as observações iniciais obtidas com a plataforma de cDNA onde estão representados 822 mensagens intrônicas, foram realizadas novas hibridizações com RNA de células LNCaP tratadas com α -amanitina, agora usando uma plataforma de *microarrays* de oligonucleotídeos desenhada pelo grupo (descrita no item 3.11.1) que possui 13.004 sondas intrônicas.

Como pode ser observado na figura 42, cerca de 39 % (3.909) das mensagens exônicas expressas foram significativamente afetadas pela inibição da RNAP II enquanto o restante, provavelmente mRNAs mais estáveis, não tiveram sua abundância significativamente alterada. Como esperado, a grande maioria das mensagens codificantes para proteínas (exônicas) alteradas pelo tratamento (96 %) tiveram seus níveis diminuídos, enquanto apenas 4 % das mensagens codificantes

para proteínas foram aumentadas (Figura 42A), resultado similar ao já observado em células HeLa (Kravchenko *et al.*, 2005).

Figura 42 – **Efeito do tratamento com α -amanitina sobre a expressão de transcritos intrônicos e exônicos.** Alteração na expressão após tratamento com α -amanitina (azul escuro) de 3.909 transcritos codificadores de proteínas (exônicos) (A), 400 transcritos totalmente intrônicos antisense (asTI) (B), 468 transcritos totalmente intrônicos senso (sTI) (C), e 522 transcritos parcialmente intrônicos antisense (asPI) (D). Todos os transcritos cujos padrões de expressão estão representados foram identificados como diferencialmente expressos pelas análises estatísticas combinadas, descritas no texto, em amostras tratadas com α -amanitina (azul escuro) em relação ao controle (veículo etanol, azul claro). Cada linha representa um transcrito, e dentro de cada amostra (separadas pelas barras coloridas indicadas) as colunas indicam as réplicas.

Frações bastante menores de transcritos expressos totalmente intrônicos antisenso (asTI, 16 %) e senso (sTI, 18 %), e transcritos parcialmente intrônicos antisenso (asPI, 19%) foram afetados pela α -amanitina (Figura 42B-D). Surpreendentemente, entre os transcritos cujos níveis foram alterados nas células tratadas, 16 % (62/400), 13 % (59/468) e 12 % (64/522) de asTI, sTI e asPI foram, respectivamente, aumentados. A Tabela Suplementar 5 (ANEXO 10) apresenta todos os transcritos cuja expressão foi afetada pelo tratamento com α -amanitina.

4.6 Sistema para modulação da expressão de *asRASSF1* e *asPPHLN1* em células humanas

A determinação das funções celulares específicas desempenhadas pelos Longos RNAs intrônicos não-codificantes identificados pelas diferentes abordagens descritas até aqui é o próximo passo na caracterização dessa classe de ncRNAs. Utilizando a estratégia descrita no item 3.9.6 já foi construído um sistema de expressão para os Longos RNAs intrônicos não-codificantes correlacionados com o grau de diferenciação tumoral de próstata *asRASSF1* e *asPPHLN1* (Reis *et al.*, 2004). Esse sistema, esquematizado na figura 14, será utilizado para superexpressar os transcritos intrônicos em estudo, e observar as alterações celulares desencadeadas.

5 DISCUSSÃO

5.1 Identificação e caracterização parcial de *RASL11A* e *ARL-9*, membros representativos de novas famílias de pequenas GTPases monoméricas.

Analisando os resultados obtidos com os alinhamentos das proteínas Rasl11 e Arl-9/Arl10 com alguns membros representantes da superfamília de pequenas GTPases monoméricas humanas e com seus ortólogos, foi possível notar que se tratam de novas subfamílias que apesar de apresentarem as principais características da superfamília, possuem certas peculiaridades. Essas características diferenciais que permitem sua separação em ramos distintos estão mantidas em muitos eucariotos, sugerindo que essas proteínas são membros de famílias gênicas evolutivamente conservadas.

A análise da seqüência do produto do gene *ARL-9*, descrito aqui, mostra que o domínio G, incluindo o *Interswitch Toggle* típico das proteínas Arf, está apenas representado de forma completa se for utilizado como iniciador de tradução um códon alternativo não-AUG. O uso de códons de iniciação alternativos, no caso CUG, já foi documentado para genes humanos funcionais, incluindo proto-oncogenes, fatores de transcrição e fatores de proliferação (Boeck & Kolakofsky, 1994; Tee & Jaffe, 2001). Em todos os casos estudados, esses códons aparecem em um contexto similar, o qual é caracterizado pela presença de uma purina na posição -3 e uma guanossina (mais raramente uma citosina) na posição +4 (Boeck & Kolakofsky, 1994). Essas características de seqüência estão também presentes na ORF que se inicia com o códon alternativo CUG no transcrito de *ARL-9*. Contudo, a confirmação de seu uso, assim como sua significância biológica só pode ser verificada através de experimentação bioquímica mais aprofundada.

A proteína Rasl11a, identificada, clonada e parcialmente caracterizada durante esse projeto apresenta na região do *switch* 1 do domínio G um resíduo de Asn na posição 59 de sua seqüência primária, ao invés de um resíduo de Thr. Essa modificação é observada também na proteína Rasl11b, em todos os ortólogos de Rasl11, mas não é restrito a essa subfamília, visto que Ris, uma família recentemente descrita de proteínas relacionadas à Ras, apresenta a mesma substituição (Wistow *et al.*, 2002). Na proteína H-Ras, membro prototípico da superfamília, o resíduo Thr35 no *switch* 1, em conjunto com a Gly60 no *switch* 2, contribuem para a estabilização da forma ligada a GTP através de pontes de hidrogênio entre os NH da cadeia principal e o oxigênio do fosfato γ do GTP (Vetter & Wittinghofer, 2001a). É possível que essa substituição de Thr para Asn nas proteínas Rasl11, preservando o caráter polar não carregado da cadeia lateral, não interfira significativamente com a estabilização do GTP e nem com a atividade GTPásica das proteínas. Estudos com mutantes de proteínas Ras revelaram que uma troca do resíduo Thr35 por um resíduo de Ser, nas proteínas Ras, reduz a atividade GTPásica intrínseca, mas mantém a capacidade de ligação com alta afinidade aos seus efetores (Spoerner *et al.*, 2001). Contudo, apenas uma caracterização bioquímica detalhada, inclusive dos efetores ainda desconhecidos das proteínas Rasl11, pode determinar as conseqüências funcionais da substituição de Thr para Asn observada.

Pequenas GTPases relacionadas a Ras geralmente apresentam na região carboxi-terminal seqüências que sinalizam para modificações que permitem sua associação com a membrana plasmática. Essa associação é freqüentemente indispensável para sua funcionalidade biológica. As proteínas Rasl11 não apresentam nenhum sinal já descrito de reconhecimento para adição de lipídeos na

região C-terminal, o que é encontrado nas proteínas Ras (Reuther & Der, 2000a), tais como sinais de farnesilação ou palmitoilação (Dai *et al.*, 1998; Sinensky, 2000). Entretanto, uma assinatura conservada de resíduos polibásicos foi encontrada, nas análises aqui descritas, em todos os ortólogos de vertebrados de Rasl11, sugerindo que essas proteínas apresentam um mecanismo de ligação à membrana plasmática baseado em resíduos carregados semelhante ao postulado para Rin e Rit, membros da família Ras humana (Lee *et al.*, 1996; Sinensky, 2000).

O tamanho de um dos transcritos observados por *Northern Blot* (~1,6kb) para *RASL11A* apresenta concordância com a sequência completa obtida por RACE-PCR (1.543bp). A outra forma de menor tamanho (~1,2kb) poderia apenas ser explicada por iniciação ou termino alternativo de transcrição, o que ainda carece de confirmação experimental.

As análises quantitativas de PCR em tempo real revelaram que o transcrito de *RASL11A* está em quantidade significativamente menor em tumores de próstata, quando comparado com sua abundância em amostras de tecido prostático não-tumoraladjacente ao tumor. Esses resultados sugerem para esse novo membro humano da superfamília de pequenas GTPases monoméricas um papel de supressor tumoral. Existem diversos trabalhos na literatura que mostram que pequenas GTPases podem agir como supressoras tumorais. Perda de heterozigose no gene *NOEY/AHR1* foi detectada em 41% dos cânceres de ovário e mama (Yu *et al.*, 1999b). Ainda nesse mesmo trabalho foi verificado que a re-expressão de epissomal Noey inibe a proliferação das células tumorais de ovário e mama, indicando que esse membro da família Ras tem atividade supressora tumoral. Superexpressão de Rig é capaz de inibir a transformação celular mediada por Ras em células NIH 3T3 (Ellis *et al.*, 2002b). A co-expressão de Rap1a é capaz de

reverter a instabilidade genômica induzida por H-Ras oncogênica (Wani *et al.*, 1997a). E recentemente foi demonstrada que a superexpressão de Rerg em células de carcinoma de mama inibe sua proliferação *in vitro* e também a formação de tumor em camundongos *nude* (Finlin *et al.*, 2001). O fato de Rasl11a ser bastante similar a Rerg aumenta as evidências do papel supressor tumoral do gene *RASL11A* em câncer de próstata.

Finalmente, o conjunto destes resultados aponta para uma nova subfamília de pequenas GTPases monoméricas, Rasl11, que pode apresentar um importante papel na supressão de tumores. Dessa forma a geração de anticorpos contra as proteínas dessa subfamília, assim como uma caracterização bioquímica de seu mecanismo de ativação, iniciadas durante esse projeto com a obtenção de expressão heteróloga em um sistema em bactérias, devem contribuir no entendimento das funções dessas proteínas durante o surgimento de tumores, assim como na fisiologia de células normais.

5.2 Transcritos intrônicos marcadores de prognóstico em câncer de próstata.

Os resultados obtidos no estudo de expressão gênica em câncer de próstata por *microarrays* de cDNA sugerem que a maioria das regiões intrônicas de genes codificantes de proteína onde estão mapeadas ESTs, ao longo de todo o genoma humano, são de fato transcricionalmente ativas, assim como as observadas em estudos específicos com os cromossomos 21 e 22 (Kapranov *et al.*, 2002; Rinn *et al.*, 2003; Kampa *et al.*, 2004). A observação de longos RNAs antisense sem *splicing* nessa parte de nosso projeto é inteiramente concordante com resultados prévios de outros grupos, que mostram mensagens de 0,4–1,3 kb antisense sem *splicing* e totalmente intrônicas transcritas a partir do cromossomo 22 (Rinn *et al.*, 2003).

Nas amostras de RNA poli-(A)+ foram detectados simultaneamente pares de transcritos intrônicos antisenso e senso sobrepostos que são correlacionados com o grau de diferenciação tumoral em seis diferentes *loci*. Essas mensagens antisenso totalmente intrônicas provavelmente tem como alvo a seqüência senso intrônica complementar presente no pré-mRNA. Dessa forma podem desempenhar suas funções regulatórias de maneira seqüência-específica no processamento das mensagens codificantes de proteína correspondentes.

Pares de transcritos senso/antisenso, assim como mensagens antisenso totalmente intrônicas, têm sido detectados também em camundongos através de análises de mapeamento em escala genômica (Kiyosawa *et al.*, 2003). Esses transcritos têm sido implicados nos fenômenos de *imprinting* e inativação do cromossomo X em mamíferos (Lee *et al.*, 1999; Sleutels *et al.*, 2000). RNAs antisenso também têm sido demonstrados como mediadores do silenciamento gênico e metilação na doença autossômica alfa-talassemia (Tufarelli *et al.*, 2003). Ainda mais importante, trabalhos antigos da literatura mostram que alguns transcritos antisenso intrônicos agem como efetores negativos de genes relacionados com câncer, tais como *cMyc* (Nepveu & Marcu, 1986) e *p53* (Khochbin & Lawrence, 1989), e como efetores positivos de *RPS14* (Tasheva & Roufa, 1995). Assim, esses pares de transcritos antisenso e senso devem fazer parte de um mecanismo geral de regulação da expressão gênica em eucariotos ainda não completamente compreendido.

De fato, através de uma abordagem de mapeamento de sítios para ligação de fatores de transcrição (TFBS), sem viés para regiões codificantes para proteína, nos cromossomos 21 e 22 foi revelado que 36% dos TFBS para Sp1, *cMyc* e *p53* são

significativamente correlacionados com RNAs não-codificantes, e os pares de mensagens geralmente são co-regulados (Cawley *et al.*, 2004).

Tanto os níveis do transcrito intrônico antisense, quanto os do senso de *RASSF1* estão significativamente diminuídos em tumores com alto GS. Ainda, foi observado que a fração de mensagens correlacionadas com o grau de diferenciação tumoral com relação a todas expressas em próstata, é consideravelmente maior de transcritos intrônicos (6.6%, [23/346]), quando comparada à fração de transcritos que mapeiam em exons de *RefSeq* (2%, [27/1,369]). Assim, da mesma forma que evidências apontando para uma maior complexidade do programa transcricional humano continuam sendo acumuladas, também se acumulam as perguntas a respeito da relevância biológica dos RNAs não-codificantes, que aparentemente respondem pela grande maioria dos transcritos. A correlação com o estágio de progressão tumoral observada nessa parte do projeto é claramente um forte argumento contra a hipótese de simples ruído transcricional. Esses resultados apresentados também apontam fortemente para a necessidade de inclusão de sondas para RNAs não-codificantes, principalmente os expressos a partir de regiões intrônicas, na busca de metodologias alternativas para diagnóstico e prognóstico molecular de câncer, até então exclusivamente focadas em produtos de origem protéica.

Finalmente, um paradigma corrente em oncogênese aponta diretamente para o acúmulo de mutações em fases abertas de leitura de mensagens codificantes para proteínas, que possuem tanto comportamento de oncoproteínas quanto de supressoras tumorais, como causas do surgimento e desenvolvimento de câncer (Mcmanus, 2003). O envolvimento de vários tipos de RNAs não-codificantes, como os microRNAs (Mcmanus, 2003; Calin *et al.*, 2004), em câncer vem se tornando

mais evidente. Assim a identificação de transcritos intrônicos correlacionados com o surgimento ou desenvolvimento de neoplasias, apresentado aqui pela primeira vez na literatura, pode e deve aumentar definitivamente a complexidade desse paradigma.

5.3 Transcritos intrônicos regulados pelo estímulo de andrógeno.

O resultado obtido a partir das análises de GO das mensagens que responderam ao tratamento é mais uma observação altamente sugestiva que os RNAs intrônicos não são um mero ruído transcricional. Na verdade, aponta para possíveis funções desempenhas por essas mensagens durante a expressão gênica desencadeada com a exposição celular ao andrógeno. Essas mensagens provavelmente agem na regulação pós-transcricional de genes envolvidos na transdução de sinal em células de próstata. E de fato, alguns desses *loci* já haviam sido demonstrados como envolvidos em eventos celulares relacionados ao crescimento e diferenciação de células da próstata. Por exemplo, a proteína codificada pelo mRNA GAS6 é um ligante do receptor tirosina quinase Axl, e o complexo Gas6/Axl apresenta atividade mitogênica em linhagens celulares humanas de câncer de próstata, modulando a via de transdução de sinal de PI3K/AKT e MEK (Sainaghi *et al.*, 2005). De maneira similar, o fator de transcrição codificado por ERG, que é membro da família ETS de genes responsáveis pela integração dos sinais que regulam o crescimento e a diferenciação celular, respostas ao estresse, e tumorigênese, foi anteriormente identificado como o proto-oncogene cujos níveis encontram-se aumentados com maior frequência em células epiteliais malignas de próstata (Petrovics *et al.*, 2005).

O envolvimento de RNAs não-codificantes em eventos de processamento pós-transcricional em moléculas de mRNA codificantes para proteína já tem sido demonstrado. *HBII-52*, um pequeno RNA nucleolar (snoRNA) foi descrito como um elemento que atua em *trans* na regulação do processamento do transcrito 5-*HT2CRA*, codificante para o receptor de serotonina (Kishore & Stamm, 2006). Alteração no processamento ocorre devido à ligação do *HBII-52* a um elemento de silenciamento do exon Vb presente no pré-mRNA de 5-*HT2CRA*, gerando uma mensagem defeituosa induzida pelo snoRNA, causando a síndrome congênita de Prader-Willi (Kishore & Stamm, 2006).

Em outro trabalho, Yan e seus colaboradores identificaram no *locus FAS* um longo RNA antisenso parcialmente intrônico que foi denominado *SAF*, naturalmente transcrito a partir da região genômica relativa ao primeiro intron do mRNA senso codificante para a proteína Fas (Yan *et al.*, 2005). Nesse estudo, os autores mostraram que a superexpressão de *SAF* pode induzir grandes alterações no processamento dos exons 3' do mRNA de *FAS*, que passa a codificar para uma proteína sem a propriedade de ficar ancorada à membrana plasmática, onde normalmente realiza suas funções (Yan *et al.*, 2005). Assim, células com expressão de *SAF* aumentada deixam de induzir apoptose mediada por Fas (Yan *et al.*, 2005). Os resultados obtidos nessa parte do projeto apontaram para novos RNAs antisenso intrônicos, que de maneira análoga a *SAF*, podem estar envolvidos em mecanismos regulatórios, ainda não totalmente elucidados. Aparentemente as mensagens identificadas atuam na modulação do processamento de pré-mRNAs de mensagens codificantes que apresentam uma complexa utilização de exons, influenciando na função desempenhada pelo produto proteico.

Os resultados apresentados aqui fornecem evidências originais que RNAs não-codificantes transcritos a partir de regiões intrônicas devem ser regulados por sinais fisiológicos que normalmente atuam sobre os transcritos codificantes de proteína, como os hormônios. A existência de um número considerável de sítios para ligação de fatores de transcrição (TFBS – *Transcription Factor Binding Site*) como Sp1, c-Myc, p53 e Creb, posteriores a regiões transcritionalmente ativas, e portanto fora da região tradicionalmente considerada como região promotora, foi demonstrada (Cawley *et al.*, 2004; Euskirchen *et al.*, 2004). Alguns desses TFBS estão orientados de maneira a possibilitar a geração de transcritos ainda não conhecidos e com orientação antisenso em relação às mensagens já estudadas (Cawley *et al.*, 2004). Cawley e colaboradores ainda observaram que frações similares de mensagens codificantes de proteína e RNAs não-codificantes, com TFBS experimentalmente determinados, são reguladas após tratamento com ácido retinóico (Cawley *et al.*, 2004). Ainda, essas observações são reforçadas pela identificação de longos RNAs não-codificantes de camundongos que respondem ao estímulo com lipopolissacarídeos (Ravasi *et al.*, 2006).

O mRNA de *MYO5A* codifica para uma proteína envolvida no tráfego intracelular de vesículas e exocitose (Watanabe *et al.*, 2005). Interessante notar nesse contexto, que isoformas de *Myo5a* são tecido-específicas e originadas a partir de uso alternativo de exons, conferindo a especificidade de cargas transportadas por essa proteína no meio intracelular (Au & Huang, 2002). Com isso, é perfeitamente concebível especular que o aumento nos níveis de *asMYO5A* em células de próstata em resposta direta ao estímulo do andrógeno, afete a abundância relativa das diferentes formas do mRNA que codifica para *Myo5a*, modulando finalmente a atividade e especificidade dessa proteína.

Apesar de não existirem evidências conclusivas a respeito dos mecanismos moleculares, é postulado que os transcritos intrônicos realizem suas funções atuando em *cis* através de pareamento de bases do tipo Watson-Crick ao pré-mRNA senso complementar (Reis *et al.*, 2005). Os resultados obtidos nessa etapa do projeto estão de acordo com conceito de que pelo menos uma fração dos transcritos intrônicos deve afetar diretamente a abundância ou o processamento de mensagens codificantes de proteína originadas da fita oposta, assim como já observado por outros autores (Lee *et al.*, 1999; Katayama *et al.*, 2005). Entretanto, não pode ser excluída a possibilidade que alguns desses longos RNAs não-codificantes possam modular processos transcricionais, pós-transcricionais e até nos produtos protéicos de outros *loci*. Para isso poderiam tanto assumir estruturas secundárias complexas, ser parcialmente complementares às seqüências alvo, quanto ser processados em seqüências funcionais menores, similares aos microRNAs (Bartel, 2004). Nesse caso, a identificação dos genes afetados seria a maior limitação na caracterização funcional de novos RNAs não-codificantes biologicamente relevantes.

Finalmente, esse conjunto completo de resultados sugere fortemente que ao menos uma parte dos RNAs intrônicos de células de mamíferos são transcricionalmente controlados e devem exercer funções regulatórias, corroborando a visão da existência de um sistema complexo, mas ainda pouco compreendido, de regulação baseado em RNAs não-codificantes em organismos superiores (Johnson *et al.*, 2005; Mattick & Makunin, 2006; Willingham & Gingeras, 2006).

5.4 Transcritos intrônicos diferencialmente expressos em tecidos humanos e de camundongos.

Hibridizações cruzadas entre espécies, similares às realizadas nessa etapa do projeto, também têm sido utilizadas em análises de expressão com diversas motivações, que vão desde ausência de plataformas de *microarrays* específicas para a espécie em estudo (Moody *et al.*, 2002; Tsoi *et al.*, 2003); para possibilitar comparações dos padrões de expressão de um conjunto de transcritos em múltiplas espécies (Enard *et al.*, 2002; Adjaye *et al.*, 2004); e até para possibilitar a identificação de genes conservados com a observação de sua expressão em diferentes estágios de desenvolvimento (Vallee *et al.*, 2006). Entretanto esse tipo de abordagem necessita de cautela, pois mesmo pequenas divergências na sequência de nucleotídeos entre as mensagens afetam diretamente seus níveis de intensidade e a especificidade da hibridização em comparações entre espécies na mesma plataforma (Gilad *et al.*, 2005).

Assim, no estudo mostrado aqui para evitar esse tipo de variação não controlável, as hibridizações de cada espécie foram realizadas em lâminas separadas. Da mesma forma, as comparações apresentadas foram apenas realizadas depois das análises estatísticas serem finalizadas independentemente em cada espécie. Ainda, a observação que todas as mensagens intrônicas expressas em camundongo foram também detectadas nos tecidos humanos correspondentes, sugere fortemente que com nosso desenho experimental e com a abordagem bastante estrita de análise estamos apontando apenas para mensagens hibridizadas com uma relativa especificidade.

Os 17 transcritos intrônicos compartilhados entre humanos e camundongos, detectados como diferencialmente expressos em fígado, rim e próstata, têm seus

perfis de expressão agrupados por tecido, revelando uma assinatura tecidual comum de transcrição intrônica entre as duas espécies. Esse resultado sugere que a expressão intrônica não é randômica, similar à expressão tecido-específica já observada para outros RNAs não-codificantes, os microRNAs (Lu *et al.*, 2005). Do conjunto de mensagens intrônicas aparentemente sintênicas identificados aqui, 10 foram confirmadas quanto a sua existência na população de RNA, e tiveram sua orientação transcricional determinada em amostras provenientes de fígado de camundongo. A determinação da orientação transcricional em paralelo de 5 dessas 10 mensagens em amostras humanas mostrou uma correspondência apenas parcial, sendo observada uma maior porcentagem (apesar do número baixo de *loci* testados) de transcrição antisense.

Assim, especula-se que talvez tanto os transcritos intrônicos senso quanto os antisense poderiam representar papéis distintos na regulação do programa transcricional de mamíferos. Sendo, provavelmente as mensagens antisense as determinantes da maior complexidade transcricional humana, através de um mecanismo de regulação ainda mais fino, que modularia os efeitos desempenhados pelas mensagens senso no controle da expressão gênica. Esse mecanismo foi postulado pela observação de um outro estudo atual onde é mostrado que a superexpressão de certos introns (região intrônica com mesma orientação transcricional dos exons de mesmo *locus*) em linhagens celulares onde a expressão da mensagem codificante de proteína correspondente não era detectável causa alteração na expressão de alguns mRNAs de maneira altamente reprodutível e específica para cada intron (Hill *et al.*, 2006). Esse resultado indica claramente uma funcionalidade para seqüências intrônicas senso não totalmente relacionada às mensagens codificantes correspondentes, talvez como proposto pelos autores,

atuando como sinais de notificação para a transcrição de outros mRNAs que participam das mesmas vias metabólicas, ou de vias relacionadas (Hill *et al.*, 2006).

Por outro lado, até o momento apenas uma clássica exceção (*Xist*) têm sido utilizada para indicar que uma baixa conservação de identidade não implica necessariamente em ausência de conservação de funcionalidade (Numata *et al.*, 2003; Pang *et al.*, 2006; Ravasi *et al.*, 2006). Casos como esse geralmente são apenas identificados acidentalmente, pois metodologias tradicionais de análise *in silico* por definição não apontam para seqüências com baixa identidade que sejam funcionalmente preservadas durante a evolução. Os resultados obtidos nessa etapa do projeto mostram uma conservação, entre duas espécies, de expressão tecido-específica para um número de RNAs intrônicos com baixa conservação de identidade, sugerindo que a expressão funcional desse tipo de mensagem não-codificante é mais comum do que o inicialmente esperado. Pistas para explicar essa observação podem ser obtidas do trabalho de Torarinsson e colaboradores onde foi mostrado que milhares de regiões genômicas sintênicas de humanos e camundongos, mas não alinháveis quanto a estrutura primária, podem conter RNAs que possuem conservação de estrutura secundária (Torarinsson *et al.*, 2006). Isso aponta fortemente para a existência de um fenômeno já observado em proteínas, a conservação de estrutura para realização de funções similares, também em moléculas de RNA.

Finalmente, sob a luz dessas observações, especula-se sobre o papel dos transcritos intrônicos antisense, identificados aqui em maior proporção em humanos em relação a camundongos, adicionando ainda mais complexidade ao modelo de regulação intrônica proposto (Hill *et al.*, 2006). Enquanto a sinalização intrônica senso ativaria as redes de genes relacionadas, os transcritos antisense poderiam

exercer um papel de modulação negativa das funções desempenhadas por cada intron (senso) de uma maneira espacialmente (tecido), e talvez temporalmente, bastante controlada. Assim, alguns transcritos antisenso poderiam ter um papel indireto na regulação da expressão gênica, tanto capturando a sinalização intrônica e a redirecionando para as diferentes vias, quanto através de impedimento para que essa seqüência adquirira a estrutura secundária correta que permite a realização de suas funções celulares.

5.5 Interação de transcritos intrônicos com proteínas ligadoras de RNA.

Os resultados obtidos nos ensaios de interação de mensagens intrônicas com proteínas ligadoras de RNA (RBPs) envolvidas em reconhecidos processos regulação fortalecem nossas hipóteses de funcionalidade para essa classe de transcritos. A existência de transcritos intrônicos candidatos a ligação funcional com PABP, EIF4E, hTRA2 e HUD, aumentam a confiabilidade do conjunto de resultados com RNAs intrônicos obtidos nesse projeto como um todo, e ainda encorajam a caracterização funcional de alguns membros dessa peculiar classe de RNAs.

Ainda, o enriquecimento significativo ($p < 0,05$) de transcritos intrônicos imunoprecipitados com PABP, EIF4E e HUD na categoria *regulation of progression through cell cycle*, assim como na categoria *signal transducer activity* com PABP e HUD são novas evidências providas pela análise de ontologia gênica (GO) que também se somam às já discutidas sobre as funções regulatórias dos Longos RNAs Intrônicos não-codificantes.

5.6 Análise de expressão de RNAs intrônicos em células LNCaP tratadas com α -amanitina

Apesar da enorme quantidade de resultados acumulados que demonstram que a maioria dos transcritos humanos são RNAs não-codificantes de proteína expressos a partir de regiões intrônicas e intergênicas do genoma (Willingham & Gingeras, 2006), até o momento não existem trabalhos que exploraram sistematicamente aos mecanismos de biogênese dessas mensagens. Na verdade, tem se assumido que os longos RNAs não-codificantes são em sua grande maioria transcritos pela RNAP II pois esses compartilham certas características estruturais em comum com os mRNAs, como poliadenilação e *capping* 5' (Erdmann *et al.*, 2001; Numata *et al.*, 2003). Ainda, a observação da ligação específica do receptor de andrógeno a regiões promotoras intrônicas (Louro *et al.*, 2006), assim como a identificação de sítios de ligação para conhecidos fatores de transcrição por RNAP II localizados em regiões intrônicas e intergências ativas (Martone *et al.*, 2003; Cawley *et al.*, 2004; Euskirchen *et al.*, 2004; Kim *et al.*, 2005), podem ser utilizadas como indícios para expressão pela RNAP II dessa classe de mensagens.

Até pouco tempo apenas três RNA polimerases nucleares eram conhecidas, sendo bem diferenciadas por suas respostas ao tratamento com α -amanitina, que não tem efeito na atividade da RNAP I, inibe fortemente a RNAP II, e levemente a RNAP III (Jacob *et al.*, 1970; Kedingner *et al.*, 1970; Lindell *et al.*, 1970). Entretanto uma quarta RNAP nuclear foi recentemente descrita (spRNAP IV), e surpreendentemente foi observado que tratamento com α -amanitina aumenta a expressão de mRNAs codificantes de proteína transcritos por ela (Kravchenko *et al.*, 2005).

Nessa etapa do projeto foi avaliada pela primeira vez a contribuição da RNAP II na expressão dos RNAs não codificantes intrônicos em células humanas, bloqueando a sua atividade com a α -amanitina e medindo os efeitos na expressão gênica com a plataforma de *microarray* de oligonucleotídeos com cerca de 13.000 sondas intrônicas. Constatamos que além de uma menor porcentagem de transcritos intrônicos (16-19%) serem sensíveis ao tratamento com α -amanitina, quando comparada com a dos exônicos (39%), uma fração significativamente maior dos transcritos intrônicos (12–16 %), em relação aos exônicos afetados pelo tratamento (4%), tiveram seus níveis aumentados. Esse comportamento observado para os transcritos intrônicos nas células tratadas com α -amanitina levanta a possibilidade de que um outro sistema de síntese esteja atuando na transcrição de muitos RNAs não codificantes. Apesar da spRNAP IV, que aparentemente é ativada pela α -amanitina (Kravchenko *et al.*, 2005), ser uma ótima candidata, a participação direta da spRNAP IV na ativação da transcrição intrônica aqui observada, ainda necessita de confirmação experimental.

6 CONCLUSÕES E CONSIDERAÇÕES FINAIS

O conjunto de resultados originais obtidos durante esse doutoramento ilustram a complexidade e contribuem para desvendar o repertório transcricional humano completo. Novos transcritos foram identificados e parcialmente caracterizados tanto em um contexto celular normal, quanto no surgimento e desenvolvimento de neoplasias de próstata.

Com a abordagem de busca de novos membros parálogos de famílias gênicas envolvidas com câncer foram identificadas duas novas subfamílias, Rasl11 e Arl9/10, da superfamília de pequenas GTPases monoméricas. As proteínas aparentemente codificadas por essas mensagens são conservadas entre os eucariotos e apresentam certas peculiaridades na sequência primária. A observação que os níveis do mRNA de *RASL11A* encontram-se significativamente diminuídos em amostras tumorais de próstata aponta para possíveis papéis supressores de tumor também para esse novo membro da superfamília. Futuros estudos funcionais, principalmente com as proteínas Rasl11, tornam-se muito interessantes para o entendimento do processo tumoral em próstata, e para validação da abordagem de busca por proteínas parálogas como estratégia para reconhecimento das possíveis vias bioquímicas alternativas, que eventualmente estejam dificultando o tratamento efetivo do câncer.

A estratégia de observar o perfil de expressão de amostras de próstata normais e tumorais, utilizando uma plataforma de *microarray* enriquecida com mensagens ainda não anotadas, acabou revelando a existência de uma nova classe de RNAs não-codificantes. Esses transcritos são longos, não sofrem *splicing*, são expressos a partir dos introns e freqüentemente com orientação oposta (antisense) ao RNA mensageiro (mRNA) de genes codificantes para proteína. O padrão de

transcrição observado para estes Longos RNAs Intrônicos não-codificantes em amostras de câncer de próstata correlaciona-se com o grau de diferenciação tumoral, sendo o primeiro relato da implicação de transcritos dessa natureza em uma doença humana. Essa descoberta motiva a elaboração de hipóteses sobre possíveis funções biológicas destes transcritos no controle fino da expressão gênica. Ainda, apontou para a possibilidade de inclusão dessas mensagens no arsenal molecular para ajudar aos médicos na determinação do prognóstico de pacientes com câncer.

A caracterização da biogênese desses ncRNAs foi iniciada nesse projeto, revelando que assim como os mRNAs codificantes para proteínas, eles respondem a estímulos fisiológicos hormonais. Tomando como um exemplo o transcrito antisenso *asMYO5A*, foi demonstrado que a ativação da expressão em resposta ao andrógeno ocorre de forma independente do mRNA, transcrito a partir da fita oposta; a ativação ocorre através da ligação do complexo andrógeno/receptor de andrógeno diretamente a um elemento de resposta intrônico, posicionado no genoma na região a 5' do local de início de transcrição intrônica. Outros resultados obtidos ainda revelam que possivelmente existe o envolvimento de diferentes RNA polimerases (RNAPs) na transcrição de uma fração significativa dessas mensagens. O aumento observado na abundância de uma porcentagem considerável dos Longos RNAs Intrônicos não-codificantes após a inibição da atividade da RNAP II é um forte indicativo do envolvimento de outra RNAP, sendo a spRNAP IV uma candidata bastante provável.

Papéis funcionais para esses transcritos intrônicos em eventos celulares normais ganharam novas evidências com os experimentos realizados durante esse projeto. A conservação de expressão específica em tecidos, observada entre humanos e camundongos, fortalece a visão de uma transcrição controlada que deve

estar envolvida em eventos característicos de cada tipo celular. A provável interação desses RNAs com proteínas ligadoras de RNA (RBPs) envolvidas em mecanismos de regulação pós-transcricional, assim como a observação de transcrição intrônica correlacionada tanto com a abundância, quanto com o padrão de utilização de exons de mensagens codificantes para proteína de mesmo *locus* em células tratadas com andrógeno, reforçam a atuação dos RNAs intrônicos no controle do programa transcricional humano que foi postulada.

De fato os resultados aqui apresentados abrem muitas possibilidades de estudos com os Longos RNAs Intrônicos não-codificantes. A futura utilização do sistema de expressão para células de mamíferos, já construído durante o projeto, para modular os níveis desses transcritos trará mais pistas, ainda mais conclusivas a respeito da funcionalidade dessa classe de mensagens, contribuindo com um considerável avanço no entendimento do controle do complexo programa transcricional humano.

7 REFERÊNCIAS BIBLIOGRÁFICAS

Abate-Shen, C. & M. M. Shen. (2000). Molecular genetics of prostate cancer. *Genes Dev*, **14**, 19, 2410-34.

Adjaye, J., R. Herwig, D. Herrmann, W. Wruck, A. Benkahla, *et al.* (2004). Cross-species hybridisation of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics*, **5**, 1, 83.

Adjei, A. A. (2001). Blocking oncogenic Ras signaling for cancer therapy. *J Natl Cancer Inst*, **93**, 14, 1062-74.

Adjei, A. A., G. A. Croghan, C. Erlichman, R. S. Marks, J. M. Reid, *et al.* (2003). A Phase I trial of the farnesyl protein transferase inhibitor R115777 in combination with gemcitabine and cisplatin in patients with advanced cancer. *Clin Cancer Res*, **9**, 7, 2520-6.

Adjei, A. A., A. Mauer, L. Bruzek, R. S. Marks, S. Hillman, *et al.* (2003). Phase II study of the farnesyl transferase inhibitor R115777 in patients with advanced non-small-cell lung cancer. *J Clin Oncol*, **21**, 9, 1760-6.

Akhtar, A., D. Zink & P. B. Becker. (2000). Chromodomains are protein-RNA interaction modules. *Nature*, **407**, 6802, 405-9.

Akker, S. A., P. J. Smith & S. L. Chew. (2001). Nuclear post-transcriptional control of gene expression. *J Mol Endocrinol*, **27**, 2, 123-31.

Albertsen, P. C., J. A. Hanley, D. F. Gleason & M. J. Barry. (1998). Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *Jama*, **280**, 11, 975-80.

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, *et al.* (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 17, 3389-402.

Au, J. S. & J. D. Huang. (2002). A tissue-specific exon of myosin Va is responsible for selective cargo binding in melanocytes. *Cell Motil Cytoskeleton*, **53**, 2, 89-102.

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 2, 281-97.

Bertone, P., V. Stolc, T. E. Royce, J. S. Rozowsky, A. E. Urban, *et al.* (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science*, **306**, 5705, 2242-2246.

Best, C. J., I. M. Leiva, R. F. Chuaqui, J. W. Gillespie, P. H. Duray, *et al.* (2003). Molecular differentiation of high- and moderate-grade human prostate cancer by cDNA microarray analysis. *Diagn. Mol. Pathol.*, **12**, 2, 63-70.

Black, D. L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem*, **72**, 291-336.

Boeck, R. & D. Kolakofsky. (1994). Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis. *EMBO J.*, **13**, 15, 3608-3617.

Boettner, B. & L. Van Aelst. (2002). The role of Rho GTPases in disease development. *Gene*, **286**, 2, 155-74.

Bos, J. L. (1997). Ras-like GTPases. *Biochim Biophys Acta*, **1333**, 2, M19-31.

Bruno, I. G., W. Jin & G. J. Cote. (2004). Correction of aberrant FGFR1 alternative RNA splicing through targeting of intronic regulatory elements. *Hum Mol Genet*, **13**, 20, 2409-20.

Brunschwig, E. B., K. Wilson, D. Mack, D. Dawson, E. Lawrence, *et al.* (2003). PMEPA1, a transforming growth factor-beta-induced marker of terminal colonocyte differentiation whose expression is maintained in primary and metastatic colon cancer. *Cancer Res.*, **63**, 7, 1568-1575.

Bushnell, D. A., P. Cramer & R. D. Kornberg. (2002). Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 Å resolution. *Proc Natl Acad Sci U S A*, **99**, 3, 1218-22.

Calin, G. A., C. G. Liu, C. Sevignani, M. Ferracin, N. Felli, *et al.* (2004). MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A*, **101**, 32, 11755-60.

Calvo, A., O. Gonzalez-Moreno, C. Y. Yoon, J. I. Huh, K. Desai, *et al.* (2005). Prostate cancer and the genomic revolution: Advances using microarray analyses. *Mutat Res*, **576**, 1-2, 66-79.

Camargo, A. A., H. P. Samaia, E. Dias-Neto, D. F. Simao, I. A. Migotto, *et al.* (2001). The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome. *Proc Natl Acad Sci U S A*, **98**, 21, 12103-8.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, *et al.* (2005). The transcriptional landscape of the mammalian genome. *Science*, **309**, 5740, 1559-63.

Caron, P. R., M. D. Mullican, R. D. Mashal, K. P. Wilson, M. S. Su, *et al.* (2001). Chemogenomic approaches to drug discovery. *Curr Opin Chem Biol*, **5**, 4, 464-70.

Carroll, P. (2001). Rising PSA after a radical treatment. *Eur Urol*, **40 Suppl 2**, 9-16.

Cavaille, J., K. Buiting, M. Kieffmann, M. Lalande, C. I. Brannan, *et al.* (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. *Proc Natl Acad Sci U S A*, **97**, 26, 14311-6.

- Cavaille, J., P. Vitali, E. Basyuk, A. Huttenhofer & J. P. Bachellerie. (2001). A novel brain-specific box C/D small nucleolar RNA processed from tandemly repeated introns of a noncoding RNA gene in rats. *J Biol Chem*, **276**, 28, 26374-83.
- Cawley, S., S. Bekiranov, H. H. Ng, P. Kapranov, E. A. Sekinger, *et al.* (2004). Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs. *Cell*, **116**, 4, 499-509.
- Chen, G. & A. J. Courey. (2000). Groucho/TLE family proteins and transcriptional repression. *Gene*, **249**, 1-2, 1-16.
- Chen, J., M. Sun, W. J. Kent, X. Huang, H. Xie, *et al.* (2004). Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res*, **32**, 16, 4812-20.
- Chirgwin, J. M., A. E. Przybyla, R. J. Macdonald & W. J. Rutter. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18**, 24, 5294-9.
- Chomczynski, P. & N. Sacchi. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 1, 156-9.
- Claessens, F. & D. T. Gewirth. (2004). DNA recognition by nuclear receptors. *Essays Biochem*, **40**, 59-72.
- Cohen, P. (1999). The development and therapeutic potential of protein kinase inhibitors. *Curr Opin Chem Biol*, **3**, 4, 459-65.
- Cox, A. D. & C. J. Der. (2002). Ras family signaling: therapeutic targeting. *Cancer Biol Ther*, **1**, 6, 599-606.
- Cox, D. N., A. Chao, J. Baker, L. Chang, D. Qiao, *et al.* (1998). A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev*, **12**, 23, 3715-27.
- Crooks, G. E., G. Hon, J. M. Chandonia & S. E. Brenner. (2004). WebLogo: a sequence logo generator. *Genome Res*, **14**, 6, 1188-90.
- Dai, Q., E. Choy, V. Chiu, J. Romano, S. R. Slivka, *et al.* (1998). Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum. *J. Biol. Chem.*, **273**, 24, 15030-15034.
- Davis, E., F. Caiment, X. Tordoir, J. Cavaille, A. Ferguson-Smith, *et al.* (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. *Curr Biol*, **15**, 8, 743-9.
- Dennis, G., Jr., B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, *et al.* (2003). DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol*, **4**, 5, P3.

Dermitzakis, E. T., A. Reymond & S. E. Antonarakis. (2005). Conserved non-genic sequences - an unexpected feature of mammalian genomes. *Nat Rev Genet*, **6**, 2, 151-7.

Desilva, U., L. Elnitski, J. R. Idol, J. L. Doyle, W. Gan, *et al.* (2002). Generation and comparative analysis of approximately 3.3 Mb of mouse genomic sequence orthologous to the region of human chromosome 7q11.23 implicated in Williams syndrome. *Genome Res*, **12**, 1, 3-15.

Dhanasekaran, S. M., T. R. Barrette, D. Ghosh, R. Shah, S. Varambally, *et al.* (2001). Delineation of prognostic biomarkers in prostate cancer. *Nature*, **412**, 6849, 822-6.

Dias-Neto, E., R. G. Correa, S. Verjovski-Almeida, M. R. Briones, M. A. Nagai, *et al.* (2000). Shotgun sequencing of the human transcriptome with ORF expressed sequence tags. *Proc. Natl. Acad. Sci. USA*, **97**, 7, 3491-3496.

Du, T. & P. D. Zamore. (2005). microPrimer: the biogenesis and function of microRNA. *Development*, **132**, 21, 4645-52.

Duret, L., F. Dorkeld & C. Gautier. (1993). Strong conservation of non-coding sequences during vertebrates evolution: potential involvement in post-transcriptional regulation of gene expression. *Nucleic Acids Res*, **21**, 10, 2315-22.

Dvir, A. (2002). Promoter escape by RNA polymerase II. *Biochim Biophys Acta*, **1577**, 2, 208-223.

Eckley, D. M. & T. A. Schroer. (2003). Interactions between the evolutionarily conserved, actin-related protein, Arp11, actin, and Arp1. *Mol. Biol. Cell*, **14**, 7, 2645-2654.

Ellis, C. A., M. D. Vos, H. Howell, T. Vallecorsa, D. W. Fults, *et al.* (2002a). Rig is a novel Ras-related protein and potential neural tumor suppressor. *Proc Natl Acad Sci U S A*, **99**, 15, 9876-81.

Ellis, C. A., M. D. Vos, H. Howell, T. Vallecorsa, D. W. Fults, *et al.* (2002b). Rig is a novel Ras-related protein and potential neural tumor suppressor. *Proc. Natl. Acad. Sci. U. S. A.*, **99**, 15, 9876-9881.

Enard, W., P. Khaitovich, J. Klose, S. Zollner, F. Heissig, *et al.* (2002). Intra- and interspecific variation in primate gene expression patterns. *Science*, **296**, 5566, 340-3.

Epstein, J. I., A. W. Partin, J. Sauvageot & P. C. Walsh. (1996). Prediction of progression following radical prostatectomy. A multivariate analysis of 721 men with long-term follow-up. *Am. J. Surg. Pathol.*, **20**, 3, 286-292.

Erdmann, V. A., M. Z. Barciszewska, M. Szymanski, A. Hochberg, N. De Groot, *et al.* (2001). The non-coding RNAs as riboregulators. *Nucleic Acids Res*, **29**, 1, 189-93.

Esquela-Kerscher, A. & F. J. Slack. (2006). Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer*, **6**, 4, 259-69.

Euskirchen, G., T. E. Royce, P. Bertone, R. Martone, J. L. Rinn, *et al.* (2004). CREB binds to multiple loci on human chromosome 22. *Mol Cell Biol*, **24**, 9, 3804-14.

Ewing, B. & P. Green. (2000). Analysis of expressed sequence tags indicates 35,000 human genes. *Nat Genet*, **25**, 2, 232-4.

Farrell, C. M. & L. N. Lukens. (1995). Naturally occurring antisense transcripts are present in chick embryo chondrocytes simultaneously with the down-regulation of the alpha 1 (I) collagen gene. *J Biol Chem*, **270**, 7, 3400-8.

Finlin, B. S., C. L. Gau, G. A. Murphy, H. Shao, T. Kimel, *et al.* (2001). RERG is a novel ras-related, estrogen-regulated and growth-inhibitory gene in breast cancer. *J Biol Chem*, **276**, 45, 42259-67.

Frazer, K. A., L. Elnitski, D. M. Church, I. Dubchak & R. C. Hardison. (2003). Cross-species sequence comparisons: a review of methods and available resources. *Genome Res*, **13**, 1, 1-12.

Futreal, P. A., A. Kasprzyk, E. Birney, J. C. Mullikin, R. Wooster, *et al.* (2001). Cancer and genomics. *Nature*, **409**, 6822, 850-2.

Garcia-Echeverria, C., P. Traxler & D. B. Evans. (2000). ATP site-directed competitive and irreversible inhibitors of protein kinases. *Med Res Rev*, **20**, 1, 28-57.

Gebuhr, T. C., S. J. Bultman & T. Magnuson. (2000). Pc-G/trx-G and the SWI/SNF connection: developmental gene regulation through chromatin remodeling. *Genesis*, **26**, 3, 189-97.

Geck, P., J. Szelei, J. Jimenez, T. M. Lin, C. Sonnenschein, *et al.* (1997). Expression of novel genes linked to the androgen-induced, proliferative shutoff in prostate cancer cells. *J Steroid Biochem Mol Biol*, **63**, 4-6, 211-8.

Gemmill, R. M., J. D. West, F. Boldog, N. Tanaka, L. J. Robinson, *et al.* (1998). The hereditary renal cell carcinoma 3;8 translocation fuses FHIT to a patched-related gene, TRC8. *Proc. Natl. Acad. Sci. U S A*, **95**, 16, 9572-9577.

Gilad, Y., S. A. Rifkin, P. Bertone, M. Gerstein & K. P. White. (2005). Multi-species microarrays reveal the effect of sequence divergence on gene expression profiles. *Genome Res*, **15**, 5, 674-80.

Girard, A., R. Sachidanandam, G. J. Hannon & M. A. Carmell. (2006). A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, **442**, 7099, 199-202.

Gleason, D. Histologic grading and clinical staging of carcinoma of the prostate. In: M. Tannenbaun (Ed.). Urologic pathology. Philadelphia: Lea & Febiger, 1977. Histologic grading and clinical staging of carcinoma of the prostate., p.171-197

- Glinsky, G. V., A. B. Glinskii, A. J. Stephenson, R. M. Hoffman & W. L. Gerald. (2004). Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest*, **113**, 6, 913-23.
- Golub, T. R., D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, *et al.* (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science*, **286**, 5439, 531-7.
- Griffiths-Jones, S., R. J. Grocock, S. Van Dongen, A. Bateman & A. J. Enright. (2006). miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*, **34**, Database issue, D140-4.
- Hardison, R. C., J. Oeltjen & W. Miller. (1997). Long human-mouse sequence alignments reveal novel regulatory elements: a reason to sequence the mouse genome. *Genome Res*, **7**, 10, 959-66.
- Harlow, E. & D. Lane. Antibodies: A Laboratory Manual. 1998
- Harris, C. J. & A. P. Stevens. (2006). Chemogenomics: structuring the drug discovery process to gene families. *Drug Discov Today*, **11**, 19-20, 880-8.
- Hill, A. E., J. S. Hong, H. Wen, L. Teng, D. T. Mcpherson, *et al.* (2006). Micro-RNA-like effects of complete intronic sequences. *Front Biosci*, **11**, 1998-2006.
- Hogenesch, J. B., K. A. Ching, S. Batalov, A. I. Su, J. R. Walker, *et al.* (2001). A comparison of the Celera and Ensembl predicted gene sets reveals little overlap in novel genes. *Cell*, **106**, 4, 413-5.
- Howe, K. J. (2002). RNA polymerase II conducts a symphony of pre-mRNA processing activities. *Biochim Biophys Acta*, **1577**, 2, 308-24.
- Huang, C., E. A. Sloan & C. F. Boerkoel. (2003). Chromatin remodeling and human disease. *Curr Opin Genet Dev*, **13**, 3, 246-52.
- Hughes, T. R., M. Mao, A. R. Jones, J. Burchard, M. J. Marton, *et al.* (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.*, **19**, 4, 342-347.
- Imamura, T., S. Yamamoto, J. Ohgane, N. Hattori, S. Tanaka, *et al.* (2004). Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochem Biophys Res Commun*, **322**, 2, 593-600.
- Isaacs, J. T., P. I. Lundmo, R. Berges, P. Martikainen, N. Kyprianou, *et al.* (1992). Androgen regulation of programmed death of normal and malignant prostatic cells. *J Androl*, **13**, 6, 457-64.
- Iseli, C., C. V. Jongeneel & P. Bucher. (1999). ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, 138-148.

Jacob, S. T., E. M. Sajdel & H. N. Munro. (1970). Specific action of alpha-amanitin on mammalian RNA polymerase protein. *Nature*, **225**, 5227, 60-2.

Jensen, R. A. (2001). Orthologs and paralogs - we need to get it right. *Genome Biol*, **2**, 8, INTERACTIONS1002.

Ji, P., S. Diederichs, W. Wang, S. Boing, R. Metzger, *et al.* (2003). MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, **22**, 39, 8031-41.

Johnson, J. M., S. Edwards, D. Shoemaker & E. E. Schadt. (2005). Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet*, **21**, 2, 93-102.

Jones, D. O., I. G. Cowell & P. B. Singh. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays*, **22**, 2, 124-37.

Jongeneel, C. V., M. Delorenzi, C. Iseli, D. Zhou, C. D. Haudenschield, *et al.* (2005). An atlas of human gene expression from massively parallel signature sequencing (MPSS). *Genome Res*, **15**, 7, 1007-14.

Kampa, D., J. Cheng, P. Kapranov, M. Yamanaka, S. Brubaker, *et al.* (2004). Novel RNAs Identified From an In-Depth Analysis of the Transcriptome of Human Chromosomes 21 and 22. *Genome Res.*, **14**, 3, 331-342.

Kapranov, P., S. E. Cawley, J. Drenkow, S. Bekiranov, R. L. Strausberg, *et al.* (2002). Large-scale transcriptional activity in chromosomes 21 and 22. *Science*, **296**, 5569, 916-919.

Kapranov, P., J. Drenkow, J. Cheng, J. Long, G. Helt, *et al.* (2005). Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res*, **15**, 7, 987-97.

Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, *et al.* (2005). Antisense transcription in the mammalian transcriptome. *Science*, **309**, 5740, 1564-6.

Kedinger, C., M. Gniazdowski, J. L. Mandel, Jr., F. Gissinger & P. Chambon. (1970). Alpha-amanitin: a specific inhibitor of one of two DNA-pendent RNA polymerase activities from calf thymus. *Biochem Biophys Res Commun*, **38**, 1, 165-71.

Khochbin, S. & J. J. Lawrence. (1989). An antisense RNA involved in p53 mRNA maturation in murine erythroleukemia cells induced to differentiate. *EMBO J.*, **8**, 13, 4107-4114.

Kim, T. H., L. O. Barrera, C. Qu, S. Van Calcar, N. D. Trinklein, *et al.* (2005). Direct isolation and identification of promoters in the human genome. *Genome Res*, **15**, 6, 830-9.

Kishore, S. & S. Stamm. (2006). The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science*, **311**, 5758, 230-2.

Kiss, T. (2002). Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell*, **109**, 2, 145-8.

Kiyosawa, H., I. Yamanaka, N. Osato, S. Kondo & Y. Hayashizaki. (2003). Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res.*, **13**, 6B, 1324-1334.

Kondo, S., N. Yamamoto, T. Murakami, M. Okumura, A. Mayeda, *et al.* (2004). Tra2 beta, SF2/ASF and SRp30c modulate the function of an exonic splicing enhancer in exon 10 of tau pre-mRNA. *Genes Cells*, **9**, 2, 121-30.

Kravchenko, J. E., I. B. Rogozin, E. V. Koonin & P. M. Chumakov. (2005). Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin. *Nature*, **436**, 7051, 735-9.

Kuge, O., C. Dascher, L. Orci, T. Rowe, M. Amherdt, *et al.* (1994). Sar1 promotes vesicle budding from the endoplasmic reticulum but not Golgi compartments. *J. Cell Biol.*, **125**, 1, 51-65.

Lander, E. S., L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, *et al.* (2001). Initial sequencing and analysis of the human genome. *Nature*, **409**, 6822, 860-921.

Lapointe, J., C. Li, J. P. Higgins, M. Van De Rijn, E. Bair, *et al.* (2004). Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A*, **101**, 3, 811-6.

Lee, C. H., N. G. Della, C. E. Chew & D. J. Zack. (1996). Rin, a neuron-specific and calmodulin-binding small G-protein, and Rit define a novel subfamily of ras proteins. *J. Neurosci.*, **16**, 21, 6784-6794.

Lee, J. T., L. S. Davidow & D. Warshawsky. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat. Genet.*, **21**, 4, 400-404.

Lee, T. I. & R. A. Young. (2000). Transcription of eukaryotic protein-coding genes. *Annu Rev Genet*, **34**, 77-137.

Lewis, B. P., C. B. Burge & D. P. Bartel. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 1, 15-20.

Liang, F., I. Holt, G. Pertea, S. Karamycheva, S. L. Salzberg, *et al.* (2000). Gene index analysis of the human genome estimates approximately 120,000 genes. *Nat Genet*, **25**, 2, 239-40.

Lindell, T. J., F. Weinberg, P. W. Morris, R. G. Roeder & W. J. Rutter. (1970). Specific inhibition of nuclear RNA polymerase II by alpha-amanitin. *Science*, **170**, 956, 447-9.

Liu, L., J. H. Yoon, R. Dammann & G. P. Pfeifer. (2002). Frequent hypermethylation of the RASSF1A gene in prostate cancer. *Oncogene*, **21**, 44, 6835-6840.

Liu, Y., G. Dehni, K. J. Purcell, J. Sokolow, M. L. Carcangiu, *et al.* (1996). Epithelial expression and chromosomal location of human TLE genes: implications for notch signaling and neoplasia. *Genomics*, **31**, 1, 58-64.

Lonard, D. M. & B. W. O'malley. (2005). Expanding functional diversity of the coactivators. *Trends Biochem Sci*, **30**, 3, 126-32.

Loots, G. G., R. M. Locksley, C. M. Blankespoor, Z. E. Wang, W. Miller, *et al.* (2000). Identification of a coordinate regulator of interleukins 4, 13, and 5 by cross-species sequence comparisons. *Science*, **288**, 5463, 136-40.

Louro, R., H. I. Nakaya, P. P. Amaral, F. Festa, M. C. Sogayar, *et al.* (2006). Androgen responsive intronic non-coding RNAs. *BMC Genomics*, **(Submitted)**, ANEXO 7.

Louro, R., H. I. Nakaya, A. C. Paquola, E. A. Martins, A. M. Da Silva, *et al.* (2004). RASL11A, member of a novel small monomeric GTPase gene family, is down-regulated in prostate tumors. *Biochem Biophys Res Commun*, **316**, 3, 618-27. ANEXO 4

Lu, J., G. Getz, E. A. Miska, E. Alvarez-Saavedra, J. Lamb, *et al.* (2005). MicroRNA expression profiles classify human cancers. *Nature*, **435**, 7043, 834-8.

Luo, J., S. Zha, W. R. Gage, T. A. Dunn, J. L. Hicks, *et al.* (2002). Alpha-methylacyl-CoA racemase: a new molecular marker for prostate cancer. *Cancer Res*, **62**, 8, 2220-6.

Luo, R. Z., X. Fang, R. Marquez, S. Y. Liu, G. B. Mills, *et al.* (2003). ARHI is a Ras-related small G-protein with a novel N-terminal extension that inhibits growth of ovarian and breast cancers. *Oncogene*, **22**, 19, 2897-909.

Maffini, M. V., P. Geck, C. E. Powell, C. Sonnenschein & A. M. Soto. (2002). Mechanism of androgen action on cell proliferation: AS3 protein as a mediator of proliferative arrest in the rat prostate. *Endocrinology*, **143**, 7, 2708-14.

Makalowski, W., J. Zhang & M. S. Boguski. (1996). Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res*, **6**, 9, 846-57.

Mangus, D. A., M. C. Evans & A. Jacobson. (2003). Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol*, **4**, 7, 223.

Mansfield, J. H., B. D. Harfe, R. Nissen, J. Obenauer, J. Srineel, *et al.* (2004). MicroRNA-responsive 'sensor' transgenes uncover Hox-like and other

developmentally regulated patterns of vertebrate microRNA expression. *Nat Genet*, **36**, 10, 1079-83.

Marchler-Bauer, A., J. B. Anderson, P. F. Cherukuri, C. Deweese-Scott, L. Y. Geer, *et al.* (2005). CDD: a Conserved Domain Database for protein classification. *Nucleic Acids Res*, **33**, Database issue, D192-6.

Margulies, E. H., M. Blanchette, D. Haussler & E. D. Green. (2003). Identification and characterization of multi-species conserved sequences. *Genome Res*, **13**, 12, 2507-18.

Martone, R., G. Euskirchen, P. Bertone, S. Hartman, T. E. Royce, *et al.* (2003). Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci U S A*, **100**, 21, 12247-52.

Maston, G. A., S. K. Evans & M. R. Green. (2006). Transcriptional Regulatory Elements in the Human Genome. *Annu Rev Genomics Hum Genet*.

Mata, J., S. Marguerat & J. Bahler. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci*, **30**, 9, 506-14.

Mattick, J. S. (2004). RNA regulation: a new genetics? *Nat Rev Genet*, **5**, 4, 316-323.

Mattick, J. S. & I. V. Makunin. (2006). Non-coding RNA. *Hum Mol Genet*, **15 Suppl 1**, R17-29.

Matys, V., E. Fricke, R. Geffers, E. Gossling, M. Haubrock, *et al.* (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res*, **31**, 1, 374-8.

Mcmanus, M. T. (2003). MicroRNAs and cancer. *Semin. Cancer Biol.*, **13**, 4, 253-258.

Miyaso, H., M. Okumura, S. Kondo, S. Higashide, H. Miyajima, *et al.* (2003). An intronic splicing enhancer element in survival motor neuron (SMN) pre-mRNA. *J Biol Chem*, **278**, 18, 15825-31.

Moody, D. E., Z. Zou & L. McIntyre. (2002). Cross-species hybridisation of pig RNA to human nylon microarrays. *BMC Genomics*, **3**, 1, 27.

Nadon, R., P. Shi, A. Skandalis, E. Woody, H. Hubschle, *et al.* (2001). Statistical inference methods for gene expression arrays. Microarrays: Optical Technologies and Informatics. *Proceedings of SPIE*, **4266**, 46-55.

Nakaya, H. I., P. P. Amaral, R. Louro, A. Lopes, A. A. Fachel, *et al.* (2006). Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription *Genome Biol*, **(Submitted)**, ANEXO 2.

- Nelson, P. S., N. Clegg, H. Arnold, C. Ferguson, M. Bonham, *et al.* (2002). The program of androgen-responsive genes in neoplastic prostate epithelium. *Proc Natl Acad Sci U S A*, **99**, 18, 11890-5.
- Nepveu, A. & K. B. Marcu. (1986). Intragenic pausing and anti-sense transcription within the murine c-myc locus. *EMBO J.*, **5**, 11, 2859-2865.
- Noguchi, M., S. Miyamoto, T. A. Silverman & B. Safer. (1994). Characterization of an antisense Inr element in the eIF-2 alpha gene. *J Biol Chem*, **269**, 46, 29161-7.
- Numata, K., A. Kanai, R. Saito, S. Kondo, J. Adachi, *et al.* (2003). Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection. *Genome Res*, **13**, 6B, 1301-6.
- Oesterling, J. E., J. Roy, A. Agha, T. Shown, T. Krarup, *et al.* (1997). Biologic variability of prostate-specific antigen and its usefulness as a marker for prostate cancer: effects of finasteride. The Finasteride PSA Study Group. *Urology*, **50**, 1, 13-8.
- Okazaki, Y., M. Furuno, T. Kasukawa, J. Adachi, H. Bono, *et al.* (2002). Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature*, **420**, 6915, 563-573.
- Ota, T., Y. Suzuki, T. Nishikawa, T. Otsuki, T. Sugiyama, *et al.* (2004). Complete sequencing and characterization of 21,243 full-length human cDNAs. *Nat Genet*, **36**, 1, 40-5.
- Oxford, G. & D. Theodorescu. (2003). Ras superfamily monomeric G proteins in carcinoma cell motility. *Cancer Lett*, **189**, 2, 117-28.
- Pabo, C. O. & R. T. Sauer. (1992). Transcription factors: structural families and principles of DNA recognition. *Annu Rev Biochem*, **61**, 1053-95.
- Paduch, M., F. Jelen & J. Otlewski. (2001). Structure of small G proteins and their regulators. *Acta Biochim Pol*, **48**, 4, 829-50.
- Page, R. D. (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*, **12**, 4, 357-8.
- Pang, K. C., M. C. Frith & J. S. Mattick. (2006). Rapid evolution of noncoding RNAs: lack of conservation does not mean lack of function. *Trends Genet*, **22**, 1, 1-5.
- Pasqualato, S., L. Renault & J. Cherfils. (2002a). Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep*, **3**, 11, 1035-41.
- Pasqualato, S., L. Renault & J. Cherfils. (2002b). Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication. *EMBO Rep.*, **3**, 11, 1035-1041.

Paule, M. R. & R. J. White. (2000). Survey and summary: transcription by RNA polymerases I and III. *Nucleic Acids Res*, **28**, 6, 1283-98.

Penalva, L. O. & J. D. Keene. (2004). Biotinylated tags for recovery and characterization of ribonucleoprotein complexes. *Biotechniques*, **37**, 4, 604, 606, 608-10.

Penalva, L. O., S. A. Tenenbaum & J. D. Keene. (2004). Gene expression analysis of messenger RNP complexes. *Methods Mol Biol*, **257**, 125-34.

Penny, G. D., G. F. Kay, S. A. Sheardown, S. Rastan & N. Brockdorff. (1996). Requirement for Xist in X chromosome inactivation. *Nature*, **379**, 6561, 131-7.

Petrovics, G., A. Liu, S. Shaheduzzaman, B. Furasato, C. Sun, *et al.* (2005). Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene*, **24**, 23, 3847-52.

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.*, **29**, 9, e45.

Potter, S. R., W. Horniger, M. Tinzl, G. Bartsch & A. W. Partin. (2001). Age, prostate-specific antigen, and digital rectal examination as determinants of the probability of having prostate cancer. *Urology*, **57**, 6, 1100-4.

Prescott, E. M. & N. J. Proudfoot. (2002). Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci U S A*, **99**, 13, 8796-801.

Ptashne, M. & A. Gann. (1997). Transcriptional activation by recruitment. *Nature*, **386**, 6625, 569-77.

Quackenbush, J. (2002). Microarray data normalization and transformation. *Nat. Genet.*, **32 Suppl**, 496-501.

Ramaswamy, S., K. N. Ross, E. S. Lander & T. R. Golub. (2003). A molecular signature of metastasis in primary solid tumors. *Nat. Genet.*, **33**, 1, 49-54.

Rao, S., D. Cunningham, A. De Gramont, W. Scheithauer, M. Smakal, *et al.* (2004). Phase III double-blind placebo-controlled study of farnesyl transferase inhibitor R115777 in patients with refractory advanced colorectal cancer. *J Clin Oncol*, **22**, 19, 3950-7.

Ravasi, T., H. Suzuki, K. C. Pang, S. Katayama, M. Furuno, *et al.* (2006). Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res*, **16**, 1, 11-9.

Reis, E. M., R. Louro, H. I. Nakaya & S. Verjovski-Almeida. (2005). As antisense RNA gets intronic. *Omics*, **9**, 1, 2-12. ANEXO 1

Reis, E. M., H. I. Nakaya, R. Louro, F. C. Canavez, A. V. Flatschart, *et al.* (2004). Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene*, **23**, 39, 6684-6692. ANEXO 3

Reuther, G. W. & C. J. Der. (2000a). The Ras branch of small GTPases: Ras family members don't fall far from the tree. *Curr. Opin. Cell Biol.*, **12**, 2, 157-165.

Reuther, G. W. & C. J. Der. (2000b). The Ras branch of small GTPases: Ras family members don't fall far from the tree. *Curr Opin Cell Biol*, **12**, 2, 157-65.

Rhodes, D. R., T. R. Barrette, M. A. Rubin, D. Ghosh & A. M. Chinnaiyan. (2002). Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res*, **62**, 15, 4427-33.

Richter, J. D. & N. Sonenberg. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature*, **433**, 7025, 477-80.

Rinn, J. L., G. Euskirchen, P. Bertone, R. Martone, N. M. Luscombe, *et al.* (2003). The transcriptional activity of human Chromosome 22. *Genes Dev.*, **17**, 4, 529-540.

Rogelj, B. & K. P. Giese. (2004). Expression and function of brain specific small RNAs. *Rev Neurosci*, **15**, 3, 185-98.

Saha, S., A. B. Sparks, C. Rago, V. Akmaev, C. J. Wang, *et al.* (2002). Using the transcriptome to annotate the genome. *Nat Biotechnol*, **20**, 5, 508-12.

Sainaghi, P. P., L. Castello, L. Bergamasco, M. Galletti, P. Bellosta, *et al.* (2005). Gas6 induces proliferation in prostate carcinoma cell lines expressing the Axl receptor. *J Cell Physiol*, **204**, 1, 36-44.

Saito, K., T. Fujiwara, J. Katahira, K. Inoue & H. Sakamoto. (2004). TAP/NXF1, the primary mRNA export receptor, specifically interacts with a neuronal RNA-binding protein HuD. *Biochem Biophys Res Commun*, **321**, 2, 291-7.

Sanchez-Carbayo, M., N. D. Socci, E. Charytonowicz, M. Lu, M. Prystowsky, *et al.* (2002). Molecular profiling of bladder cancer using cDNA microarrays: defining histogenesis and biological phenotypes. *Cancer Res.*, **62**, 23, 6973-6980.

Sanchez-Diaz, P. & L. O. Penalva. Post-transcriptional meets post-genomic. The saga of RNA binding proteins in a new era. 2006.

Sandelin, A., P. Bailey, S. Bruce, P. G. Engstrom, J. M. Klos, *et al.* (2004). Arrays of ultraconserved non-coding regions span the loci of key developmental genes in vertebrate genomes. *BMC Genomics*, **5**, 1, 99.

Shaffer, P. L., A. Jivan, D. E. Dollins, F. Claessens & D. T. Gewirth. (2004). Structural basis of androgen receptor binding to selective androgen response elements. *Proc Natl Acad Sci U S A*, **101**, 14, 4758-63.

Shatkin, A. J. & J. L. Manley. (2000). The ends of the affair: capping and polyadenylation. *Nat Struct Biol*, **7**, 10, 838-42.

She, X., Z. Jiang, R. A. Clark, G. Liu, Z. Cheng, *et al.* (2004). Shotgun sequence assembly and recent segmental duplications within the human genome. *Nature*, **431**, 7011, 927-30.

Shendure, J. & G. M. Church. (2002). Computational discovery of sense-antisense transcription in the human and mouse genomes. *Genome Biol.*, **3**, 9, research0044.1-research0044.14.

Shoemaker, D. D., E. E. Schadt, C. D. Armour, Y. D. He, P. Garrett-Engle, *et al.* (2001). Experimental annotation of the human genome using microarray technology. *Nature*, **409**, 6822, 922-927.

Silverman, T. A., M. Noguchi & B. Safer. (1992). Role of sequences within the first intron in the regulation of expression of eukaryotic initiation factor 2 alpha. *J Biol Chem*, **267**, 14, 9738-42.

Sinensky, M. (2000). Recent advances in the study of prenylated proteins. *Biochim. Biophys. Acta*, **1484**, 2-3, 93-106.

Singh, D., P. G. Febbo, K. Ross, D. G. Jackson, J. Manola, *et al.* (2002). Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell*, **1**, 2, 203-209.

Siomi, H. & G. Dreyfuss. (1997). RNA-binding proteins as regulators of gene expression. *Curr Opin Genet Dev*, **7**, 3, 345-53.

Sleutels, F., D. P. Barlow & R. Lyle. (2000). The uniqueness of the imprinting mechanism. *Curr. Opin. Genet. Dev.*, **10**, 2, 229-233.

Sleutels, F., R. Zwart & D. P. Barlow. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, **415**, 6873, 810-3.

Smith, C. W. & J. Valcarcel. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem Sci*, **25**, 8, 381-8.

Sorek, R. & H. M. Safer. (2003). A novel algorithm for computational identification of contaminated EST libraries. *Nucleic Acids Res.*, **31**, 3, 1067-1074.

Spoerner, M., C. Herrmann, I. R. Vetter, H. R. Kalbitzer & A. Wittinghofer. (2001). Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl. Acad. Sci. U. S. A.*, **98**, 9, 4944-4949.

Strausberg, R. L., K. H. Buetow, M. R. Emmert-Buck & R. D. Klausner. (2000). The cancer genome anatomy project: building an annotated gene index. *Trends Genet*, **16**, 3, 103-6.

- Swinnen, J. V. & G. Verhoeven. (1998). Androgens and the control of lipid metabolism in human prostate cancer cells. *J Steroid Biochem Mol Biol*, **65**, 1-6, 191-8.
- Tamayo, P., D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, *et al.* (1999). Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA*, **96**, 6, 2907-2912.
- Tasheva, E. S. & D. J. Roufa. (1995). Regulation of human RPS14 transcription by intronic antisense RNAs and ribosomal protein S14. *Genes Dev.*, **9**, 3, 304-316.
- Tee, M. K. & R. B. Jaffe. (2001). A precursor form of vascular endothelial growth factor arises by initiation from an upstream in-frame CUG codon. *Biochem. J.*, **359**, Pt 1, 219-226.
- Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin & D. G. Higgins. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, **25**, 24, 4876-82.
- Thompson, J. D., D. G. Higgins & T. J. Gibson. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 22, 4673-80.
- Torarinsson, E., M. Sawera, J. H. Havgaard, M. Fredholm & J. Gorodkin. (2006). Thousands of corresponding human and mouse genomic regions unalignable in primary sequence contain common RNA structure. *Genome Res*, **16**, 7, 885-9.
- Tsoi, S. C., J. M. Cale, I. M. Bird, V. Ewart, L. L. Brown, *et al.* (2003). Use of human cDNA microarrays for identification of differentially expressed genes in Atlantic salmon liver during *Aeromonas salmonicida* infection. *Mar Biotechnol (NY)*, **5**, 6, 545-54.
- Tufarelli, C., J. A. Stanley, D. Garrick, J. A. Sharpe, H. Ayyub, *et al.* (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat. Genet.*, **34**, 2, 157-165.
- Tusher, V. G., R. Tibshirani & G. Chu. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA*, **98**, 9, 5116-5121.
- Tycowski, K. T., A. Aab & J. A. Steitz. (2004). Guide RNAs with 5' caps and novel box C/D snoRNA-like domains for modification of snRNAs in metazoa. *Curr Biol*, **14**, 22, 1985-95.
- Vallee, M., C. Robert, S. Methot, M. F. Palin & M. A. Sirard. (2006). Cross-species hybridizations on a multi-species cDNA microarray to identify evolutionarily conserved genes expressed in oocytes. *BMC Genomics*, **7**, 113.

Van 'T Veer, L. J., H. Dai, M. J. Van De Vijver, Y. D. He, A. A. Hart, *et al.* (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, **415**, 6871, 530-536.

Van Der Vlag, J. & A. P. Otte. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat Genet*, **23**, 4, 474-8.

Varambally, S., S. M. Dhanasekaran, M. Zhou, T. R. Barrette, C. Kumar-Sinha, *et al.* (2002). The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*, **419**, 6907, 624-9.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, *et al.* (2001). The sequence of the human genome. *Science*, **291**, 5507, 1304-1351.

Vetter, I. R. & A. Wittinghofer. (2001a). The guanine nucleotide-binding switch in three dimensions. *Science*, **294**, 5545, 1299-1304.

Vetter, I. R. & A. Wittinghofer. (2001b). The guanine nucleotide-binding switch in three dimensions. *Science*, **294**, 5545, 1299-304.

Wani, M. A., N. C. Denko & P. J. Stambrook. (1997a). Expression of Rap 1 suppresses genomic instability of H-ras transformed mouse fibroblasts. *Somat. Cell Mol. Genet.*, **23**, 2, 123-133.

Wani, M. A., N. C. Denko & P. J. Stambrook. (1997b). Expression of Rap 1 suppresses genomic instability of H-ras transformed mouse fibroblasts. *Somat Cell Mol Genet*, **23**, 2, 123-33.

Watanabe, M., K. Nomura, A. Ohyama, R. Ishikawa, Y. Komiya, *et al.* (2005). Myosin-Va regulates exocytosis through the submicromolar Ca²⁺-dependent binding of syntaxin-1A. *Mol Biol Cell*, **16**, 10, 4519-30.

Williamson, A. R. (1999). The Merck Gene Index project. *Drug Discov Today*, **4**, 3, 115-122.

Willingham, A. T. & T. R. Gingeras. (2006). TUF love for "junk" DNA. *Cell*, **125**, 7, 1215-20.

Wistow, G., S. L. Bernstein, S. Ray, M. K. Wyatt, A. Behal, *et al.* (2002). Expressed sequence tag analysis of adult human iris for the NEIBank Project: steroid-response factors and similarities with retinal pigment epithelium. *Mol. Vis.*, **8**, 185-195.

Woolfe, A., M. Goodson, D. K. Goode, P. Snell, G. K. McEwen, *et al.* (2005). Highly conserved non-coding sequences are associated with vertebrate development. *PLoS Biol*, **3**, 1, e7.

Wutz, A., O. W. Smrzka, N. Schweifer, K. Schellander, E. F. Wagner, *et al.* (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature*, **389**, 6652, 745-9.

- Xu, L. L., N. Shanmugam, T. Segawa, I. A. Sesterhenn, D. G. Mcleod, *et al.* (2000). A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate. *Genomics*, **66**, 3, 257-63.
- Yan, M. D., C. C. Hong, G. M. Lai, A. L. Cheng, Y. W. Lin, *et al.* (2005). Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas. *Hum. Mol. Genet.*, **14**, 11, 1465-1474.
- Yekta, S., I. H. Shih & D. P. Bartel. (2004). MicroRNA-directed cleavage of HOXB8 mRNA. *Science*, **304**, 5670, 594-6.
- Yelin, R., D. Dahary, R. Sorek, E. Y. Levanon, O. Goldstein, *et al.* (2003). Widespread occurrence of antisense transcription in the human genome. *Nat. Biotechnol.*, **21**, 4, 379-386.
- Yu, Y., F. Xu, H. Peng, X. Fang, S. Zhao, *et al.* (1999a). NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas. *Proc Natl Acad Sci U S A*, **96**, 1, 214-9.
- Yu, Y., F. Xu, H. Peng, X. Fang, S. Zhao, *et al.* (1999b). NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas. *Proc. Natl. Acad. Sci. U. S. A.*, **96**, 1, 214-219.
- Zohn, I. M., S. L. Campbell, R. Khosravi-Far, K. L. Rossman & C. J. Der. (1998). Rho family proteins and Ras transformation: the RHOad less traveled gets congested. *Oncogene*, **17**, 11 Reviews, 1415-38.

8 LISTA DE ANEXOS

Anexos inclusos no CD-ROM suplementar a esta Tese

ANEXO 1: Artigo publicado - Reis, E. M., R. Louro, H. I. Nakaya & S. Verjovski-Almeida. (2005). As antisense RNA gets intronic. *Omics*, **9**, 1, 2-12.

ANEXO 2: Artigo submetido - Nakaya, H. I., P. P. Amaral, R. Louro, A. Lopes, A. A. Fachel, *et al.* (2006). Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription *Genome Biol.*

ANEXO 3: Artigo publicado - Reis, E. M., H. I. Nakaya, R. Louro, F. C. Canavez, A. V. Flatschart, *et al.* (2004). Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene*, **23**, 39, 6684-6692.

ANEXO 4: Artigo publicado - Louro, R., H. I. Nakaya, A. C. Paquola, E. A. Martins, A. M. Da Silva, *et al.* (2004). RASL11A, member of a novel small monomeric GTPase gene family, is down-regulated in prostate tumors. *Biochem Biophys Res Commun*, **316**, 3, 618-27.

ANEXO 5: Tabela Suplementar 1 - Transcritos exônicos regulados por andrógeno.

ANEXO 6: Tabela Suplementar 2 - Transcritos intrônicos regulados por andrógeno.

ANEXO 7: Artigo Submetido - Louro, R., H. I. Nakaya, P. P. Amaral, F. Festa, M. C. Sogayar, *et al.* (2006). Androgen responsive intronic non-coding RNAs. *BMC Genomics*.

ANEXO 8: Tabela Suplementar 3 - Assinatura de expressão de transcritos intrônicos conservada em tecidos humanos e de camundongos.

ANEXO 9: Tabela Suplementar 4 - Categorias de GO enriquecidas de transcritos intrônicos imunoprecipitados com RBPs.

ANEXO 10: Tabela Suplementar 5 - Transcritos exônicos e intrônicos alterados pelo tratamento com α -amanitina.

ANEXO 11: Súmula Curricular.

Mini Review

As Antisense RNA Gets Intronic

EDUARDO M. REIS, RODRIGO LOURO, HELDER I. NAKAYA,
and SERGIO VERJOVSKI-ALMEIDA

ABSTRACT

Recent work describing the transcriptional output of the human genome points to the existence of a significant number of non-coding RNA transcripts coming from intronic regions, with a fraction of these being oriented antisense relative to the protein-coding mRNA of the known gene. In this article, we survey the main findings of the large-scale expression analysis projects that led to the identification of antisense intronic messages and which demonstrate their ubiquitous expression in the human genome. We review the current knowledge on long, unspliced, intronic antisense transcripts, a new class of non-coding RNAs, recently described by our group to be correlated with the degree of tumor differentiation in prostate cancer, which we postulate is involved in the fine tuning of gene expression in eukaryotes. Possible mechanisms of antisense intronic transcript biogenesis and function in gene expression regulation are discussed, as is their involvement in human diseases. While there is still no conclusive evidence demonstrating a functional role for these long, intronic antisense messages, the far-reaching implications of their existence for the mechanisms regulating gene expression certainly warrant further experimentation.

INTRODUCTION

The recent advent of microarray technology has paved the way for a refined high-throughput mapping of the transcriptional activity in the human genome (Shoemaker et al., 2001; Kapranov et al., 2002; Rinn et al., 2003; Bertone et al., 2004; Kampa et al., 2004). A pioneering study from Kapranov and colleagues revealed a 10-fold excess of transcriptionally active regions along chromosome 21 and 22 (Kapranov et al., 2002) than originally predicted by mapping of known genes (Lander et al., 2001; Venter et al., 2001). Using microarrays that probed the genomic DNA along all the extensions of chromosomes 21 and 22 (an approach that is essentially not biased towards the protein-coding regions), comparable fractions of transcriptional activity were detected within exons or introns of annotated genes, and nearly half of these intronic transcripts were expressed antisense to their respective well-characterized introns (Rinn et al., 2003;

Department of Biochemistry, Institute of Chemistry, University of São Paulo, São Paulo, Brazil.

Kampa et al., 2004). More recently, Bertone et al. (2004) analyzed the transcriptional activity in liver cells with poly(A) mRNA pooled from a number of individuals, using a whole genome tiling array. These authors found that 22% (3,095) of the 13,889 transcriptionally active regions detected by the tiling array were intronic. Remarkably, an equal fraction of transcripts was found to arise from either the sense (1,566) or the antisense (1,529) strand of the corresponding gene (Bertone et al., 2004).

Microarray-based evidence of ubiquitous transcriptional activity streaming from intronic genomic segments was also recently reported for *Drosophila melanogaster* (Stolc et al., 2004). In that work, it was observed that 41% of probes representing the full complement of intronic and intergenic regions of the *D. melanogaster* genome are transcriptionally active (Stolc et al., 2004). The expressed intronic and intergenic sequences are more likely to be evolutionarily conserved than nonexpressed ones, and about 15% of them appeared to be developmentally regulated (Stolc et al., 2004).

The widespread occurrence of overlapping sense/antisense transcription has also been detected following informatics analyses of the collection of complete and partial transcribed messages generated from cloned cDNA libraries and mapped to human (Shendure and Church, 2002; Yelin et al., 2003; Chen et al., 2004) or mouse (Shendure and Church, 2002; Kiyosawa et al., 2003) genomes. However, the full complement of human unspliced transcripts that map to totally intronic regions was not fully appreciated in those reports because the authors selected (i) transcripts with evidence of splicing (Shendure and Church, 2002); (ii) pairs of sense-antisense messages for which at least one exon was colinear on the genome sequence (Yelin et al., 2003); or (iii) only ESTs where both a poly(A) signal and a poly(A) tail were present (Chen et al., 2004).

To search for yet unidentified totally intronic RNA transcripts in the collection of human transcribed sequences available at GenBank, we aligned a set of 4,686,886 human ESTs from dbEST to the human genome sequence (UCSC build 34). Based on alignment coordinates, we merged overlapping sequences, defining a set of 1,713,905 EST clusters. Most of these (1,324,536 clusters, i.e., 77%) are colinear with known human genes (i.e., 28,031 RefSeqs and 130,982 complete/partial mRNAs deposited in GenBank), and 1,091,233 of these clusters overlap at least one known exon. Remarkably, an additional 233,303 clusters (18% of all EST clusters co-linear with known genes) are totally contained within intronic regions (Nakaya et al., manuscript in preparation). The 18% fraction of intronic regions in relation to the total number of transcribed regions in the human genome, as reported here for transcription arising from all tissues available in GenBank, is in agreement with the 22% fraction of intronic regions in relation to the total transcription identified in liver tissue by Bertone et al. (2004).

While a fraction of the totally intronic EST clusters may represent new exons of rare alternatively spliced forms of known genes (when mapping to the sense strand), it is conceivable that a significant number of new antisense messages are represented in this dataset. Nonetheless, direct experimental and independent evidence of transcription is required, since a fraction of these intronic ESTs might derive from contaminant genomic DNA (Sorek and Safer, 2003). The recently published data discussed above indicate that expression of natural antisense transcripts is more ubiquitous in the human genome than anticipated. Such transcripts are likely to be expressed at very low levels when compared to protein-coding transcripts (Kampa et al., 2004). This is probably the reason for the limited amount of data on the functional role of naturally occurring antisense RNAs.

INTRONIC RNA IN HUMAN DISEASES

To investigate the expression level of intronic messages in human tissues, a subset of approximately 2,000 totally intronic EST clusters identified in our informatics analysis was selected for microarray experiments, along with an additional 2,000 clusters from exonic segments of known genes. A representative cDNA clone from each selected cluster was used for construction of microarrays enriched in intronic transcripts. Hybridization of these intronic microarrays with 27 prostate tumors and corresponding adjacent normal tissue revealed that in prostate, the fraction of expressed messages arising from exonic or intronic transcripts were similar (Reis et al., 2004). Moreover, the expression levels of 23 intronic non-coding transcripts are correlated (p value ≤ 0.001) to the degree of prostate tumor differentiation. Intronic *RASSF1*, the most corre-

lated gene, is expressed both as sense and antisense messages as shown by strand-specific reverse transcriptase assay (Reis et al., 2004; Fig. 1). In addition to antisense *RASSF1*, a number of antisense intronic messages were shown by RACE-PCR and strand-specific Northern blotting to be long (0.6–1.1 kb) and unspliced (Reis et al., 2004).

Recent reports suggest that the expression pattern of short (19–25-nt) intronic micro RNAs may be associated with the pathogenesis of human B cell chronic lymphocytic leukemias (Calin et al., 2004a,b). Also, expression of an antisense transcript from a truncated form of *LUC7L* was shown to cause silencing of an adjacent α -globin gene in an inherited form of anemia (α -thalassemia) (Tufarelli et al., 2003). Nevertheless, our result is the first report implicating natural intronic antisense transcripts in a human disease (Reis et al., 2004). It is not a settled issue if in tumors the expression of intronic antisense RNAs is a true mechanism of regulation or just reflects errors in promoter recognition/transcription initiation. Nevertheless, discovery of intronic antisense messages correlated to the degree of malignancy in prostate tumors (Reis et al., 2004) has an impact on the molecular diagnosis of cancer in general, arguing for inclusion of non-coding intronic transcripts into the arsenal of tools used for molecular diagnostics, so far almost exclusively populated by microarrays that contain only exonic protein-coding transcripts.

BIOGENESIS OF INTRONIC ANTISENSE RNA

In sets of experiments combining chromatin immunoprecipitation and tiling microarray analyses, different groups have mapped, *in vivo* and in an unbiased manner, DNA binding sites of well-characterized human transcription factors (TFs) such as NF- κ B (Martone et al., 2003), Sp1, c-Myc, p53 (Cawley et al., 2004) and Creb (Euskirchen et al., 2004). Surprisingly, all these reports point to the existence of a large number of transcription factor binding sites (TFBSs) that do not lie in the canonical 5' ends of well-characterized protein-coding genes, but instead map within or near the 3' exons of known genes. In two of these reports, a large fraction of the experimentally determined TFBSs was positioned upstream to transcriptionally active intronic regions (Cawley et al., 2004; Euskirchen et al., 2004), some of which generate transcripts oriented antisense to the known gene (Cawley et al., 2004). Noteworthy, in the study from Cawley et al., (2004) was the finding that comparable fractions of protein-coding genes and non-coding RNAs having an experimentally determined upstream TFBS region were regulated following exposure to retinoic acid. Taken together, these results suggest that at least a fraction of the intronic antisense transcripts that compose the human transcriptome are expressed and regulated by similar molecular mechanisms as protein-coding genes.

The observation that non-coding RNA transcription is regulated by TFs such as Sp1, c-Myc, p53, which are common to protein-coding genes, points to the recruitment of RNA polymerase II (Pol II) to form the transcription complex of intronic messages. In fact, *in vitro* transcription analysis has shown that RNA Pol II recognizes an Inr element in the first intron of eukaryotic initiation factor 2a (eIF2a) and generates a transcript in the opposite direction of eIF-2a mRNA (Silverman et al., 1992). RNA Pol II was also shown to direct transcription of polyadenylated and nonpolyadenylated antisense-oriented transcripts that initiate in the first intron and span the first exon in the C-MYC and N-MYC loci (Nepveu and Marcu, 1986; Krystal et al., 1990). To illustrate that the transcription of non-coding RNAs by RNA Pol II is not restricted to intronic antisense messages, it was also recently shown that micro RNAs are mainly, if not exclusively, transcribed by RNA Pol II (Lee et al., 2004).

NON-CODING INTRONIC RNA AND REGULATION OF GENE EXPRESSION

Non-coding RNAs that are derived from the introns of protein-coding genes and non-protein-coding genes represent a fundamental advance in the genetic operating system of higher organisms, defining how the trajectories of differentiation and development and the differences among individuals and species are genomically programmed (Mattick, 2004). A number of mechanisms of gene expression regulation have been proposed to involve double-stranded RNA (dsRNA) in eukaryotes (Mattick, 2003). While most of these models

FIG. 1. Experimental demonstration of an intronic antisense *RASSF1* transcript. *RASSF1* intronic sequences are drawn as light gray thin boxes, exonic sequences as dark gray boxes or arrows, and the intronic antisense *RASSF1* transcript (asRNA) as dashed arrows. Antisense orientation of the intronic *RASSF1* transcript was confirmed using strand-specific reverse transcription (SSRT, middle dotted-line panel), followed by PCR amplification (lower panels) as described (Reis et al., 2004). SSRT did generate either forward (light gray continuous line) or reverse (light gray dashed line) cDNA templates. Briefly, 100 ng of DNase-treated poly A(+) RNA extracted from a prostate tumor sample was used. Reverse transcription was primed by specific oligonucleotides (0.4–0.9 μ M) that were complementary to either the sense (RT-R gray half-arrow) or antisense (RT-F gray half-arrow) intronic region of *RASSF1* unspliced pre-mRNA transcript. RNA and primers were denatured at 70°C for 15 min, and SSRT reactions were carried at 50°C to avoid RNA secondary structure effects. Two SSRT negative controls were performed in parallel to monitor either the absence of genomic DNA contaminants (RT-minus Control) or the absence of reverse transcription randomly primed with RNA secondary structures (primer-minus Control). Subsequent PCR reactions (at right lower panel) were performed using a pair of primers (black half arrows) designed internally to both cDNA templates (forward or reverse), in a segment where both templates overlap. Using the same templates resulting from the SSRT reactions, an extra PCR control was performed (at left lower panel) where another pair of primers (black half arrows) designed to span only the intronic sense reverse transcription product was used. Pairs of short thin light gray lines (either solid or dashed) are representing PCR-generated products. Transcription factor binding sites (TFBS), transcription factor (TF) and RNA polymerase II (Pol II) are indicated in the upper panel.

describe dsRNAs comprising at least one exonic partner, we discuss below possible mechanisms where long antisense transcripts overlapping intronic regions might be involved.

Micro-RNAs and RNA interference

Micro-RNAs are endogenous ~22-nt RNAs that can play important regulatory roles in animals and plants by targeting mRNAs for cleavage or translational repression (Bartel, 2004). Most miRNA genes come from regions of the genome quite distant from previously annotated genes, implying that they derive from independent transcription units (Lagos-Quintana et al., 2001). Only a minority of these miRNAs originates from intronic regions of genes (Bartel, 2004); those miRNAs encoded in intronic regions have the same orientation as the predicted mRNAs, suggesting that they are processed from intronic segments following splicing of the premature mRNA (Bartel, 2004). This arrangement provides a convenient mechanism for the coordinated expression of a miRNA and a protein-encoding transcript (Bartel, 2004). However, it remains unclear if long intronic transcripts have any effect on the biogenesis and/or regulation of micro RNAs originating in the same loci.

Another known mechanism of regulation based on double-stranded RNA intermediates is RNA interference (RNAi), the oldest and most ubiquitous antiviral system exclusively found in eukaryotes (Brantl, 2002). When exposed to foreign genetic material (RNA or DNA) many organisms mount a highly specific RNAi reaction in the cytoplasm, generating a sequence-specific double-stranded RNA (dsRNA), to counter attack and silence the invading nucleic-acid sequence by triggering its specific degradation, a mechanism that has been extensively reviewed elsewhere (Brantl, 2002; Mello and Conte, 2004). In mammals, long dsRNAs (longer than 30 bp) introduced in the cytoplasm activate a general interferon-like response that stalls translation and activates non-specific mRNA degradation in invaded cells rather than triggering sequence-specific RNAi interference (Elbashir et al., 2001). To avoid launching such unspecific cellular response, we speculate that in humans the natural long intronic antisense RNA transcripts might be retained and predominantly reside in the nucleus rather than in the cytoplasm, where they could find their matching pre-mRNA pair and exert their sequence-specific biological roles, as discussed below. In fact, eukaryotic cells retain long (>100 bp) dsRNA in the nucleus by a process involving mRNA editing by nuclear adenosine deaminases (ADARs), followed by binding of these hyper-edited RNAs to 54-kD nuclear RNA-binding protein (p54nrb) (Kumar and Carmichael, 1997; Zhang and Carmichael, 2001; Peters et al., 2003). ADAR enzymes catalyze the hydrolytic deamination of adenosines to inosines on each strand of long perfect RNA duplexes, a mechanism that has been reviewed elsewhere (Bass 2002). However, this form of editing has not yet been demonstrated for any mammalian dsRNAs, so the generality of this mechanism remains to be determined.

We speculate that long intronic antisense RNA messages may modulate gene expression by intervening in different steps of mRNA transcription and/or processing (Fig. 2).

Alternative splicing and antisense RNA

We postulate here that naturally expressed non-coding RNAs with antisense direction relative to introns of protein-coding genes could be involved in alternative splicing of human genes by pairing with the sense strand of pre-mRNAs. Correct excision of introns from a pre-mRNA and the joining of exons are directed

FIG. 2. Proposed mechanisms of regulation of gene expression involving intronic antisense RNA transcripts. (A) Transcription of protein-coding pre-mRNA (target gene) and non-coding intronic antisense message (asRNA). (B,C) Putative roles of intronic antisense RNA transcripts in modulating the expression pattern of target genes: modulation of the abundance of alternative splicing isoforms (B); modulation of the methylation status of upstream regulatory regions of target genes (left) or modulation of steady-state RNA levels by transcriptional interference (right) (C). Dark gray boxes and arrows represent sense orientation and dashed arrows, antisense orientation. Intronic sequences are drawn as thin boxes, exonic sequences as thick boxes or arrows, and the alternatively spliced exon as a hatched box. Intronic antisense RNA transcript is represented as dashed arrows. Transcription factor binding sites (TFBS), transcription factor (TF), RNA polymerase II (Pol II), spliceosome complex (SC), and DNA methyltransferase (DMt) are indicated.

INTRONIC ANTISENSE RNA

by spliceosomes that recognize splice sites at intron/exons boundaries or sequence elements located inside intronic or exonic regions of pre-mRNA (Black, 2003). Splicing is known to be modulated by intronic splicing enhancer (ISE) and silencer (ISS) elements (Akker et al., 2001) through binding of regulatory factors to ISE and ISS that favor or block formation of the spliceosome (Bruno et al., 2004). These can be located at hundreds of nucleotides away from the regulated exon (Black, 2003).

Although no direct evidence has been obtained relating long intronic antisense transcripts to regulation of splicing, some experiments performed *in vitro* point in this direction. Synthetic antisense oligonucleotides targeted at introns of mutant pre-mRNAs from *HBB* or *CFTR* presenting aberrant splice sites did restore correct splicing and translation (Sierakowska et al., 1996; Friedman et al., 1999; Lacerra et al., 2000). An example of targeting intronic silencing sequence (ISS) elements with antisense synthetic oligonucleotides is found in the work of Bruno et al. (2004) with the fibroblast growth factor receptor 1 (*FGFR1*) gene. This gene has an alternatively spliced alpha-exon, which is aberrantly skipped in human glioblastomas. The use of antisense synthetic oligonucleotides that are complementary to either upstream or downstream ISS elements flanking the alpha-exon, increased alpha-exon inclusion from 10% up to 70% *in vivo* (Bruno et al., 2004). The effect was dose dependent, sequence specific and reproducible in several human cell lines (Bruno et al., 2004). In addition, targeting of the intronic splicing enhancer (ISE) of *SMN* pre-mRNA with synthetic antisense oligonucleotides produced exon 7 exclusion in *SMN* mRNA (Miyaso et al., 2003). These data demonstrate that intronic elements are important for proper spliceosome assembly and may be affected by binding of antisense nucleotides.

Finally, a clear example of naturally occurring RNA masking is the inhibition of alternative splicing of one of two functionally antagonistic splice forms of the thyroid hormone receptor *ErbAa* mRNA (*ErbAa2* not *ErbAa1*) by expression of the antisense transcript, *Rev-ErbAa* that overlaps only *ErbAa2* (Munroe and Lazar, 1991; Hastings et al., 1997, 2000).

Taken together, the results discussed above suggest that natural intronic antisense RNAs could use a similar molecular machinery to regulate, *in vivo*, alternative splicing patterns of human genes, modulating the accessibility of the spliceosome complex to regulatory sequences located within intronic segments of pre-mRNAs.

Methylation status of target genes

Chromatin remodeling and methylation are processes that modify the transcriptional regulation of gene expression through rearrangement of nucleosomes or by chemical modifications on nucleosides of DNA (Huang et al., 2003).

It has been shown that Polycomb-group proteins play crucial roles in chromatin rearrangements via histone deacetylation (van der Vlag and Otte, 1999; Gebuhr et al., 2000). In these Polycomb-group proteins, in MOF histone acetyltransferase (Akhtar et al., 2000) and in other proteins that are important in similar remodeling mechanisms, a conserved domain has been identified, and described to have RNA binding properties, called chromodomain (Akhtar et al., 2000; Jones et al., 2000). Importantly, Akhtar et al. (2000) have suggested that in *Drosophila* X chromosome male dosage compensation phenomenon, the MOF histone acetyltransferase chromodomain interacts *in vivo* with roX, a non-coding RNA, which may contribute to the integration of MOF into the dosage compensation complex (DCC) at the male X chromosome.

Wassenegger et al. demonstrated in tobacco plants the occurrence of RNA-directed *de novo* methylation of specific sequences of genome-integrated copies of potato spindle tuber viroid (PSTVd) in response to an artificial PSTVd RNA inoculation, demonstrating that a mechanism of *de novo* methylation of genes might exist that can be induced and targeted in a sequence-specific manner by their own mRNA (Wassenegger et al., 1994). This phenomenon was predicted to occur in all higher eukaryotes (Wassenegger et al., 1994); however, there were few studies of X chromosome inactivation and genomic imprinting, generally associated with DNA methylation/demethylation, suggesting an RNA-dependent mechanism (Lee et al., 1999; Sleutels et al., 2000, 2002).

In the last two years direct evidence for antisense RNA dependent methylation and gene silencing has been obtained. For *HBA2* gene which is silenced and its CpG island becomes completely methylated, this

INTRONIC ANTISENSE RNA

phenomenon is mediated by an antisense transcript *LUC7L* expressed in the *HBA2* locus in a patient affected by alpha-thalassemia (Tufarelli et al., 2003); for *SPHK1*, its CpG island is a target for tissue-specific DNA demethylation directed by a partially intronic endogenous non-coding RNA transcript *KHPS1a* expressed from the antisense strand at the 5'-end of the gene (Imamura et al., 2004).

While DNA-RNA complementarity seems to be a common feature for *cis*-acting antisense RNA-directed DNA methylation, the minimum size of the complementary region is still controversial. In a recent work, Pelisser and Wassenegger (2000) demonstrated that such mechanism occurs even when the region of complementarity encompasses a length of only 30 nt. Imamura et al. (2004) showed that for *SPHK1*, different antisense constructs either directly overlapping the tissue-dependent differentially methylated region (T-DMR) or binding only at the vicinity of the T-DMR were similarly able to change *KHPS1* CpG island pattern of methylation. This observation indicates that *cis*-acting antisense-mediated modulation of DNA methylation extends beyond the region of RNA-DNA duplex.

Transcriptional interference

Two different RNAs that are transcribed concomitantly from opposite strands in the same genomic *locus* can exhibit transcriptional interference due to transcriptional collision, a phenomenon that was investigated by Prescott and Proudfoot (2002) using *Saccharomyces cerevisiae* as a model. In this experiment, the steady-state level of mRNA was reduced by rearranging two genes in a way that the direction of the transcription from both genes is convergent (Prescott and Proudfoot, 2002). As soon as the two transcriptional units begin to overlap, elongation mediated by RNA polymerase II is restricted, resulting in a severe reduction of mRNA accumulation at the steady-state (Prescott and Proudfoot, 2002). Evidence of transcriptional interference has also been obtained for antisense partners at the insulin-like growth factor type-2 receptor/antisense Igf2r RNA (Igf2r/AIR) (Wutz et al., 1997), at the eukaryotic initiation factor 2a (eIF2a) (Silverman et al., 1992; Noguchi et al., 1994) and at alpha 1 (I) collagen (Farrell and Lukens, 1995) loci. Thus, expression of intronic natural antisense transcripts could be similarly regulated through a process of transcriptional interference.

CONCLUSION

As the amount of evidence accumulates that a significant fraction of transcription occurs in intronic non protein-coding regions of the genome, so are the questions about the role of these transcripts in the regulation of gene expression. These include developmental, tissue, and pathway specificity of intronic transcription, their involvement in human pathogenesis and their precise mechanism of modulating alternative splicing and protein synthesis. To address these questions, the availability of novel and sensitive methods to detect the expression of rare antisense intronic messages, such as strand-specific tiling microarrays (Bertone et al., 2004; Kampa et al., 2004) will be particularly useful. Development of *in vitro* models to test molecular and phenotypic modifications associated with subtle changes in the steady-state levels of intronic transcripts will also be required. In this matter, we expect that nuclear overexpression in cultured cells of intronic antisense transcripts driven by regulated promoters will be important to evaluate the effects of antisense intronic transcription in DNA methylation status, mRNA accumulation, alternative splicing and protein translation of target genes. We also anticipate that overexpression or silencing of intronic antisense transcripts in transfected cells will be important tools to assess possible effects of antisense RNA on phenotype, which may provide further clues to understand the complex network of RNA-dependent gene expression regulation and its involvement in human diseases.

ACKNOWLEDGMENTS

This work was financed by fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and by a grant and fellowships from Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP).

REFERENCES

- AKHTAR, A., ZINK, D., and BECKER, P.B. (2000). Chromodomains are protein–RNA interaction modules. *Nature* **407**, 405–409.
- AKKER, S.A., SMITH, P.J., and CHEW, S.L. (2001). Nuclear post-transcriptional control of gene expression. *J Mol Endocrinol* **27**, 123–131.
- BARTEL, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297.
- BASS, B.L. (2002). RNA editing by adenosine deaminases that act on RNA. *Annu Rev Biochem* **71**, 817–846.
- BERTONE, P., STOLC, V., ROYCE, T.E., et al. (2004). Global identification of human transcribed sequences with genome tiling arrays. *Science* **306**, 2242–2246.
- BLACK, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291–336.
- BRANTL, S. (2002). Antisense-RNA regulation and RNA interference. *Biochim Biophys Acta* **1575**, 15–25.
- BRUNO, I.G., JIN, W., and COTE, G.J. (2004). Correction of aberrant FGFR1 alternative RNA splicing through targeting of intronic regulatory elements. *Hum Mol Genet* **13**, 2409–2420.
- CALIN, G.A., LIU, C.G., SEVIGNANI, C., et al. (2004a). MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci USA* **101**, 11755–11760.
- CALIN, G.A., SEVIGNANI, C., DUMITRU, C.D., et al. (2004b). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci USA* **101**, 2999–3004.
- CAWLEY, S., BEKIRANOV, S., NG, H.H., et al. (2004). Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs. *Cell* **116**, 499–509.
- CHEN, J., SUN, M., KENT, W.J., et al. (2004). Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res* **32**, 4812–4820.
- ELBASHIR, S.M., HARBORTH, J., LENDECKEL, W., et al. (2001). Duplexes of 21–nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494–498.
- EUSKIRCHEN, G., ROYCE, T.E., BERTONE, P., et al. (2004). CREB binds to multiple loci on human chromosome 22. *Mol Cell Biol* **24**, 3804–3814.
- FARRELL, C.M., and LUKENS, L.N. (1995). Naturally occurring antisense transcripts are present in chick embryo chondrocytes simultaneously with the down-regulation of the alpha 1 (I) collagen gene. *J Biol Chem* **270**, 3400–3408.
- FRIEDMAN, K.J., KOLE, J., COHN, J.A., et al. (1999). Correction of aberrant splicing of the cystic fibrosis transmembrane conductance regulator (CFTR) gene by antisense oligonucleotides. *J Biol Chem* **274**, 36193–36199.
- GEBUHR, T.C., BULTMAN, S.J., and MAGNUSON, T. (2000). Pc-G/trx-G and the SWI/SNF connection: developmental gene regulation through chromatin remodeling. *Genesis* **26**, 189–197.
- HASTINGS, M.L., INGLE, H.A., LAZAR, M.A., et al. (2000). Post-transcriptional regulation of thyroid hormone receptor expression by cis-acting sequences and a naturally occurring antisense RNA. *J Biol Chem* **275**, 11507–11513.
- HASTINGS, M.L., MILCAREK, C., MARTINCIC, K., et al. (1997). Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res* **25**, 4296–4300.
- HUANG, C., SLOAN, E.A. and BOERKOEL, C.F. (2003). Chromatin remodeling and human disease. *Curr Opin Genet Dev* **13**, 246–252.
- IMAMURA, T., YAMAMOTO, S., OHGANE, J., et al. (2004). Non-coding RNA directed DNA demethylation of Sphk1 CpG island. *Biochem Biophys Res Commun* **322**, 593–600.
- JONES, D.O., COWELL, I.G. and SINGH, P.B. (2000). Mammalian chromodomain proteins: their role in genome organisation and expression. *Bioessays* **22**, 124–137.
- KAMPA, D., CHENG, J., KAPRANOV, P., et al. (2004). Novel RNAs Identified From an In-Depth Analysis of the Transcriptome of Human Chromosomes 21 and 22. *Genome Res* **14**, 331–342.
- KAPRANOV, P., CAWLEY, S.E., DRENKOW, J., et al. (2002). Large-scale transcriptional activity in chromosomes 21 and 22. *Science* **296**, 916–919.
- KIYOSAWA, H., YAMANAKA, I., OSATO, N., et al. (2003). Antisense transcripts with FANTOM2 clone set and their implications for gene regulation. *Genome Res* **13**, 1324–1334.
- KRYSTAL, G.W., ARMSTRONG, B.C. and BATTEY, J.F. (1990). N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. *Mol Cell Biol* **10**, 4180–4191.
- KUMAR, M. and CARMICHAEL, G.G. (1997). Nuclear antisense RNA induces extensive adenosine modifications and nuclear retention of target transcripts. *Proc Natl Acad Sci USA* **94**, 3542–3547.
- LACERRA, G., SIERAKOWSKA, H., CARESTIA, C., et al. (2000). Restoration of hemoglobin A synthesis in erythroid cells from peripheral blood of thalassemic patients. *Proc Natl Acad Sci USA* **97**, 9591–9596.

- LAGOS-QUINTANA, M., RAUHUT, R., LENDECKEL, W., et al. (2001). Identification of novel genes coding for small expressed RNAs. *Science* **294**, 853–858.
- LANDER, E.S., LINTON, L.M., BIRREN, B., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921.
- LEE, J.T., DAVIDOW, L.S. and WARSHAWSKY, D. (1999). Tsix, a gene antisense to Xist at the X-inactivation centre. *Nat Genet* **21**, 400–404.
- LEE, Y., KIM, M., HAN, J., et al. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J* **23**, 4051–4060.
- MARTONE, R., EUSKIRCHEN, G., BERTONE, P., et al. (2003). Distribution of NF-kappaB-binding sites across human chromosome 22. *Proc Natl Acad Sci USA* **100**, 12247–12252.
- MATTICK, J.S. (2003). Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms. *Bioessays* **25**, 930–939.
- MATTICK, J.S. (2004). RNA regulation: a new genetics? *Nat. Rev. Genet.* **5**, 316–323.
- MELLO, C.C. and CONTE, D., JR. (2004). Revealing the world of RNA interference. *Nature* **431**, 338–342.
- MIYASO, H., OKUMURA, M., KONDO, S., et al. (2003). An intronic splicing enhancer element in survival motor neuron (SMN) pre-mRNA. *J Biol Chem* **278**, 15825–15831.
- MUNROE, S.H. and LAZAR, M.A. (1991). Inhibition of c-erbA mRNA splicing by a naturally occurring antisense RNA. *J Biol Chem* **266**, 22083–22086.
- NEPVEU, A. and MARCU, K.B. (1986). Intragenic pausing and anti-sense transcription within the murine c-myc locus. *EMBO J* **5**, 2859–2865.
- NOGUCHI, M., MIYAMOTO, S., SILVERMAN, T.A., et al. (1994). Characterization of an antisense Inr element in the eIF-2 alpha gene. *J Biol Chem* **269**, 29161–29167.
- PELISSIER, T. and WASSENEGGER, M. (2000). A DNA target of 30 bp is sufficient for RNA-directed DNA methylation. *RNA* **6**, 55–65.
- PETERS, N.T., ROHRBACH, J.A., ZALEWSKI, B.A., et al. (2003). RNA editing and regulation of *Drosophila* 4f-rnp expression by sas-10 antisense readthrough mRNA transcripts. *RNA* **9**, 698–710.
- PRESCOTT, E.M. and PROUDFOOT, N.J. (2002). Transcriptional collision between convergent genes in budding yeast. *Proc Natl Acad Sci USA* **99**, 8796–8801.
- REIS, E.M., NAKAYA, H.I., LOURO, R., et al. (2004). Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. *Oncogene* **23**, 6684–6692.
- RINN, J.L., EUSKIRCHEN, G., BERTONE, P., et al. (2003). The transcriptional activity of human Chromosome 22. *Genes Dev* **17**, 529–540.
- SHENDURE, J. and CHURCH, G.M. (2002). Computational discovery of sense-antisense transcription in the human and mouse genomes. *Genome Biol* **3**, research0044.1–research0044.14.
- SHOEMAKER, D.D., SCHADT, E.E., ARMOUR, C.D., et al. (2001). Experimental annotation of the human genome using microarray technology. *Nature* **409**, 922–927.
- SIERAKOWSKA, H., SAMBADE, M.J., AGRAWAL, S., et al. (1996). Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc Natl Acad Sci USA* **93**, 12840–12844.
- SILVERMAN, T.A., NOGUCHI, M. and SAFER, B. (1992). Role of sequences within the first intron in the regulation of expression of eukaryotic initiation factor 2 alpha. *J Biol Chem* **267**, 9738–9742.
- SLEUTELS, F., BARLOW, D.P. and LYLE, R. (2000). The uniqueness of the imprinting mechanism. *Curr Opin Genet Dev* **10**, 229–233.
- SLEUTELS, F., ZWART, R. and BARLOW, D.P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature* **415**, 810–813.
- SOREK, R., and SAFER, H.M. (2003). A novel algorithm for computational identification of contaminated EST libraries. *Nucleic Acids Res* **31**, 1067–1074.
- STOLC, V., GAUHAR, Z., MASON, C., et al. (2004). A gene expression map for the euchromatic genome of *Drosophila melanogaster*. *Science* **306**, 655–660.
- TUFARELLI, C., STANLEY, J.A., GARRICK, D., et al. (2003). Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet* **34**, 157–165.
- VAN DER VLAG, J., and OTTE, A.P. (1999). Transcriptional repression mediated by the human polycomb-group protein EED involves histone deacetylation. *Nat Genet* **23**, 474–478.
- VENTER, J.C., ADAMS, M.D., MYERS, E.W., et al. (2001). The sequence of the human genome. *Science* **291**, 1304–1351.
- WASSENEGGER, M., HEIMES, S., RIEDEL, L., et al. (1994). RNA-directed de novo methylation of genomic sequences in plants. *Cell* **76**, 567–576.

- WUTZ, A., SMRZKA, O.W., SCHWEIFER, N., et al. (1997). Imprinted expression of the Igf2r gene depends on an intronic CpG island. *Nature* **389**, 745–749.
- YELIN, R., DAHARY, D., SOREK, R., et al. (2003). Widespread occurrence of antisense transcription in the human genome. *Nat Biotechnol* **21**, 379–386.
- ZHANG, Z., and CARMICHAEL, G.G. (2001). The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **106**, 465–475.

Address reprint requests to:
Sergio Verjovski-Almeida
Departamento de Bioquímica
Instituto de Química
Universidade de São Paulo
05508–900 São Paulo, SP, Brasil

E-mail: verjo@iq.usp.br

Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription

Helder I. Nakaya, Paulo P. Amaral, Rodrigo Louro, André Lopes, Angela A. Fachel, Yuri B. Moreira, Tarik A. El-Jundi, Aline M. da Silva, Eduardo M. Reis and Sergio Verjovski-Almeida*

Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo, SP, Brazil

*Corresponding author

Dr. Sergio Verjovski-Almeida

Tel: +55-11-3091-2173

Fax: +55-11-3091-2186

e-mail: hnakaya@iq.usp.br

e-mail: ppamaral@iq.usp.br

e-mail: rolouro@iq.usp.br

e-mail: alopes@iq.usp.br

e-mail: afachel@iq.usp.br

e-mail: ymoreira@iq.usp.br

e-mail: tarikj@iq.usp.br

e-mail: almsilva@iq.usp.br

e-mail: emreis@iq.usp.br

e-mail: verjo@iq.usp.br

Running head: Intronic expression patterns in the human genome

Abstract

Background: RNAs transcribed from intronic regions of genes are involved in a number of processes related to post-transcriptional control of gene expression. However, the complement of human genes in which introns are transcribed, the number of intronic transcriptional units and their tissue expression patterns are not known.

Results: Survey of the mRNA and EST public databases revealed more than 55,000 Totally Intronic Noncoding (TIN) RNAs transcribed from the introns of 11,679 protein-coding genes; the latter correspond to 74% of all unique RefSeq genes. Guided by this information, we designed an oligoarray platform containing ~20,000 probes for randomly-selected intronic transcripts and the corresponding 7,464 protein-coding genes. Using this platform, we identified exonic and intronic expression signatures for human liver, prostate and kidney. For each tissue, the most highly expressed antisense TIN RNAs were transcribed from the introns of protein-coding genes enriched in the “Regulation of transcription” GO category. RNA Polymerase II inhibition resulted in increased expression of a fraction of the intronic RNAs in cell cultures, suggesting that other RNA polymerases may be involved in their biosynthesis; “Regulation of transcription” was again enriched among the genes that had up-regulated intronic messages. A subset of intronic and protein-coding signatures transcribed from the same genomic *loci* has correlated expression patterns, suggesting that intronic RNAs regulate the abundance or the pattern of exon usage in protein-coding messages.

Conclusions: We have identified diverse intronic RNA expression patterns, suggesting that they have regulatory roles. This gene-oriented approach should permit further comparative analysis of intronic transcription under various physiological and pathological conditions, thus advancing current knowledge about the biological functions of these noncoding RNAs.

Background

The 5 million expressed sequence tags (ESTs) deposited into public sequence databases probably constitute the best representation of the human transcriptome. Human EST data have been extensively used to identify novel genes *in silico* [1, 2] and novel exons of protein-coding genes [3-6]. Informatics analyses of the EST collection mapped to the human genome have also shown that the occurrence of overlapping sense/antisense transcription is widespread [7-9]. However, the complement of unspliced human transcripts that map exclusively to introns was not appreciated in those reports because the authors selected (i) transcripts with evidence of splicing [7]; (ii) pairs of sense-antisense messages for which at least one exon was colinear on the genome sequence [8]; or (iii) only ESTs where both a polyadenylation signal and a poly(A) tail were present [9].

A detailed analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs revealed that 15,815 are noncoding RNAs (ncRNAs), of which 71% are unspliced/single exon, indicating that ncRNA is a major component of the transcriptome [10]. The recent finishing and detailed annotation of the euchromatic sequence of the human genome has identified 20,000–25,000 protein-coding genes [11]; however, noncoding messages were not assessed [11]. Extrapolation from the numbers for chromosome 7 leads to an estimate of 3,700 human ncRNAs [12], and two databases of human and murine noncoding RNAs are available [13, 14]. Nevertheless, there has been no comprehensive count and mapping of human noncoding RNAs.

Examples of long (0.6–2 kb) intronic noncoding RNAs involved in different biological processes are described in the literature; they participate in the transcriptional or post-transcriptional control of gene expression [15, 16], and in the regulation of exon-skipping [17] and intron retention [18]. In addition, microarray experiments performed by our group have revealed a set of long intronic ncRNAs the expression of which is correlated to the degree of malignancy in prostate cancer [19]. Introns are also the sources of short ncRNAs that have been characterized as microRNAs [20] and small nucleolar RNAs [21]. Biogenesis and function are better understood for microRNAs than for other ncRNAs; they may regulate as many as one-third of human genes [20], and tissue-specific expression signatures have been identified in different human cancers [22]. However, the complement and biological function of most of the complex and diverse ncRNA output, both the short and the long ncRNAs, remain to be determined.

Different types of noncoding RNA genes can be transcribed by either RNA polymerase I, II or III [15]. Recently, a fourth nuclear RNA polymerase (RNAP) consisting of an isoform of the human single-polypeptide mitochondrial RNAP, named spRNAP IV, was found to transcribe a small fraction of mRNAs in human cells [23]. Surprisingly, α -amanitin

up-regulates the transcription of protein-coding mRNAs by this polymerase [23]. The role of spRNAP IV in the transcription of ncRNAs has not been investigated.

Here we report a search for hitherto unidentified exclusively intronic unspliced RNA transcripts in the collection of transcribed human sequences available at GenBank. The characterization comprises the identification and distribution analysis of 55,000 long intronic ncRNAs over the introns of protein-coding genes and the detection of a higher frequency of alternatively-spliced exons of genes that have intronic transcription. An oligoarray with 44,000 elements representing exons of protein-coding genes and the corresponding actively-transcribed introns was employed to assess intronic transcription in different human tissues. Robust tissue signatures of exonic and intronic expression were detected in human kidney, prostate and liver. We found that in each tissue, the most highly-expressed exclusively intronic antisense RNAs were transcribed from a group of protein-coding genes that is significantly enriched in the “Regulation of transcription” GO category. A subset of partially intronic antisense ncRNAs and the corresponding overlapped protein-coding exons showed a correlated pattern of tissue expression, indicating that intronic RNAs may have a role in regulating abundance or alternative exon-splicing events. Finally, we found that a significant fraction of wholly or partially intronic ncRNAs is insensitive to RNA Polymerase II inhibition by α -amanitin, and another fraction is even up-regulated when RNA polymerase II transcription is blocked, suggesting that a portion of long ncRNAs may be transcribed by spRNAP IV. We conclude that oligoarray-based gene-oriented analysis of intronic transcription is a powerful tool for identifying novel potentially functional noncoding RNAs.

Results

Defining a comprehensive reference dataset of spliced protein-coding genes

In order to analyze the complex distribution of transcriptionally active regions on a genome-wide scale, we started by mapping the set of well-annotated 22,458 RefSeq transcripts to the human genome sequence. We excluded 1,184 unspliced RefSeq, and 601 RefSeq that were wholly intronic to another RefSeq. When the spliced RefSeq transcripts mapping to the same *locus* were merged, we identified a set of 15,783 non-redundant spliced RefSeq units. Thus, a total of 4,890 RefSeq representing isoforms of the same genes were merged into these units. In addition, the GenBank mRNA sequence dataset was mapped to the genome in order to document splice variants present in that set but not in the non-redundant RefSeq data. For this purpose, 161,993 human mRNAs from GenBank were mapped to the human genome, as described under Methods. Initially, they were clustered into a total of 45,137 transcriptional units mapping to unique *loci* in the genome (Table 1). A detailed analysis of

the mapping coordinates of these mRNA clusters with respect to the non-redundant RefSeq dataset revealed that 11,361 spliced and unspliced clusters mapped outside the non-redundant RefSeq dataset, representing less well-characterized human transcripts. As expected, most of the mRNA clusters (14,575) were spliced and mapped to exons of RefSeq genes in the sense direction (Table 1). In addition, 2,559 spliced mRNA clusters mapped in the antisense direction with respect to the non-redundant RefSeq dataset, suggesting that 16% of the RefSeq genes have spliced natural antisense transcripts that overlap at least one of their exons. Among these antisense messages, 1,414 are already annotated as RefSeq transcripts. Such genomic organization of sense-antisense gene pairs seems to have been conserved throughout vertebrate evolution [7, 8, 24, 25]. When the unspliced mRNA clusters were included, we found a total of 4,231 antisense messages with overlaps to exons in RefSeq genes, indicating that as many as 27% of the latter have antisense counterparts. This is in line with the prediction that over 20% of human transcripts might form sense-antisense pairs [9].

Most interestingly, we found 7,507 spliced and unspliced mRNA clusters that are entirely intronic to the non-redundant RefSeq genes. While 5,002 (67%) of these mapped in the sense direction and may represent new exons of the corresponding genes, 2,505 (33%) mapped exclusively to the introns of RefSeq genes in the antisense direction. Owing to the fragmented nature of the transcript data in GenBank, some of these intronic antisense messages may originate from the 3' or 5' ends of overlapping sense-antisense transcripts of adjacent genes. However, most of them could represent independent antisense transcriptional units, which became more evident when data from the public EST repository were taken into account, as described below.

Identification of long unspliced totally intronic transcripts

We performed an extensive search for evidence of intronic transcription in the human dbEST collection (GenBank) comprising 5,340,464 ESTs. Ambiguously-mapping EST sequences were filtered as described under Methods, then the genomic coordinates of overlapping EST sequences were used to merge 4,762,523 human ESTs into a set of 332,946 non-redundant EST clusters (Table 2). In order to avoid sequences that may have been derived from genomic contamination in the EST dataset, 210,181 EST singlets were excluded from further analyses; so only 34,398 spliced and 88,367 unspliced EST clusters were considered (Table 2). For each of these clusters, a consensus contig sequence was derived from the aligned genomic sequence (Figure 1). As expected, most ESTs (3,616,644) were grouped into 16,241 spliced EST contigs mapping to exons of the RefSeq reference dataset (Table 2). In addition, a small number of spliced EST clusters mapped to introns of the RefSeq genes. They may constitute fragments of novel exons in these genes,

since the median exon length in these spliced EST contigs is 233 nt, similar to the median length of exons in the RefSeq reference dataset (141 nt).

The most interesting finding was that 55,139 unspliced EST contigs formed by grouping 190,583 ESTs mapped entirely to the introns of genes in the RefSeq dataset (Table 2). A marked feature of these unspliced wholly-intronic EST contigs is their low protein-coding potential; *in silico* analysis of the coding potential using the normalized ESTScan2 score [26] predicted that 98% of them are probably noncoding transcripts, supporting the idea that they represent a separate class of noncoding RNAs. To check whether ESTScan2 predicted the coding potential of such a fragmented sequence dataset correctly, we created a virtual dataset *in silico* composed of 55,139 exonic fragments from RefSeq genes with exactly the same lengths as the 55,139 wholly-intronic EST contigs. ESTScan2 correctly predicted that 70% of these *in silico*-generated virtual exonic fragments have coding potential. This supports the inference that since only a very few (~2%) of the wholly-intronic EST contigs are predicted by ESTScan2 to have a protein-coding potential, most of the RNAs in this class (98%) are indeed noncoding messages.

Inspection of the length distribution curves (Figure 1) of the wholly-intronic EST contigs reveals messages with lengths well over 1,000 nt. The median length (573 nt) is 4.1 times greater than the median length of exons (141 nt) in the RefSeq reference dataset. On the basis of these findings, we call these transcriptional units long Totally Intronic Noncoding (TIN) transcripts.

Most mammalian snoRNAs [21] and a large fraction of microRNAs [27] are derived from introns in protein-coding and noncoding genes transcribed by RNAP II. To address the possibility that some of the TIN transcripts are the sources of these known small RNAs, we compared the human genomic coordinates of TIN sequences to those of 346 snoRNAs (<http://www-snorna.biotech.fr/>) and 383 microRNAs [28]. We found that 98 snoRNA or microRNA transcripts (14%) mapped to 86 TIN EST contigs, which may well be the sources of these small RNAs. The 86 TIN EST contigs comprise a very small portion (0.2%) of the TIN transcript dataset. We postulate that the large remaining set could be the source of new snoRNAs and microRNAs as well as of new types of ncRNAs.

Identification of long unspliced partially-intronic transcripts

A set of unspliced Partially Intronic Noncoding (PIN) EST contigs was identified. A PIN contig was defined as an unspliced contig that overlaps an exon of a RefSeq gene and extends at least 30 bases over both ends of the exon (Figure 1). In total, 12,592 PIN EST contigs (median length 719 nt) were identified, of which 5,992 (48%) appeared to be antisense to the corresponding RefSeq gene. This was determined by matching the PIN contig sequences to ESTs from high-quality directionally-cloned EST libraries [7] to

Transcriptionally Active Regions (TARs) in whole-genome strand specific tiling arrays [29], and to the publicly-available unspliced full-length mRNA dataset from GenBank. An estimated 90% of PIN transcripts have no or limited protein-coding potential as determined by ESTScan2 analysis.

Most RefSeq genes have intronic transcription

Overall, we found that at least 11,679 RefSeq genes, corresponding to 74% of all spliced human genes in the reference dataset, have transcriptionally active introns to which TIN or PIN EST contigs were mapped. If we were to consider TIN or PIN EST singlets, the fraction of RefSeq genes with intronic transcription would increase to 86% of all RefSeq genes.

TIN and PIN transcripts are potential alternative splicing regulators

We found that the average frequency of exon skipping for genes in the RefSeq reference dataset that show evidence of PIN transcripts is 0.23, and the average frequency of exon skipping for exons immediately 3' to TIN transcripts is 0.22. These frequencies are significantly ($p < 0.0001$) higher than the average frequency of exon skipping (0.14) in the overall set of RefSeq genes (data not shown).

Next, we examined both the distribution of exon-skipping frequency across the different exons of protein-coding genes (Figure 2A) and the abundance of unspliced TIN EST contigs across the different introns of the same genes (Figure 2B). A higher frequency of exon skipping was detected closer to the 5'-ends of protein-coding genes (Figure 2A), and a concomitantly higher abundance of unspliced TIN EST contigs was detected in the first 2 introns of these genes (Figure 2B). The median length of the first intron of most genes with TIN transcripts (60–63%) is similar to the median length of all other introns in the same genes (Figure 2B). One would expect a random distribution of TIN EST contigs across the different introns of a gene if TINs were transcribed by spurious RNA Polymerase II transcription. In contrast, we found an uneven distribution, which suggests that TIN transcription may frequently be influenced by proximity to the gene promoter and might be regulated and driven by a so far uncharacterized mechanism favoring the first introns. It should be noted that in a fraction (29–33%) of genes with TIN transcription, the first intron is longer than the others, resulting in a significant correlation between TIN contigs and intron length [Supplementary Figure S1]. The hypothesis is that more information is conveyed in the longer intronic regions of these particular genes (see Discussion).

Design and overall performance of a gene-oriented intronic oligonucleotide array platform

The analyses described so far have indicated active sites of wholly and partially intronic transcription of noncoding messengers (TIN and PIN transcription) within protein-coding genes. Guided by this information, we designed a 44k oligoarray with randomly-selected protein-coding genes along with the corresponding intronic transcripts. This permitted large-scale detection of human intronic expression in a gene-oriented manner. A total of 9,097 probes from the commercially available set of Agilent 60-mer probes (Figure 3A, probe 5) were used, representing different exons in 7,464 unique randomly-selected protein-coding genes, along with custom-designed intronic probes for the antisense or sense strand, as shown in Figure 3A. Probes for 7,520 TIN transcripts (Figure 3A, probes 3 and 4) and 4,977 antisense PIN transcripts (Figure 3A, probe 1) were designed as described below. A probe representing each PIN-overlapped protein-coding exon was also included (Figure 3A, probe 2). All protein-coding genes and all intronic transcripts in the oligoarray are represented in the UniGene dataset at GenBank.

We opted to use the 60-mer Agilent oligoarray technology to construct this custom-designed array because the probe characteristics and the hybridization and washing protocols in this platform have been optimized to attain reproducible results [30]. Therefore, probe design followed Agilent recommendations with respect to GC content and T_m , as detailed under Methods, to ensure a homogeneous and effective hybridization of fluorescent targets. In fact, the reproducibility of expression in our experiments was fairly high, as evaluated by the correlation coefficients obtained for the two-color raw intensities within each slide and the correlation coefficients of inter-slide comparisons. These correlation coefficients ranged from 0.914 to 0.981 for intra-slide and from 0.915 to 0.949 for inter-slide comparisons.

Probe specificity was ensured by selecting 60-mer sequences with a homopolymeric stretch no longer than 6 bases; in addition, probes should not have 8 or more bases derived from repetitive regions of the genome. The selected probes have a low probability of cross-hybridization, as estimated by a BLAST search against the sequences of all transcribed human messages using the following criteria. All probes have 100% matches to the transcript sequences they represent, which translates into a best-match BLAST bit-score of 119. A bit-score high-end cutoff for the second-best match of each selected probe was set at 42.1, which would correspond to cross-hybridization with a maximum match of 21 bases with no gaps. This high-end cutoff level was determined from the bit-scores of the second-best hits for all the Agilent-designed commercial probes for protein-coding genes included in our platform; it is a conservative cutoff that includes 90% of the Agilent-optimized probes [Supplementary Figure S2]. Commercial probes with bit-score cross-hybridization matches higher than 42.1 were included because Agilent have tested each of their probes individually

for absence of cross-hybridization [30]. Since we did not test individual probes, we opted to use this conservative high-end cutoff parameter for the intronic probes.

Negative controls in the oligoarray (1,198 Agilent commercial control probes, see Methods) included sequences from adenovirus E1A transcripts, synthetically generated mRNAs, *Arabidopsis* genes and control probes designed not to hybridize to targets because of secondary structure. The hybridization and washing stringency conditions optimized by Agilent ensured that the raw signal intensities for these negative controls (median 34.3) in our experiments were low. For each experiment, the average negative control intensity plus 2 SD was used as a low-limit cutoff to call the expressed and not-expressed genes.

Figure 3B shows the distribution of average intensities in the microarray experiments for genes called not-expressed (below the low-limit cutoff) and for protein-coding, antisense or sense TIN and antisense PIN expressed transcripts. The distribution is skewed towards higher intensities for protein-coding transcripts and the median intensity is 351. The distribution of intensities is very similar for all types of intronic transcripts, and is skewed towards lower intensities when compared to that of protein-coding genes (Figure 3B). Nevertheless, the median intensities (134 for antisense TIN, 126 for antisense PIN and 135 for sense TIN transcripts) were sufficiently above that of the negative controls to permit a considerable number of expressed intronic transcripts to be detected in all tissues. Discrimination between expressed and not-expressed transcripts may be more critical for intronic messages than for protein-coding ones, and a larger fraction of false-negatives may be present in the intronic data. Our results corroborate previous tiling microarray measurements in chromosomes 21 and 22, which showed that ncRNAs were generally expressed at lower levels than protein-coding ones [31].

Partially and wholly intronic noncoding transcripts expressed in three human tissues

Gene expression profiles for human prostate, kidney and liver were obtained with the 44k intronic oligoarrays. Arrays were hybridized with amplified Cy3- and Cy5-labeled cRNA obtained by *in vitro* linear amplification of poly(A)-containing RNAs using T7-RNA polymerase. Figure 4 shows the number of protein-coding, TIN and PIN probes with signals greater than the negative control average plus 2 SD in at least one of the three tissues examined, and in each separate tissue. It can be seen that while 74% of protein-coding messages were expressed, only 30% of antisense TIN and 48% of antisense PIN transcripts were expressed in at least one tissue. A similar fraction of sense TIN transcription (36%) was observed, underscoring the natural transcription of sense intronic transcriptional units that has been observed elsewhere [29, 32].

It can be seen that 50–69% protein-coding transcripts were expressed in each individual tissue, while 14–32% antisense and sense TIN and 20–45% antisense PIN

transcripts were detected (Figure 4). This reveals that the abundance of intronic transcripts was lower than that of protein-coding messages, in terms of both the diversity of messages per tissue (Figure 4) and the relative distribution of signal intensities (Figure 3B).

The distribution along human chromosomes of the number of TIN RNA transcriptional units expressed in liver (Figure 5, gray bars) clearly agreed with the distribution computed by informatics analysis based on the entire GenBank EST dataset (Figure 5, black bars). Both distributions generally follow that of the number of RefSeq genes in each chromosome (Figure 5, red bars). There are a few exceptions; for example, chromosomes 10 and 13 seem to contain a higher fraction of expressed TIN RNA transcriptional units than protein-coding RefSeq genes, and chromosomes 19 and X have lower ratios of intronic transcriptional units to protein-coding genes. Interestingly, X chromosome inactivation (XCI) depends on a single noncoding sense-antisense transcript pair, *Xist* and *Tsix*, transcribed from a single *locus* on chromosome X. At the onset of XCI, *Xist* RNA accumulates on one of the two Xs, coating and silencing the chromosome in *cis*, a phenomenon controlled by a transient heterochromatic state that regulates transcription [33].

Figure 6 shows the distribution of sense and antisense TIN transcripts simultaneously expressed from the same *locus* as a function of the fraction of transcripts expressed in each of the three tissues. Considering only the top 10% most highly expressed sense and antisense TIN transcripts in each tissue, only 1–5% were detected simultaneously from both strands of the same introns in protein-coding genes. Among the top 50% intensities, over 83–90% of intronic transcription events are specific to one strand. Even when 100% of the expressed transcripts were considered, 63–79% were found to be expressed exclusively from one strand. This suggests that most of the sense and antisense messages are independent transcriptional units. It is apparent that the most highly-expressed intronic transcripts are strand-specific, which again suggests a regulated cellular process.

Antisense TIN transcripts are enriched in introns of genes related to regulation of transcription

We selected the top 40% most highly-expressed antisense TIN transcripts in each tissue and identified the protein-coding genes to which these transcripts map. The Gene Ontology (GO) annotation of these protein-coding genes was compared with the BiNGO tool [34] to the entire list of protein-coding genes in the array that showed evidence of antisense TIN transcription. The GO category *Regulation of transcription, DNA-dependent* (GO: 006355) was found to be significantly ($p < 0.05$) enriched in all three tissues. A typical GO

enrichment analysis is shown for prostate in Figure 7A; similar results for kidney and liver are shown in [Supplementary Figure S3].

Among the top 40% most highly expressed antisense TIN transcripts mapping to 712 protein-coding genes in the prostate, 108 (15%) belong to *Regulation of transcription, DNA-dependent* (Figure 7B). Analogous results were obtained for liver and kidney, where 75 out of 430 (17%) and 125 out of 853 (15%) of the genes, respectively, belong to *Regulation of transcription, DNA-dependent*. A total of 128 unique genes related to *Regulation of transcription* were found in common among the 40% most highly-expressed antisense TIN transcripts in prostate, kidney or liver. Most of these (72 genes, i.e. 56%) were expressed in all three tissues (Figure 7B), while some were shared between two tissues and a few were only expressed in one. The *Regulation of transcription* GO category includes various DNA-binding proteins such as transcription factors, zinc finger and nuclear receptor genes. The entire list of genes identified in Figure 7B can be found in Table S1. Similar analyses with the top 40% highly-expressed sense TIN and antisense PIN transcripts did not identify any enriched GO category.

A similar analysis using the 40% most highly-expressed protein-coding genes showed an entirely different set of significantly ($p < 0.05$) enriched GO categories; between 10 and 15 significantly enriched categories were detected in each tissue, and none was related to *Regulation of transcription* [Supplementary Figure S4]. The most significantly enriched GO categories in all three tissues include genes involved in RNA and protein biosynthesis, ribosome biosynthesis, mRNA processing and initiation of translation.

Many TIN and PIN RNAs are insensitive to RNAP II inhibition or are even up-regulated by α -amanitin

We treated human prostate cancer-derived LNCaP cells with the RNA polymerase II inhibitor α -amanitin for 24h, and used the 44k oligoarray to assess its effect on the expression of protein-coding and noncoding intronic RNA. Differentially expressed transcripts (Figure 8) were identified by combining two statistical approaches, the SAM method with $FDR < 2\%$ [35] and a Signal-to-Noise Ratio (SNR) analysis with bootstrap permutation ($p < 0.05$) [36]. About 38% (3,764) of the expressed protein-coding messages were significantly affected by RNAP II inhibition, while the remaining presumably more stable mRNAs were not. As expected, most (96%) of the affected protein-coding messages were down-regulated, but 4% were up-regulated. We found that 96 protein-coding RNAs were up-regulated at least two-fold. Kravchenko *et al.* [23] found that a similar number of protein-coding RNAs (70 transcripts) were up-regulated two-fold or more by α -amanitin in HeLa cells, in experiments with Affymetrix microarrays representing ~20,000 protein-coding transcripts.

Markedly fewer of the expressed TIN antisense (12%) and sense (14%) transcripts were affected by α -amanitin. Similar fractions of antisense (16%, 45/285) and sense (14%, 50/354) TIN transcripts were up-regulated in α -amanitin treated cells (Figure 8). PIN antisense transcript levels exhibited an expression pattern rather different from that of protein-coding transcripts when RNAP II was inhibited: only 16% were affected, of which 10% (41/395) were up-regulated. Interestingly, 2.5-4 times as many TIN and PIN RNAs as protein-coding messages (4%) were up-regulated by α -amanitin (Figure 8).

Intriguingly, the intronic messages (both TIN and PIN transcripts) with increased abundance in cells with blocked RNAP II transcription were transcribed from the introns of protein-coding genes that are again significantly ($p < 0.05$) enriched in the *Regulation of transcription* GO category (Figure 9). A complete list of the noncoding intronic and protein-coding transcripts that were up-regulated upon exposure to α -amanitin is shown in Table S2.

We consider that the stringent criteria used, combining two statistical methods to identify the differentially expressed transcripts, may be conservative. Therefore, the proportion of intronic messages that are up-regulated following α -amanitin treatment may be even greater than those reported here. In any case, the number of intronic ncRNAs insensitive to inhibition, or up-regulated upon α -amanitin treatment, is likely to be thousands when extrapolated to all the intronic transcripts found in human cells. Considering only the 55,139 wholly intronic EST clusters, over a thousand are predicted to be up-regulated if at least 13% are affected by 24h of RNAP II inhibition.

Tissue signatures of TIN and PIN expression

Tissue-specific signatures of intronic expression were determined for prostate tumor, normal kidney and normal liver. A total of 449 antisense TIN (Figure 10A), 600 sense TIN (Figure 10B) and 492 antisense PIN transcripts (Figure 10C) were identified, using a combination of two statistical approaches (see Methods for details). A complete list of the intronic transcripts identified in tissue signatures, and the corresponding spliced protein-coding genes mapping to the same genomic *loci*, is provided in Tables S3-5. These tissue signatures comprise hundreds of different transcripts (Figure 10A-C), mapping to introns of genes with diverse functions, and no particular GO category enrichment could be detected.

A tissue signature containing 2,961 protein-coding transcripts was also identified (Figure 10D). Analysis of GO enrichment (not shown) revealed that in liver the protein-coding tissue signature is enriched in GO categories related to urea cycle (GO: 006594), cysteine metabolism (GO: 006534), cholesterol biosynthesis (GO: 008203) and prostaglandin metabolism (GO: 006693), while in kidney it is enriched in the GO categories related to sodium and potassium ion transport (GO: 006834 and GO: 006813, respectively).

In the prostate, no relevant GO categories were enriched, but prostate-specific genes such as *KLK3* and *TMEPAI* were found.

We searched for co-regulated intronic and protein-coding pairs of messages that were simultaneously expressed from the same genomic *locus* in the same tissue, in order to identify noncoding RNAs potentially involved in modulating gene expression in a *cis*-acting manner. For this purpose, we initially cross-referenced the tissue signature of antisense PIN RNAs (Figure 10C) with the protein-coding signature (Figure 10D) to determine whether both signatures contained PIN-overlapped exons of the protein-coding gene transcribed from the opposite strand in the same genomic *locus* (see Figure 3, probe 2). Considering all three tissues, we found 105 gene *loci* in which antisense PIN RNAs and PIN RNA-overlapped protein-coding exon pairs were simultaneously detected in both tissue signatures (Figure 11). The tissue expression patterns of PIN RNA and PIN RNA-overlapped exon pairs were similar in a subset of 54 *loci* (Table S6 and Figure 11A, left and central panels). Interestingly, the 3' exon of the protein-coding transcript in this subset (Figure 11A, right panel) follows the same pattern. This is the predominant pattern in the tissue signature. Conceivably, the similar relative levels of antisense PIN RNA and protein-coding exons indicate that the intronic RNA has a functional role in modulating the transcription or transcript stability of the corresponding protein-coding gene. Alternatively, the levels of antisense PIN RNA and protein-coding message in each tissue may be similar because a common factor simultaneously modulates the transcription of both types of message from the same *locus*.

In a smaller subset of 10 *loci*, the 3' exon of the protein-coding transcript (Figure 11B, right panel) does not follow the pattern of tissue expression of the PIN RNA and the corresponding PIN-overlapped exon of the protein-coding gene (Table S6 and Figure 11B, left and central panels). In addition, the PIN RNA (Table S6 and Figure 11C, left panel) in 7 *loci* has an inverted expression pattern relative to that of the PIN RNA-overlapped exon (Figure 11C, central panel). In some tissues, there is an inverted pattern in the relative levels of PIN-overlapped exon and the 3' exon of the protein-coding gene for these two sets (Figure 11B and C, central and right panels), suggesting that the protein-coding message is alternatively spliced in a tissue-dependent manner. The similar levels of PIN RNAs and PIN-overlapped exons in Figure 11B (central and right panels) suggest that in these cases the PIN RNA may be involved in exon retention of the protein-coding gene, whereas the inverted pattern observed in Figure 11C (central and right panels) suggests that the PIN RNA may favor skipping of the overlapped exon. The effect of intronic RNAs on splicing has been documented in a recent report, where overexpression of a naturally-occurring antisense PIN RNA (*Saf* transcript) mapping to the first intron of *Fas* caused the retention of an alternative *Fas* exon that was complementary to the antisense PIN transcript [17].

An analogous cross-reference of tissue signatures from intronic and protein-coding messages (Figure 10D) was performed using the antisense and sense TIN RNA tissue signatures (Figures 10A and B). Among the three tissues, we compiled 153 gene *loci* in which pairs of antisense or sense TIN RNAs and the 3' protein-coding exon were simultaneously detected in the tissue signatures (Figure 12). A similar tissue expression pattern of antisense TIN RNA and the 3' protein-coding exon pair was detected in a subset of 41 *loci* (Table S7 and Figure 12A). For 18 pairs of antisense TIN/3' exon an inverted expression pattern was observed (Figure 12B). Similar direct or inverted expression patterns were found for 71 (direct) or 23 (inverted) pairs when sense TIN and 3' protein-coding exons from the same *locus* were cross-referenced (Figure 12C and D, respectively). The slightly higher proportion ($71/94 = 0.76$) of directly correlated sense TIN/3' exon pairs compared to antisense TIN/3' exon pairs ($41/59 = 0.69$) may have resulted from internal priming of intronic segments of premature mRNAs containing stretches of poly-A during cRNA amplification when the target was being prepared. As for co-regulated PIN RNAs, the correlated expression of TIN RNAs and the 3' exons of protein-coding transcripts may suggest that these noncoding RNAs have a role in modulating the transcription rate or the stability of the corresponding protein-coding RNA.

Discussion

Long intronic unspliced transcripts in humans

In this work we have evaluated the contribution of introns in the human genome to the production of noncoding RNAs by gathering data on expressed intronic sequences from public databases, and in parallel by measuring expression with intronic oligoarrays. We focused on the unspliced messages that map wholly (TIN) or partially (PIN) to intronic regions and found that most of the genes defined by RefSeq sequences (74%) have intronic transcription. This fraction is likely to prove even greater since intronic expression has not yet been assessed in different developmental stages and physiological conditions. While some of the unspliced intronic ESTs (mapping to the sense strand) may represent hitherto-overlooked exons of alternatively spliced forms of known genes, a significant number of the sense and antisense transcripts in this dataset is likely to derive from novel independent transcriptional units. This is supported by the low protein-coding potential and long length of the TIN and PIN EST contig sequences (medians of 573 nt and 719 nt, respectively), well above the typical lengths of exons on protein-coding genes (median 141 nt).

The median length of the 55,000 TIN RNAs identified in all chromosomes in our analysis is in line with the lengths observed in previous reports by RACE analysis of non-

annotated transcripts from 10 human chromosomes (average length 680 nt, range 173–4,650 [37]). Almost none of the TIN EST sequences (0.2%) matched known small nucleolar RNAs or microRNAs. Nevertheless, it remains possible that some long TIN messages are precursors of yet-undiscovered small RNAs.

We found no correlation between intron size and the abundance of mapped TIN unspliced EST contigs for most of the genes (~60%) that showed evidence of intronic transcription, suggesting that most intronic transcription does not occur by chance. In addition, the consistent correlation between approximately 30% of TIN contigs and intron length might support the ‘genomic design’ hypothesis [38], in the sense that transcription of the longer introns in tissue and development-specific genes could carry regulatory information [38]. In effect, we see a more abundant expression of intronic antisense messages in genes with regulatory functions (see discussion below).

Intronic transcripts may stabilize protein-coding transcripts or regulate their alternative splicing

Most of the PIN and TIN RNAs selected in the tissue-specific signatures have the same tissue expression patterns as the corresponding protein-coding genes. This might indicate that transcription of some PIN and TIN RNAs is linked to a *cis*-acting stabilization of the corresponding protein-coding transcript [39-41]. Intronic transcripts may also act in *trans*, for example by controlling regional chromatin architecture as demonstrated for some specific long ncRNAs [33, 42, 43]. Overexpression of complete introns in the *CFTR* gene affects the expression of a large number of protein-coding messages in *trans*, many of them related to *CFTR* function [44].

A few PIN RNAs selected in the tissue-specific signatures showed tissue expression patterns that correlated with the corresponding protein-coding exon overlapped by the PIN transcript. However, the expression pattern of an exon closer to the 3' end of the same protein-coding gene was not correlated, suggesting that alternatively spliced isoforms of the protein-coding transcripts were tissue-specific. Splicing is known to be modulated by the binding of intronic splicing enhancer (ISE) and silencer (ISS) elements to regulatory factors, favoring or blocking spliceosome formation [45]. Some PIN RNAs might regulate the skipping or retention of exons by interacting either with splice signals or with ISE and ISS elements in pre-mRNAs. In fact, there are examples of control of exon skipping by artificially introduced oligonucleotides in human cells [46], and some promising therapeutic strategies rely on antisense oligonucleotides that modulate exon-skipping [47]. As for exon retention, a recent report has identified a long antisense noncoding transcript named *Saf*, which maps as a partially intronic transcript to the first intron of *Fas*, a gene encoding an apoptotic protein [17]. Overexpression of *Saf* in Jurkat cells induced the expression of different

alternatively spliced *Fas* isoforms, in which an alternative exon overlapped by *Saf* was retained and non-adjacent 3' exons were skipped, indicating that *cis*-acting antisense intronic RNAs have a regulatory function [17].

Our present microarray analysis is conservative, in that we only selected transcripts that had correlated patterns of intronic and protein-coding messages and were also simultaneously present in tissue-signatures. A more direct experimental approach, for example over-expressing or suppressing specific PIN transcripts and measuring their effect on the splicing pattern of the overlapped exons, might reveal novel candidates for antisense RNA regulators of exon usage that possibly contribute to a ubiquitous and under-appreciated mechanism of alternative splicing regulation.

Alternative splicing affects more than 70% of human protein-coding genes [48], in which exon skipping is the most frequent event [49]. We found *in silico* evidence that *cis*-acting intronic transcription influences alternative splicing, i.e. a higher incidence of noncoding transcription in the first introns along with higher skipping frequency of the first exons in the protein-coding genes. In addition, the frequency of skipping for exons close to or overlapped by intronic transcripts was significantly higher than the average frequency of exon skipping in the overall set of human genes. In fact, exon skipping can be artificially induced by introducing antisense oligonucleotides that map to intron/exon junctions [46, 47] or to wholly intronic regions [45].

The higher incidence of transcription in the first intron, closer to the gene promoter, might have other functional implications such as the impairment of transcription by transcriptional interference [50, 51], regulation of gene promoter usage [52], or regulation of the initiation of RNA Polymerase II (RNAP II) transcription [15]. In the latter case, ncRNAs are known to function as co-activators; protein-binding ncRNAs are expected to provide a broad and diverse way of controlling mRNA transcription [15, 53]. We speculate that a high fraction of the intronic transcripts, especially the sense TIN RNAs, may act in *trans*, being parts of multi-component RNA-protein complexes that regulate gene expression. There are thousands of potential RNA regulators, which may effectively amplify the complexity of a human genome with a limited number of protein-coding genes [11, 54] through RNA-RNA, RNA-DNA, or RNA-protein interactions.

Biogenesis of TIN and PIN transcripts

We evaluated the contribution of RNAP II to the biosynthesis of intronic ncRNAs in human cells by blocking its activity with α -amanitin and measuring the levels of protein-coding and noncoding intronic messages. Remarkably, a considerable fraction (10–16%) of the wholly intronic or partially intronic antisense transcripts was up-regulated, a fraction 2.5–4 fold higher than that observed for protein-coding messages (4%). In addition, fewer intronic (12–

16%) than protein-coding (38%) transcripts were sensitive to RNAP II inhibition. Importantly, the sense TIN RNAs responded quite similarly to antisense TIN RNAs with respect to RNAP II inhibition, suggesting that these ncRNAs share similar properties that are different from protein-coding messages. The refractory behavior of intronic transcript expression after α -amanitin treatment, and the apparent up-regulation of many intronic transcripts, suggests that a different transcriptional system may be involved in the biosynthesis of these long wholly intronic ncRNAs. A reasonable candidate is spRNAP IV, which is activated by α -amanitin [23], though the mechanism involved remains elusive. Further experimentation is warranted to verify this hypothesis.

Advantages of a gene-oriented intronic array platform

Experimental analysis using genome tiling arrays has permitted unbiased probing of transcribed regions in the human genome [31, 32, 37, 55, 56]. Probing chromosomes 21 and 22 revealed 5.3 kb of novel transcribed sequences within or overlapping the intronic regions of well-characterized genes, of which 2.7 kb (51%) are antisense to the protein-coding genes [31]. Tiling arrays of the whole human genome have extended these analyses, detecting messages in liver that map to 1,529 and 1,566 novel transcriptionally active intronic regions arising, respectively, from the antisense or the sense strands of the corresponding gene [29]. Genome tiling arrays for 10 human chromosomes revealed interlaced networks of both poly A⁺ and poly A⁻ annotated transcripts and unannotated transcripts of unknown function [37]. It has become apparent that introns as well as intergenic regions constitute major sources of non-protein-coding RNAs [41], and tiling arrays promise to help unravel the complex cellular program of intronic transcription.

Different physiological and pathological conditions are yet to be probed by tiling arrays, and the amount and complexity of the information generated by high-density whole human genome tiling arrays may make the experiments difficult to perform. In this context, we believe that a gene-oriented intronic array that samples the intronic noncoding regions of the genome from which there is previous evidence of transcription, along with the corresponding protein-coding regions, will help to identify the particular gene families, biological processes or functional gene categories of greatest relevance to any physiological condition under study. In the present case, we have shown that genes involved in the regulation of transcription are enriched among those biological processes with the highest intronic transcription. High-density tiling arrays of selected chromosome regions containing these transcriptional factor genes should permit further detailed studies of intronic expression patterns. The information gathered from such a combined approach should help accelerate the acquisition of information about the emerging diverse roles of intronic messages.

Tissue-specific intronic expression and enrichment of genes related to regulation of transcription

We showed that different repertoires of long TIN and PIN RNAs are preferentially transcribed in human prostate, liver and kidney. Tissue-specific expression signatures provide strong evidence that intronic transcripts are physiologically relevant. Expression signatures of microRNAs have been reported to classify human cancers [22], adding to the evidence that different ncRNAs are tissue-specific and functionally important.

The present finding that the most abundant wholly intronic antisense RNAs are transcribed from introns of genes related to the regulation of transcription provides a clue to their functional relevance. A high degree of conservation is expected in those intronic genomic regions that are under strong selective constraints. In fact, conserved genomic regions have been identified by several different approaches in the introns of genes involved in transcriptional regulation [57-61]: identification of non-transcribed ultraconserved sequences [57], multispecies conserved sequences [58], sequences conserved in vertebrates but highly divergent among chimpanzees and humans [59], short blocks of multiple-copy sequences (pyknons) [60], or conserved regions without transposon insertions [61]. Our results add to these findings by showing that conserved intronic DNA segments of genes involved in transcriptional regulation are the sources of one of the most abundant intronic RNAs in three different human tissues. The possibility that regulatory genes are controlled by ncRNAs transcribed at the same *loci* is appealing. It would represent an additional mechanism for regulating the regulators, in a rather sophisticated system for fine-tuning eukaryotic gene expression.

Conclusion

Our approach has used an oligoarray-based gene-oriented intronic platform as a practical and effective compromise between a biased array that only probes the protein-coding messages, and whole human genome tiling arrays. This approach has identified potentially functional intronic RNAs that are most abundantly transcribed from introns of genes involved in transcriptional regulation. Further comparative analysis of intronic transcription under a different number of physiological and pathological conditions should advance current knowledge about the diverse biological roles of these noncoding RNAs in the control of gene expression.

Materials and Methods

Cross-referencing of genomic coordinates of transcripts from different sequence datasets

The analyzed sequence dataset comprises all human RefSeq, mRNA and EST sequences, of which the genome coordinates were downloaded from the Genome Browser web page (<http://genome.ucsc.edu>) (hg17; NCBI Build 35, March/2005). First, sequences with poor alignment quality (coverage < 0.70 and identity < 0.90) or mapped to more than one genomic region were removed. Second, we discarded sequences with complicated rearrangement patterns such as T-cell receptor and immunoglobulin genes. ESTs and mRNAs that aligned to two or more non-overlapping RefSeq were filtered out as suspected chimeras. Sequencing errors in transcripts aligned to the genome sequence lead to gaps that are interpreted as introns by our parser. To avoid these falsely-identified introns, we joined adjacent exons whenever an intron of less than 30 bases was detected.

A bioinformatics tool was developed to handle the over 5 million human ESTs efficiently. Essentially, this tool consists of a package of scripts written in Perl that uses files of genome mapping coordinates directly obtained from the UCSC genome browser. The use of coordinates avoids the computationally intensive and parameter-dependent problems of alignment-based programs. EST sequences with overlapping exons were merged into EST clusters using the genomic mapping coordinates. RefSeq and mRNA sequences were processed separately and split into four sets according to the genomic strand to which they mapped, and further sub-divided into spliced and unspliced groups of messages. Sequences from the same strand in each sub-group were merged into a transcriptional unit when their exons overlapped at the same genomic *locus*. The mRNA dataset was aligned against the RefSeq dataset to identify additional splice variants, intronic and antisense transcripts represented in the mRNA collection, as detailed in Table 1.

From the combined data described above, a reference dataset was defined comprising the set of 15,783 spliced non-redundant RefSeq transcriptional units plus the evidence of additional splice variants obtained for each transcriptional unit from all mRNA sequences mapping to the same *locus*. This RefSeq reference dataset was compared to the total set of EST clusters, in order to define those that were exonic or wholly intronic to genes in the reference dataset. For wholly intronic RNAs, we recorded the relative position of the RefSeq intron to which the RNA mapped with respect to the total number of introns of the respective RefSeq reference gene.

Partially intronic EST contigs were identified in a later step, by searching for evidence of two or more overlapping unspliced EST sequences that mapped to an exon and covered the intronic regions flanking the exon on each side by more than 30 bases.

Exon skipping frequencies

For each exon of a gene from the RefSeq reference dataset, we counted the number of times that it mapped to an exon (# in exon) or an intron (# in intron) in all the mRNA sequences from the same sub-group (mRNAs and RefSeqs from the same locus and on the same genomic strand). Exon skipping (ES) frequency was given by:

$$ES = 1 - [\# \text{ in intron} / (\# \text{ in intron} + \# \text{ in exon})].$$

Design of intronic 44k microarray

Oligonucleotide probes were designed for 7,520 wholly (TIN) and 4,977 partially (PIN) intronic noncoding RNAs picked randomly from the list of unspliced EST contigs with most abundant ESTs representing each type of intronic transcript. First, for each PIN or TIN RNA, we selected all 60-mer sequences that satisfied a series of conditions [30] as follows: probes should not have 8 or more bases derived from repetitive regions of the genome or homopolymeric stretches of 7 or more bases (low complexity); and they should have a GC content of 35–55% and a melting temperature (T_m) of 68–76°C. In order to reduce cross-hybridization, each 60-mer sequence was searched by BLAST against a specific database comprising all human genomic regions for which the mRNA or EST data give any evidence of transcription. Those 60-mer sequences for which the second best hits against this database had bit-scores equal to or lower than 42.1 were carried forward to the next step. They were mapped back to their respective targets and one probe was selected closer to the 3' end of each target, assuming that PIN and TIN RNAs were antisense to protein-coding genes. For each target, a second probe was selected for the opposite strand by taking the reverse complementary sequence of the selected 60-mer. For PIN RNAs, the probe on the opposite strand corresponds to the exon of the gene where the partially intronic message overlaps. To measure the transcriptional level of the protein-coding genes to which PIN and TIN RNAs mapped, we included in our array 14,074 elements corresponding to 7,464 unique Agilent-designed probes contained in the Whole Human Genome Oligo Microarray set (matched by their Gene_Name annotations to the RefSeq reference dataset genes to which the PIN and TIN RNAs mapped), together with the set of 2,256 positive and negative control Agilent commercial probes (IS-44290-1-V1_eQC-V1) designed for the Agilent human expression oligoarrays. Our custom intronic 44k microarrays were printed by Agilent Technologies. A list of all probes is available at GEO under accession number GPL4051.

Human tissue samples

Total RNA was purified from two pools of normal human liver, each with samples from 5 individuals, and from normal kidney tissues obtained from 17 individuals. Four pools of normal kidney were prepared (3 with 4 samples and 1 with 5 samples). In addition, two

prostate tumor samples were used. All samples were obtained from patients who signed informed consent, and approval was received from the ethics committees of the hospitals. Total RNA was purified using Trizol (Life Technologies) according to the manufacturer's instructions, followed by treatment with DNase I following the "on-column DNase digestion" protocol of the Qiagen RNeasy kit (Qiagen) to remove potential genomic DNA contamination. All RNA samples were checked for purity using a ND-1000 spectrophotometer (NanoDrop Technologies) and for integrity by electrophoresis on a 2100 BioAnalyzer (Agilent Technologies).

α -Amanitin experiments with LNCaP cells

The prostate carcinoma cell line LNCaP was obtained from ATCC and maintained in RPMI 1640 medium (Gibco) supplemented with 10% (vol/vol) fetal calf serum (FCS), 3 mM L-glutamine, 100 μ g/ml streptomycin and 100 U/ml penicillin, at 37 °C and 5% CO₂. For RNAP II inhibition experiments, 8 x 10⁵ LNCaP cells were plated in p60 dishes and cultured for two days, after which the medium was replaced by fresh medium with or without (mock) 50 μ g/ml α -amanitin (Roche). After 24 h, the cells were washed once with ice-cold PBS, harvested, pelleted and stored at -80 °C. Total RNA was isolated using Qiagen RNeasy kit and treated with DNase I following the "on-column DNase digestion" protocol (Qiagen). RNA quality was checked as described above. Two biological replicas were processed separately.

Sample labeling and microarray hybridization procedures

Cy5- and Cy3-labeled cRNA was obtained using 300 ng total RNA as template for amplification of poly-A RNA by T7-RNA polymerase with the Agilent Low RNA Input Fluorescent Linear Amplification kit. The T7-polymerase amplified cRNA labeling approach advantageously replaces the reverse-transcriptase cDNA-labeling used in early microarray experiments, because T7-RNA polymerase labeling of cRNA preserves the strand orientation of the original mRNA template. Reverse-transcriptase labeling can eventually generate a complementary cDNA second strand and cause artifactual labeling of a target with the opposite sense to that of the original message. For LNCaP cell line samples (mock-treated or α -amanitin-treated cells), 500 ng total RNA was used and a control *in vitro* synthesized mRNA (Agilent RNA Spike-In kit) was spiked into the amplification and labeling assay. For kidney tissue samples, the 4 pools from normal individuals were considered as replicas, and each pair was labeled with either Cy3 or Cy5. Each liver sample pool, prostate tissue sample or LNCaP cell line sample was separately labeled in replicate with Cy3 or Cy5. Hybridization of 750 ng each of Cy3- and Cy5-labeled cRNA was performed with an Agilent *in situ* Hybridization kit-plus, as recommended by the manufacturer, using a total of

six 44k intronic microarrays. Slides were washed and processed according to the Agilent 60-mer Oligo Microarray Processing protocol and scanned on a GenePix 4000B scanner (Molecular Devices). Data were extracted from the images with ArrayVision 8.0 (Imaging Research Inc.). Cy5- and Cy3-derived intensity data from the same sample were corrected for intensity-dependent dye biases [62] using a Lowess function implemented in the R package (www.r-project.org). The different experiments with human tissues were normalized by the 40% trimmed mean intensity of all the spots in each slide that were above the mean plus 2 SD intensity of 1,198 negative controls. The experiments with the LNCaP cells (mock- and α -amanitin-treated cells) were normalized by the 40% trimmed mean intensity of 300 control probes from a specific probe set on the array that reports the signals from labeled targets generated from the synthetic spiked-in mRNA.

Statistical analyses

For tissue-specific expression profiles, the Significance Analysis of Microarray (SAM) approach was employed using as parameters: multi-class response, 1000 permutations, K-Nearest Neighbors Imputer, and false discovery rate (FDR) $\leq 0.00001\%$. Analysis of variance (ANOVA), implemented in the SpotFire Decision Site for Functional Genomics (SpotFire Inc.) with cutoff $p \leq 0.001$, was also used. Gene sets identified by SAM or ANOVA were combined in order to identify a more restricted set of genes that showed statistically significant changes of expression in a tissue by both analyses.

For RNAP II inhibition experiments, the Significance Analysis of Microarray (SAM) approach was employed, using as parameters: two-class unpaired response, t-statistic, 1000 permutations, K-Nearest Neighbors Imputer, and false discovery rates (FDR) ranging from 0.2% to 2%; a signal-to-noise ratio (SNR) analysis with 10k permutations ($p < 0.05$) was performed. Gene sets identified by SAM or SNR were combined in order to identify a more restricted set of genes that showed statistically significant changes of expression upon α -amanitin treatment by both analyses.

GO enrichment analyses

We used BiNGO, the Biological Network Gene Ontology plug-in tool [34] from the Cytoscape package (<http://www.cytoscape.org/>). We used the Hypergeometric statistical test with Benjamini & Hochberg's FDR multiple testing correction, choosing a significance level of 0.05. We used as the reference dataset all genes that were present in our 44k intronic microarray, as follows: for protein-coding genes, we used all Gene_IDs of protein-coding probes in the array; for TIN and PIN RNAs, we used all Gene_IDs to protein-coding genes

for which there were TIN or PIN RNA probes in the array mapping to the corresponding genomic *loci*.

Acknowledgments

The authors thank Camila Egidio for help with testing the Agilent microarray protocol. The authors also thank Dra. Marcia Kubrusly (Hospital das Clínicas, Universidade de São Paulo) and Dr. Marcello Barcinski (Instituto Nacional de Câncer, Rio de Janeiro) for providing the tissue samples. This work was supported by a grant from Fundação de Amparo a Pesquisa do Estado de São Paulo, FAPESP to SVA and AMDS and by fellowships from FAPESP and Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPq, Brasil.

Declaration of competing interests

The authors declare that they have no competing interests.

Additional files

Supplementary file 1: **Figure S1**. Abundance of wholly intronic noncoding transcription in RefSeq genes.

Supplementary file 2: **Figure S2**. Distribution of BLAST bit-score for the second best hit of the 60-mer oligonucleotide probes from the microarray.

Supplementary file 3: **Figure S3**. Most highly expressed TIN transcripts map to genes related to regulation of transcription.

Supplementary file 4: **Figure S4**. Gene ontology analyses with the most highly expressed protein-coding transcripts in three different human tissues.

Supplementary file 5: **Table S1**. List of 149 probes representing antisense TIN RNAs from 128 Gene IDs of genes related to Regulation of Transcription.

Supplementary file 6: **Table S2**. Protein-coding and intronic non-coding RNAs up-regulated upon alpha-amanitin treatment

Supplementary file 7: **Table S3**. Tissue signatures of 492 antisense PIN RNAs.

Supplementary file 8: **Table S4**. Tissue signatures of 449 antisense TIN RNAs.

Supplementary file 9: **Table S5**. Tissue signatures of 600 sense TIN RNAs.

Supplementary file 10: **Table S6**. Comparison of tissue signatures between antisense PIN RNAs and exons of protein-coding genes.

Supplementary file 11: **Table S7**. Comparison of tissue signatures between TIN RNAs and exons of protein-coding genes.

References

1. Reis EM, Ojopi EP, Alberto FL, Rahal P, Tsukumo F, Mancini UM, Guimaraes GS, Thompson GM, Camacho C, Miracca E *et al*: **Large-scale transcriptome analyses reveal new genetic marker candidates of head, neck, and thyroid cancer.** *Cancer Res* 2005, **65**(5):1693-1699.
2. Ferguson DA, Chiang JT, Richardson JA, Graff J: **eXPRESSiON: an in silico tool to predict patterns of gene expression.** *Gene Expr Patterns* 2005, **5**(5):619-628.
3. Gupta S, Zink D, Korn B, Vingron M, Haas SA: **Genome wide identification and classification of alternative splicing based on EST data.** *Bioinformatics* 2004, **20**(16):2579-2585.
4. Thanaraj TA, Clark F, Muilu J: **Conservation of human alternative splice events in mouse.** *Nucleic Acids Res* 2003, **31**(10):2544-2552.
5. Kan Z, Rouchka EC, Gish WR, States DJ: **Gene structure prediction and alternative splicing analysis using genomically aligned ESTs.** *Genome Res* 2001, **11**(5):889-900.
6. Modrek B, Lee CJ: **Alternative splicing in the human, mouse and rat genomes is associated with an increased frequency of exon creation and/or loss.** *Nat Genet* 2003, **34**(2):177-180.
7. Shendure J, Church GM: **Computational discovery of sense-antisense transcription in the human and mouse genomes.** *Genome Biol* 2002, **3**(9):research0044.0041-research0044.0014.
8. Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A, Diber A, Biton S, Tamir Y, Khosravi R *et al*: **Widespread occurrence of antisense transcription in the human genome.** *Nat Biotechnol* 2003, **21**(4):379-386.
9. Chen J, Sun M, Kent WJ, Huang X, Xie H, Wang W, Zhou G, Shi RZ, Rowley JD: **Over 20% of human transcripts might form sense-antisense pairs.** *Nucleic Acids Res* 2004, **32**(16):4812-4820.
10. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H *et al*: **Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.** *Nature* 2002, **420**(6915):563-573.
11. **Finishing the euchromatic sequence of the human genome.** *Nature* 2004, **431**(7011):931-945.
12. Scherer SW, Cheung J, MacDonald JR, Osborne LR, Nakabayashi K, Herbrick JA, Carson AR, Parker-Katirae L, Skaug J, Khaja R *et al*: **Human chromosome 7: DNA sequence and biology.** *Science* 2003, **300**(5620):767-772.
13. Pang KC, Stephen S, Engstrom PG, Tajul-Arifin K, Chen W, Wahlestedt C, Lenhard B, Hayashizaki Y, Mattick JS: **RNAdb--a comprehensive mammalian noncoding RNA database.** *Nucleic Acids Res* 2005, **33**(Database issue):D125-D130.
14. Griffiths-Jones S, Moxon S, Marshall M, Khanna A, Eddy SR, Bateman A: **Rfam: annotating non-coding RNAs in complete genomes.** *Nucleic Acids Res* 2005, **33**(Database issue):D121-124.

15. Goodrich JA, Kugel JF: **Non-coding-RNA regulators of RNA polymerase II transcription.** *Nat Rev Mol Cell Biol* 2006, **7**:612-616.
16. Khochbin S, Lawrence JJ: **An antisense RNA involved in p53 mRNA maturation in murine erythroleukemia cells induced to differentiate.** *EMBO J* 1989, **8**(13):4107-4114.
17. Yan MD, Hong CC, Lai GM, Cheng AL, Lin YW, Chuang SE: **Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas.** *Hum Mol Genet* 2005, **14**(11):1465-1474.
18. Krystal GW, Armstrong BC, Battey JF: **N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts.** *Mol Cell Biol* 1990, **10**(8):4180-4191.
19. Reis EM, Nakaya HI, Louro R, Canavez FC, Flatschart AV, Almeida GT, Egidio CM, Paquola AC, Machado AA, Festa F *et al*: **Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer.** *Oncogene* 2004, **23**(39):6684-6692.
20. Du T, Zamore PD: **microPrimer: the biogenesis and function of microRNA.** *Development* 2005, **132**(21):4645-4652.
21. Kiss T: **Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.** *Cell* 2002, **109**(2):145-148.
22. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA *et al*: **MicroRNA expression profiles classify human cancers.** *Nature* 2005, **435**(7043):834-838.
23. Kravchenko JE, Rogozin IB, Koonin EV, Chumakov PM: **Transcription of mammalian messenger RNAs by a nuclear RNA polymerase of mitochondrial origin.** *Nature* 2005, **436**(7051):735-739.
24. Dahary D, Elroy-Stein O, Sorek R: **Naturally occurring antisense: transcriptional leakage or real overlap?** *Genome Res* 2005, **15**(3):364-368.
25. Kiyosawa H, Yamanaka I, Osato N, Kondo S, Hayashizaki Y: **Antisense transcripts with FANTOM2 clone set and their implications for gene regulation.** *Genome Res* 2003, **13**(6B):1324-1334.
26. Iseli C, Jongeneel CV, Bucher P: **ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences.** *Proc Int Conf Intell Syst Mol Biol* 1999:138-148.
27. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A: **Identification of mammalian microRNA host genes and transcription units.** *Genome Res* 2004, **14**(10A):1902-1910.
28. Griffiths-Jones S: **The microRNA Registry.** *Nucleic Acids Res* 2004, **32**(Database issue):D109-111.
29. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S *et al*: **Global identification of human transcribed sequences with genome tiling arrays.** *Science* 2004, **306**(5705):2242-2246.
30. Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR *et al*: **Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer.** *Nat Biotechnol* 2001, **19**(4):342-347.

31. Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, Drenkow J, Piccolboni A, Bekiranov S, Helt G *et al*: **Novel RNAs Identified From an In-Depth Analysis of the Transcriptome of Human Chromosomes 21 and 22.** *Genome Res* 2004, **14**(3):331-342.
32. Kapranov P, Drenkow J, Cheng J, Long J, Helt G, Dike S, Gingeras TR: **Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays.** *Genome Res* 2005, **15**(7):987-997.
33. Sun BK, Deaton AM, Lee JT: **A transient heterochromatic state in Xist preempts X inactivation choice without RNA stabilization.** *Mol Cell* 2006, **21**(5):617-628.
34. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21**(16):3448-3449.
35. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *Proc Natl Acad Sci USA* 2001, **98**(9):5116-5121.
36. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA *et al*: **Molecular classification of cancer: class discovery and class prediction by gene expression monitoring.** *Science* 1999, **286**(5439):531-537.
37. Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G *et al*: **Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution.** *Science* 2005, **308**(5725):1149-1154.
38. Vinogradov AE: **Compactness of human housekeeping genes: selection for economy or genomic design?** *Trends Genet* 2004, **20**(5):248-253.
39. Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ *et al*: **Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs.** *Cell* 2004, **116**(4):499-509.
40. Reis EM, Louro R, Nakaya HI, Verjovski-Almeida S: **As antisense RNA gets intronic.** *Omic* 2005, **9**(1):2-12.
41. Mattick JS, Makunin IV: **Non-coding RNA.** *Hum Mol Genet* 2006, **15 Suppl 1**:R17-29.
42. Sanchez-Elsner T, Gou D, Kremmer E, Sauer F: **Noncoding RNAs of trithorax response elements recruit Drosophila Ash1 to Ultrabithorax.** *Science* 2006, **311**(5764):1118-1123.
43. Mayer C, Schmitz KM, Li J, Grummt I, Santoro R: **Intergenic transcripts regulate the epigenetic state of rRNA genes.** *Mol Cell* 2006, **22**(3):351-361.
44. Hill AE, Hong JS, Wen H, Teng L, McPherson DT, McPherson SA, Levasseur DN, Sorscher EJ: **Micro-RNA-like effects of complete intronic sequences.** *Front Biosci* 2006, **11**:1998-2006.
45. Bruno IG, Jin W, Cote GJ: **Correction of aberrant FGFR1 alternative RNA splicing through targeting of intronic regulatory elements.** *Hum Mol Genet* 2004, **13**(20):2409-2420.
46. De Angelis FG, Sthandier O, Berarducci B, Toso S, Galluzzi G, Ricci E, Cossu G, Bozzoni I: **Chimeric snRNA molecules carrying antisense sequences against the**

- splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Delta 48-50 DMD cells.** *Proc Natl Acad Sci U S A* 2002, **99**(14):9456-9461.
47. McClorey G, Fletcher S, Wilton S: **Splicing intervention for Duchenne muscular dystrophy.** *Curr Opin Pharmacol* 2005, **5**(5):529-534.
 48. Johnson JM, Castle J, Garrett-Engle P, Kan Z, Loerch PM, Armour CD, Santos R, Schadt EE, Stoughton R, Shoemaker DD: **Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays.** *Science* 2003, **302**(5653):2141-2144.
 49. Clark F, Thanaraj TA: **Categorization and characterization of transcript-confirmed constitutively and alternatively spliced introns and exons from human.** *Hum Mol Genet* 2002, **11**(4):451-464.
 50. Prescott EM, Proudfoot NJ: **Transcriptional collision between convergent genes in budding yeast.** *Proc Natl Acad Sci U S A* 2002, **99**(13):8796-8801.
 51. Martens JA, Laprade L, Winston F: **Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene.** *Nature* 2004, **429**(6991):571-574.
 52. Geirsson A, Lynch RJ, Paliwal I, Bothwell AL, Hammond GL: **Human trophoblast noncoding RNA suppresses CIITA promoter III activity in murine B-lymphocytes.** *Biochem Biophys Res Commun* 2003, **301**(3):718-724.
 53. Willingham AT, Orth AP, Batalov S, Peters EC, Wen BG, Aza-Blanc P, Hogenesch JB, Schultz PG: **A strategy for probing the function of noncoding RNAs finds a repressor of NFAT.** *Science* 2005, **309**(5740):1570-1573.
 54. Mattick JS: **RNA regulation: a new genetics?** *Nat Rev Genet* 2004, **5**(4):316-323.
 55. Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP, Gingeras TR: **Large-scale transcriptional activity in chromosomes 21 and 22.** *Science* 2002, **296**(5569):916-919.
 56. Rinn JL, Euskirchen G, Bertone P, Martone R, Luscombe NM, Hartman S, Harrison PM, Nelson FK, Miller P, Gerstein M *et al*: **The transcriptional activity of human Chromosome 22.** *Genes Dev* 2003, **17**(4):529-540.
 57. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS, Haussler D: **Ultraconserved elements in the human genome.** *Science* 2004, **304**(5675):1321-1325.
 58. Sironi M, Menozzi G, Comi GP, Cagliani R, Bresolin N, Pozzoli U: **Analysis of intronic conserved elements indicates that functional complexity might represent a major source of negative selection on non-coding sequences.** *Hum Mol Genet* 2005, **14**(17):2533-2546.
 59. Pollard KS, Salama SR, Lambert N, Lambot MA, Coppens S, Pedersen JS, Katzman S, King B, Onodera C, Siepel A *et al*: **An RNA gene expressed during cortical development evolved rapidly in humans.** *Nature* 2006, **443**(7108):167-172.
 60. Rigoutsos I, Huynh T, Miranda K, Tsirigos A, McHardy A, Platt D: **Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes.** *Proc Natl Acad Sci U S A* 2006, **103**(17):6605-6610.

61. Simons C, Pheasant M, Makunin IV, Mattick JS: **Transposon-free regions in mammalian genomes.** *Genome Res* 2006, **16**(2):164-172.
62. Peixoto BR, Vencio RZ, Egidio CM, Mota-Vieira L, Verjovski-Almeida S, Reis EM: **Evaluation of reference-based two-color methods for measurement of gene expression ratios using spotted cDNA microarrays.** *BMC Genomics* 2006, **7**:35.

Accession numbers

Related microarray data are deposited at Gene Expression Omnibus (GEO) under accession numbers [GenBank:GSE5452, GenBank:GSE5453].

Legends to the figures:

Figure 1. Length distribution of exons from RefSeq genes and of partially (PIN) and wholly (TIN) intronic noncoding transcripts. The curves show the length distribution of three different classes of transcripts reconstructed from genomic mapping and assembly of RefSeq and ESTs from GenBank. Exons of protein-coding RefSeq (red line), TIN (black line) and PIN (blue line) contig sequences. TIN and PIN contigs resulted from assembly of all GenBank unspliced ESTs (in gold) that cluster to a given intronic region in a genomic *locus*, as shown in the scheme above the curves.

Figure 2. Frequency of exon skipping and abundance of wholly intronic noncoding transcription in RefSeq genes. (A) Distribution of exon skipping events along spliced RefSeq genes with 7, 8, 9 or 10 exons. Full squares indicate the average frequency of skipping per exon for genes with evidence of TIN RNAs mapping to their introns. Open squares indicate the average frequency of skipping per exon for genes with no evidence in GenBank that TIN RNAs map to their introns. A significantly higher ($p < 0.002$) frequency of exon skipping was observed for RefSeq genes with TIN RNA transcription. (B) Distribution of TIN transcripts among the introns of RefSeq sequences with 7, 8, 9 or 10 introns selected from GenBank as being outside the 95% confidence level of significance (not correlated) in a Pearson correlation analysis between the abundance of TIN contigs per intron and the intron size (in nt). Bars indicate the median intron size (nt) for this selected set of genes. Triangles indicate the number of TIN contigs per intron for RefSeq genes for the same set.

Figure 3. Design and overall performance of the 44k gene-oriented intronic oligoarray. (A) Schematic view of the 44k intronic oligoarray probe design. Probe 1 is for the antisense Partially Intronic Noncoding (PIN) transcripts (dark blue arrow). Probes 3 and 4 are for antisense and sense Totally Intronic Noncoding (TIN) transcripts (black and hashed black arrows, respectively). Sense exonic probes 2 and 5 are for the protein-coding transcripts (red block and red arrow). Note that the latter were not systematically designed for an exon near the TIN message; in most instances a distant, 3' exon of the gene has been probed instead. (B) Average signal intensity distribution for antisense TIN (solid black line), sense TIN (dashed line), antisense PIN (blue line), or sense protein-coding exonic (red line) probes. Average intensities from 6 different hybridization experiments with three different human tissues, namely liver, prostate and kidney, are shown. Only probes with intensities above the average negative controls plus 2 SD were considered. The average intensity distribution for probes below this low-limit detection cutoff is shown in the curve marked as "Not expressed RNAs" (gray line).

Figure 4. Number of protein-coding, TIN and PIN transcripts expressed in three human tissues. Different types of transcripts are shown in each panel, and are color-coded as in Figure 3: protein-coding exonic (red bars), antisense TIN (black bars), antisense PIN (blue bars) or sense TIN transcripts (hashed black bars). The total number of probes present in the microarray for each type of transcript is shown with bars marked as (M). The number of transcripts expressed in at least one of the three tissues tested is shown with bars marked as (One). Transcripts exclusively expressed in each of the three tissues are shown with bars marked as (L), Liver; (P), Prostate; or (K) Kidney. The percentage of expressed transcripts relative to the total number of transcripts probed in the array is indicated at the top of each bar.

Figure 5. Genomic distribution of intronic RNAs. Relative chromosome sizes (blue bars) and the fractional number of GenBank Refseq genes (red bars) mapped per chromosome are shown. The distribution along the chromosomes of wholly intronic sequence contigs resulting from mapping and assembly of all ESTs in GenBank relative to the RefSeq reference dataset is shown (black bars). The distribution along the chromosomes of intronic RNAs expressed in human liver, as detected by oligoarray hybridizations, is shown as gray bars. The numbers on the Y axis refer to the fractional distribution in each chromosome.

Figure 6. Sense-antisense TIN transcript pairs simultaneously detected at different ranges of signal intensities for each of three different tissues. The percentages of TIN transcript pairs simultaneously transcribed from the same genomic *locus* in both the sense and

antisense orientations (full symbols), and detected at different ranges of signal intensities, are shown for each of three different tissues: liver (diamonds), prostate (triangles) and kidney (squares). The percentages of TIN messages transcribed in each tissue from only one of the two DNA strands (sense or antisense) are shown as open symbols.

Figure 7. Most highly expressed TIN transcripts map to genes related to regulation of transcription. TIN RNA expression data from three different human tissues (prostate, liver and kidney) were used to select the protein-coding genes to which the top 40% most highly expressed TIN transcripts map. The BiNGO program was used to identify significantly ($p \leq 0.05$) enriched GO terms within the set of selected protein-coding genes. (A) GO-enriched categories for prostate are shown. GO category “Regulation of transcription, DNA-dependent” (GO:006355) is the most significantly enriched. Similar results were obtained for liver and kidney (see Figure S3). (B) Venn diagram for the 128 unique protein-coding genes belonging to GO:006355 category “Regulation of transcription, DNA-dependent”. The number of genes in each tissue where intronic transcription was detected is shown in parenthesis; the numbers of coincident and dissimilar genes among kidney, prostate and liver are shown in the circles.

Figure 8. Effect of RNA Polymerase II inhibitor α -amanitin on the abundance of protein-coding, antisense TIN, sense TIN and antisense PIN RNAs. (A–D) Lines on each panel represent various transcripts for which the expression levels differed significantly between α -amanitin-treated prostate cells and untreated control cells. Each sample replica is shown in one column. Transcripts were selected by a SAM two-class test ($FDR < 0.2\%$ to 2%) combined with a signal-to-noise test ($p \leq 0.05$). For each line, expression intensities were normalized between the two conditions and colored as a function of the number of standard deviations from the mean value. (A) 3,764 significantly affected protein-coding transcripts; (B) 285 significantly affected antisense TIN transcripts; (C) 354 significantly affected sense TIN transcripts; (D) 395 significantly affected antisense PIN transcripts.

Figure 9. Genes with increased intronic transcription in the presence of the RNA Polymerase II inhibitor α -amanitin are enriched in the “Regulation of transcription” GO category. Gene ontology analysis was performed on protein-coding genes that were shown in the experiment of Figure 8 to have up-regulated expression of antisense PIN transcripts and sense and antisense TIN transcripts upon exposure to α -amanitin. Significantly ($p \leq 0.05$) enriched GO terms are shown.

Figure 10. Expression signature of intronic and protein-coding transcripts in human liver, prostate and kidney. (A–D) Transcripts with significantly different levels among prostate, kidney and liver samples were selected by a SAM multi-class test ($\text{FDR} < 0.0001\%$) combined with an ANOVA test ($p \leq 0.001$) and hierarchically clustered as described in the Methods. In each panel the selected transcripts are shown in the lines and sample replicas in the columns. For each line, expression intensities among the three tissues were normalized within each type of probe and colored as a function of the number of standard deviations from the mean value. (A) Tissue expression signature of 449 antisense TIN transcripts. (B) Tissue expression signature of 600 sense TIN transcripts. (C) Tissue expression signature of 492 antisense PIN transcripts. (D) Tissue expression signature of 2,961 protein-coding transcripts.

Figure 11. Expression signatures of antisense PIN RNAs and corresponding PIN RNA–overlapped exon pairs relative to their 3' protein-coding exons. A subset of 71 pairs of antisense PIN RNAs and corresponding PIN RNA-overlapped exons were identified among the tissue signatures shown in Figure 10 as having correlated patterns of expression: (A) 54 pairs were identified in which the 3' exon of the protein-coding transcript (right panel) follows a similar expression pattern to that of the PIN RNA/PIN RNA-overlapped exon pair (left and central panels); (B) 10 pairs were identified in which the 3' exon of the protein-coding transcript (right panel) does not follow the pattern of tissue expression of the PIN RNA and the corresponding PIN RNA-overlapped exon (left and central panels); (C) 7 pairs in which the PIN RNA (left panel) has an expression pattern inverted in relation to that of the PIN RNA-overlapped exon (central panel). Each line represents a genomic *locus* covered by three different types of probes (antisense PIN RNA, PIN RNA-overlapped protein-coding exon and 3' protein-coding exon). For each line, expression intensities among the three tissues were normalized within each type of probe and colored as a function of the number of standard deviations from the mean value.

Figure 12. Expression signatures of wholly intronic RNAs relative to their 3' protein-coding exons. Cross-referencing of the tissue signatures shown in Figure 10 identified subsets of TIN RNAs that have correlated patterns of expression relative to the 3' protein-coding exon signature from the corresponding genomic *loci*: (A) 41 pairs were identified in which the 3' exon of the protein-coding transcript (right panel) follows a similar expression pattern to that of the antisense TIN RNA (left panel); (B) 18 pairs were identified in which the 3' exon of the protein-coding transcript (right panel) follows a pattern of tissue expression inverted in relation to that of the antisense TIN RNA (left panel); (C) 71 pairs were identified in which

the 3' exon of the protein-coding transcript (right panel) follows a similar expression pattern as that of the sense TIN RNA (left panel); (D) 23 pairs were identified where the 3' exon of the protein-coding transcript (right panel) follows a pattern of tissue expression inverted in relation to that of the sense TIN RNA (left panel). For each line in each panel, expression intensities among the three tissues were normalized within each type of probe and colored as a function of the number of standard deviations from the mean value.

Table 1. Evidence of intronic transcription in the human mRNA/RefSeq GenBank dataset

	mRNA clusters with overlap to exons of non-redundant RefSeq dataset ^a		mRNA clusters wholly intronic to non-redundant RefSeq dataset		mRNA clusters not mapped to RefSeq dataset	Total
	antisense direction	sense direction	antisense direction	sense direction		
Spliced mRNA clusters ^b	2,559 (1,414) ^c	14,575 (14,369 ^d)	1,049 (378)	780 (223)	4,181 (0)	23,144 (16,384)
Unspliced mRNA clusters ^b	1,672 (26)	7,463 (87)	1,456 (56)	4,222 (87)	7,180 (927)	21,993 (1,183)
Total	4,231 (1,440)	22,038 (14,456)	2,505 (434)	5,002 (310)	11,361 (927)	45,137 (17,567)

^a The non-redundant dataset comprises 15,783 spliced RefSeq units. This was defined by mapping to the human genome sequence the total of 22,458 RefSeq sequences from GenBank, excluding 1,184 unspliced RefSeq and 601 RefSeq that were wholly intronic to another RefSeq and merging the remaining 20,673 spliced RefSeq sequences that mapped to the same locus into 15,783 spliced non-redundant RefSeq units (a total of 4,890 RefSeq that represent isoforms of the same gene were thus merged into these units).

^b mRNA clusters were obtained by mapping to the human genome sequence a total of 161,993 mRNA sequences followed by merging sequences with exon overlapping coordinates (see Methods for details), resulting in a non-redundant set of 45,137 mRNA clusters. This set was aligned to the non-redundant RefSeq dataset and each mRNA cluster was classified as exonic, wholly intronic or mapping outside of any spliced non-redundant RefSeq unit. Sense/antisense orientation was annotated.

^c For each class, the number of mRNA clusters containing at least one RefSeq is shown in parenthesis.

^d Excluding from the 15,783 spliced non-redundant RefSeq dataset a total of 1,414 RefSeq that map in the antisense direction with respect to another RefSeq.

Table 2. Classification of GenBank ESTs with respect to their genome mapping coordinates in relation to the set of non-redundant spliced RefSeq sequences

	EST clusters with overlap to exons of RefSeq genes ^a	EST clusters wholly intronic to RefSeq genes	EST clusters mapped outside of RefSeq genes	Total
Spliced EST contigs	16,241	8,013	10,144	34,398
Number of exons of spliced EST contigs (median)	10	2	3	
Total number of spliced ESTs in contigs	3,616,644	162,841	241,049	4,020,534
Number of spliced ESTs per contig (median)	91	3	4	
Unspliced EST contigs	4,030	55,139	29,198	88,367
Total number of unspliced ESTs in contigs	56,752	190,583	140,091	387,426
Number of unspliced ESTs per contig (median)	4	2	2	
Spliced EST singlets	1,053	6,205	6,631	13,889
Unspliced EST singlets	3,539	121,091	71,662	196,292
Total non-redundant EST clusters (contigs + singlets)	24,863	190,448	117,635	332,946
Total ESTs	3,677,988	480,720	459,433	4,618,141

^(a)The reference dataset comprises 15,783 spliced non-redundant RefSeq units plus the evidence of additional splice variants obtained for each transcriptional unit from all mRNA sequences mapping to the same locus.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

Figure 7

Figure 8

Figure 9

Figure 9

Figure 10

Figure 10

Figure 11

Figure 12

Figure 12

Additional files provided with this submission:

Additional file 11 : supplementary-table-s7.xls : 66Kb
<http://genomebiology.com/imedia/1358344519118565/sup11.XLS>

Additional file 10 : supplementary-table-s6.xls : 46Kb
<http://genomebiology.com/imedia/6796966351185651/sup10.XLS>

Additional file 9 : supplementary-table-s5.xls : 170Kb
<http://genomebiology.com/imedia/5664767221185651/sup9.XLS>

Additional file 8 : supplementary-table-s4.xls : 132Kb
<http://genomebiology.com/imedia/1491049845118565/sup8.XLS>

Additional file 7 : supplementary-table-s3.xls : 143Kb
<http://genomebiology.com/imedia/1936197431185651/sup7.XLS>

Additional file 6 : supplementary-table-s2.xls : 104Kb
<http://genomebiology.com/imedia/3233271031185651/sup6.XLS>

Additional file 5 : supplementary-table-s1.xls : 59Kb
<http://genomebiology.com/imedia/1813966979118565/sup5.XLS>

Additional file 4 : nakaya-et-al-genome-biology-figure-s4.pdf : 707Kb
<http://genomebiology.com/imedia/1683757471118565/sup4.PDF>

Additional file 3 : nakaya-et-al-genome-biology-figure-s3.pdf : 162Kb
<http://genomebiology.com/imedia/2126655824118565/sup3.PDF>

Additional file 2 : nakaya-et-al-genome-biology-figure-s2.pdf : 264Kb
<http://genomebiology.com/imedia/8524757541185651/sup2.PDF>

Additional file 1 : nakaya-et-al-genome-biology-figure-s1.pdf : 128Kb
<http://genomebiology.com/imedia/6976206551185651/sup1.PDF>

Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer

Eduardo M Reis^{1,2}, Helder I Nakaya¹, Rodrigo Louro¹, Flavio C Canavez¹, Áurea VF Flatschart¹, Giuliana T Almeida¹, Camila M Egidio¹, Apuã C Paquola², Abimael A Machado², Fernanda Festa¹, Denise Yamamoto¹, Renato Alvarenga¹, Camille C da Silva¹, Glauber C Brito², Sérgio D Simon^{3,4}, Carlos A Moreira-Filho^{3,4}, Katia R Leite⁵, Luiz H Camara-Lopes⁵, Franz S Campos⁶, Etel Gimba⁶, Giselle M Vignal⁶, Hamza El-Dorri¹, Mari C Sogayar¹, Marcello A Barcinski^{4,6}, Aline M da Silva¹ and Sergio Verjovski-Almeida^{*,1}

¹Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo, SP, Brasil; ²Laboratório de Bioinformática, Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo, SP, Brasil; ³Divisão de Oncologia, Hospital Israelita Albert Einstein, 05651-901 São Paulo, SP, Brasil; ⁴Instituto de Ciências Biomédicas, Universidade de São Paulo, 05508-900 São Paulo, SP, Brasil; ⁵Laboratório de Patologia Cirúrgica e Molecular, Hospital Sírio Libanês, 01308-050 São Paulo, SP, Brasil; ⁶Instituto Nacional de Câncer/MS, 20230-130 Rio de Janeiro, RJ, Brasil

A large fraction of transcripts are expressed antisense to introns of known genes in the human genome. Here we show the construction and use of a cDNA microarray platform enriched in intronic transcripts to assess their biological relevance in pathological conditions. To validate the approach, prostate cancer was used as a model, and 27 patient tumor samples with Gleason scores ranging from 5 to 10 were analyzed. We find that a considerably higher fraction (6.6%, [23/346]) of intronic transcripts are significantly correlated ($P \leq 0.001$) to the degree of prostate tumor differentiation (Gleason score) when compared to transcripts from unannotated genomic regions (1%, [6/539]) or from exons of known genes (2%, [27/1369]). Among the top twelve transcripts most correlated to tumor differentiation, six are antisense intronic messages as shown by orientation-specific RT-PCR or Northern blot analysis with strand-specific riboprobe. Orientation-specific real-time RT-PCR with six tumor samples, confirmed the correlation ($P = 0.024$) between the low/high degrees of tumor differentiation and antisense intronic *RASSF1* transcript levels. The need to use intron arrays to reveal the transcriptome profile of antisense intronic RNA in cancer has clearly emerged. *Oncogene* advance online publication, 28 June 2004; doi:10.1038/sj.onc.1207880

Keywords: prostate cancer; Gleason score; intronic transcripts; antisense RNA; microarray

Introduction

Microarray technology has permitted a refined high-throughput mapping of the transcriptional activity in

the human genome (Shoemaker *et al.*, 2001; Kapranov *et al.*, 2002; Rinn *et al.*, 2003; Kampa *et al.*, 2004), revealing a two- to ten-fold excess of transcripts along chromosome 22 (Kapranov *et al.*, 2002; Rinn *et al.*, 2003) than originally predicted by mapping of known genes (Lander *et al.*, 2001; Venter *et al.*, 2001). Using microarrays that probed the genomic DNA along all the extension of chromosomes 21 and 22, an approach that is essentially not biased towards the coding regions, comparable fractions of transcriptional activity were detected within exons or introns of annotated genes, and nearly half of these intronic transcripts were expressed antisense to their respective well-characterized introns (Rinn *et al.*, 2003; Kampa *et al.*, 2004). In other recent work, widespread occurrence of antisense transcription in the human genome has been detected (Shendure and Church, 2002; Yelin *et al.*, 2003). However, intronic non-spliced transcripts were not appreciated in these latter reports, because the authors either selected transcripts with evidence of splicing (Shendure and Church, 2002) or pairs of sense-antisense messages for which at least one exon was collinear on the genome sequence (Yelin *et al.*, 2003). In fact, large-scale transcription profiling has mainly been performed with exon arrays, which rely on the annotations of the human genome to choose which genomic protein-coding regions to sample.

Our goal was to study the transcriptional profile of intronic messages in cancer, using prostate tumors with different degrees of differentiation as a model. Therefore, we constructed a cDNA microarray that is enriched with a collection of partial transcripts that map to intronic segments of known genes. We used a bioinformatics approach to select these from nearly one million ESTs that were generated by the Human Cancer Genome Project with the ORESTES technique (Dias-Neto *et al.*, 2000) from 24 distinct types of cancer and normal adjacent tissues, having sampled over 60% of the human genes (Camargo *et al.*, 2001; Brentani *et al.*,

*Correspondence: S Verjovski-Almeida; E-mail: verjo@iq.usp.br
Received 16 March 2004; revised 7 May 2004; accepted 7 May 2004

2003). The resulting intron array was used to measure the expression profile of 27 prostate tumor samples with different degrees of differentiation (Gleason score, GS 5 to 10). Our results show that a considerable fraction of tumor-correlated genes are intronic fragments, pointing to the biological relevance of these messages in complex diseases such as cancer.

Results

Construction of an intron array

To identify a set of intronic transcripts represented in the public human EST databases that could be used for the construction of an intron array, we have compared the relative genome mapping coordinates of ORESTES reads with those from a curated set of known human transcripts (see *Materials and Methods* section for details). We found 150 236 unspliced ORESTES that were totally contained within 37 899 unique intronic regions of well-annotated human genes (RefSeq) (Table 1) and 136 628 ORESTES that mapped to 45 400 unannotated genomic regions. From these sets, we randomly sampled 619 and 1019 clones, respectively, representing 2% of each of these unique genomic regions, to be deposited in our cDNA microarray. In addition, 2048 gene fragments were selected representing known exons of RefSeq genes previously described as involved in cancer.

Intronic transcripts are expressed in prostate

To verify which fraction of these putative novel transcripts was expressed in prostate, we have measured their abundance by microarray hybridization with poly(A)⁺ mRNA from tumor and normal adjacent tissues from 27 patients. As shown in Table 1, we detected expression of comparable fractions of transcripts that mapped to intronic regions of RefSeq, to unannotated regions of the human genome, and to known exons of RefSeq genes. Finding that most of the unspliced transcripts that are totally contained in introns were expressed in prostate (Table 1) strongly

argues against a conservative approach of *ab initio* excluding from a transcriptome analysis the intronic unspliced messages as putative contaminants (Sorek and Safer, 2003). It is apparent that screening of the human transcriptome for diseases with commercial exon microarrays, primarily containing messages of previously characterized full-length genes, does not address the possible change in the levels of intronic messages with putative regulatory roles in pathological conditions.

Non-coding intronic transcript levels correlate to the degree of prostate tumor differentiation

To search for intronic and novel transcripts whose expression would correlate to the degree of tumor differentiation (GS), we analyzed the expression profile of the 27 prostate tumor samples with different GS from 5 to 10. A Self-Organizing Map and an unsupervised hierarchical clustering analysis revealed that all high-GS (9 and 10) samples clustered together, indicating that a distinct expression profile of high-GS tumors was present (Figures 1a, b). Bootstrap resampling was used to confirm that grouping of high-GS samples in the hierarchical clustering analysis is extremely robust (Figure 1b). In fact, recent reports have identified genes that are strongly correlated with the tumor GS in prostate cancer (Singh *et al.*, 2002; Best *et al.*, 2003; Lapointe *et al.*, 2004). We then used Pearson correlation and bootstrapping to identify messages significantly correlated to GS among the low- (5 and 6) and high-GS samples. We identified a classifier set of 56 genes for which the expression level was significantly correlated (P -value ≤ 0.001) to the low-/high-GS class distinction (Figure 2a). We found that within this set, 27 are known exons of RefSeq genes, 23 are intronic transcripts and 6 are ESTs that map to unannotated genomic sequence.

A cross-validation leave-one-out procedure was used to evaluate the robustness of the low-/high-GS gene classifier. Genes that appear in at least 60% of the classifiers are listed in Figures 2b and c. Noteworthy, seven out of the twelve gene fragments shown in Figures 2b and c map to intronic regions of RefSeq genes. The other five are known exons of RefSeq genes, and three

Table 1 Transcriptional activity in prostate of gene fragments mapping to different segments of the human genome and represented by ORESTES sequences

ORESTES mapping to:	Number of unique genomic regions	Number of ORESTES clones		
		Total in the collection	Total in the microarray	Expressed in prostate
Introns of RefSeq genes	37 899	150 236	619	346 (56%)
Unannotated regions of the genome	45 400	136 628	1019	539 (53%)
Known exons of RefSeq genes	29 444	390 263	2048	1369 (67%)

'Total in the collection' refers to the number of clones in the collection generated by the Human Cancer Genome Project (Dias-Neto *et al.*, 2000). 'Expressed in prostate' refers to the gene fragments whose background-subtracted intensities were at least 3 s.d. above the average microarray slide background intensity in at least 5 prostate samples (out of 47). The percentage of expressed fragments with respect to the total number of fragments of that category in the microarray is indicated in parenthesis.

of these, *EIF4E*, *TMEPA1* and *RNF139*, have been associated with progression of different types of cancer (Gemmell *et al.*, 1998; Brunschwig *et al.*, 2003; Ramaswamy *et al.*, 2003).

When the low-/high-GS classifier was applied to the expression profile of sixteen intermediate-grade GS 7 tumor samples, using a supervised method (van 't Veer *et al.*, 2002), two groups of patients were identified (Figure 3). Thus, while most of the GS 7 tumor samples exhibited a positive correlation to the average expression profile of the low-GS tumors, four GS 7 samples (patients 17, 19, 47 and 3) were identified that are negatively correlated to the low-GS profile. These results are in line with the fact that patients with intermediate-grade GS 7 tumors have a heterogeneous clinical outcome (Epstein *et al.*, 1996).

Characterization of cancer-related intronic transcripts

To characterize the seven predicted intronic transcripts identified by the low-/high-GS gene classifier, we performed orientation-specific reverse transcription (RT) followed by PCR, which showed that six of these were transcribed antisense to the corresponding intron (Figure 2d). Interestingly, we found that for these six genes an intronic sense message overlapping the intronic antisense sequence was also present in the mRNA sample (Figure 2d). The implications of this finding will be discussed below.

We found that the remaining predicted intronic message was not expressed antisense, but rather it is a novel exon of *TLE3* (Figure 4a), a member of the Notch signaling pathway involved in epithelial differentiation (Liu *et al.*, 1996). Northern blot showed a 5.53 kbp band, which is compatible with a full-length alternatively spliced *TLE3* transcript (Figure 4b). The additional exon introduces an early stop codon that truncates TLE carboxyl-terminal WD-40 domain and leaves its amino-terminal Q-rich domain; these truncated forms are expressed by most metazoans and may negatively regulate full-length TLE proteins, perhaps by sequestering them in non-productive complexes (Chen and Courey, 2000).

We used RACE-PCR and/or strand-specific Northern blot to confirm the orientation of three out of the six correlated intronic messages. RACE-PCR determined that they were 0.6 to 1.1 kbp long transcripts that mapped to introns of different genes in different

Figure 1 Unsupervised classification of 27 prostate tumor samples. (a) Self-organizing map (Tamayo *et al.*, 1999). (b) Hierarchical clustering (Sanchez-Carbayo *et al.*, 2002). Colored circles below each panel indicate the Gleason score (GS) of tumor samples, which are shown solely as a reference, since the classification was unsupervised. Note that all high-GS samples (GS 9 and 10) are clustered together in **a** and **b**, which indicates that there is a distinct pattern of differential expression underlying the dataset, when all 3576 ORESTES gene fragments in the microarray (rows in **b**) and the 27 tumor samples (columns in **b**) are analyzed together. Expression level of each gene is represented by the number of standard deviations above (red) or below (blue) the average value for that gene across all samples. The numbers shown in the tree indicate the robustness of the clustering (in a 10000 bootstrap resampling) as described under Supplementary Materials and methods available on line

Figure 2 Intronic transcripts correlated to the degree of prostate tumor differentiation. **(a)** Left panel shows the expression matrix of 56 genes correlated to the low-/high-GS class distinction (P -value ≤ 0.001). Genes (columns) are ordered by their Pearson correlation and patients (rows) by their correlation (r) to the low-GS profile, which is shown in the right panel. Intronic gene fragments have their names in blue. Expression level of each gene is represented by the number of standard deviations above (red) or below (blue) the average value for that gene across all samples. **(b)** and **(c)** Top 12 genes in the low-/high-GS classifier. Panel **b** shows the correlated genes (P -value ≤ 0.001) appearing in at least 60% of the classifiers (fraction indicated in the x-axis), as identified by a leave-one-out cross-validation test; dark-green bars correspond to known exons of RefSeq genes, and light-green bars are intronic transcripts. Panel **c** shows the expression level of each corresponding gene (row) in each tumor sample (column), as in panel **a**. **(d)** Intronic messages are detected in both sense and antisense strands. We used mRNA from DU145 prostate tumor cell line to perform orientation-specific reverse transcription (+ RT) using either an antisense or a sense specific primer, followed by PCR using a nested pair of primers for an amplicon within the intronic region of each indicated gene. Controls omitting either the reverse transcriptase (– RT/+ primer) or omitting the primer during reverse transcription (+ RT/– primer) were performed. An additional control is shown for an amplicon within a known exon of *TMEPA1*, where only the sense transcript was detected

Figure 3 Expression levels of the 56 correlated genes measured on intermediate-grade (GS 7) tumor samples. Left panel shows the expression matrix of 56 genes correlated to the low-/high-GS class distinction that were identified in the experiment described in Figure 2a. Genes (columns) are ordered by their Pearson correlation. Sixteen GS-7 patient samples were analyzed (rows) and they are ordered by their correlation (r) to the low-GS profile, which is shown in the right panel. Intronic gene fragments have their names in blue. Expression level of each gene is represented by the number of standard deviations above (red) or below (blue) the average value for that gene across all samples

Figure 4 Characterization of cancer-related intronic transcripts. Mapping of each original double-stranded ORESTES cDNA clone that is on the microarray is represented in blue, and chromosome coordinates are indicated. (**a** and **b**) Alternatively spliced form of *TLE3*, which was confirmed by RACE-PCR (red block, in panel **a**). In panel **b**, a 5.53 kbp transcript was detected by Northern blot in three different prostate cancer cell lines, using a probe complementary to the sense strand of the novel *TLE3* exon. (**c–g**) Characterization of antisense intronic transcripts for *RASSF1*, *PPHLN1* and *APIGBP1*. Red arrows show transcripts obtained by orientation-specific RACE-PCR with a colon cDNA library. Yellow blocks show transcripts obtained by non-directional RACE-PCR with a lung cDNA library. Panels **d** and **f** show strand-specific Northern blots using probes complementary to the antisense transcript from *RASSF1* (panel **d**) or *PPHLN1* (panel **f**)

chromosomes (Figures 4c, e and g, yellow bars). Interestingly, these are unspliced messages that do not have coding potential, as determined by ESTScan (Iseli *et al.*, 1999). For *RASSF1* and *APIGBP1*, an orientation-specific RACE-PCR further verified the presence of an antisense message (Figures 4c and g, red arrows). Antisense orientation of *RASSF1* and *PPHLN1* was confirmed by strand-specific Northern blot (Figures 4d and f). RACE-PCR products of the remaining 3 intronic transcripts did not extend the ORESTES cDNA clones with the libraries used.

Validation by Real-Time PCR of down-regulation of *RASSF1* intronic antisense messages

Double-stranded cDNA microarrays are unable to discriminate between sense and antisense overlapping messages. Thus, finding differential levels of expression in overlapping intronic transcripts can reflect a change in either one or in both messages. To discriminate between the two possibilities, we used quantitative orientation-specific real-time RT-PCR to confirm the differential expression of *RASSF1*, the most correlated

intronic transcript. Figure 5a shows that both antisense and sense intronic transcripts are significantly down-regulated in high-GS as compared to low-GS tumor samples (P -values = 0.024 and = 0.014, respectively). *RASSF1* hypermethylation correlates with the GS of

Figure 5 *RASSF1* expression is down regulated in high-GS tumors. Orientation-specific reverse transcription followed by Real-time quantitative PCR was used for the measurement of mRNA levels. Relative fractions were calculated with respect to the average mRNA level across all samples. (a) Antisense (dashed bars) and sense (full bars) *RASSF1* intronic transcript levels were measured in 3 low- (grey bars) and 3 high- (black bars) GS tumor samples. (b) Protein-encoding *RASSF1* exon transcript levels, measured as in a. (c) Protein-encoding *RASSF1* exon transcript levels measured in GS 7 tumor samples that were either negatively correlated (black bars) or positively correlated (grey bars) to the average low-GS expression profile in the microarray experiments of Figure 3. Standard error bars from 3 replicate experiments are shown in a–c

prostate tumors (Liu *et al.*, 2002). In line with that, we show here that *RASSF1* protein-encoding exonic transcripts are significantly down-regulated in high-GS tumor samples (P -value = 0.009) (Figure 5b). We also confirmed in the experiment of Figure 5c that *RASSF1* protein-encoding exonic transcript level is lower (P -value = 0.002) in GS 7 tumors whose expression profiles in the microarray were found to be negatively correlated with the average low-GS pattern (patients 17, 19, 47 and 3) (see Figure 3).

Discussion

Our microarray expression study suggests that the majority of the over 37 000 unique intronic regions of known genes, which were mapped by ORESTES sequences throughout the chromosomes, have indeed transcriptional activity. This adds to the finding of significant intronic transcriptional activities mapped to chromosomes 21 and 22 (Kapranov *et al.*, 2002; Rinn *et al.*, 2003; Kampa *et al.*, 2004). Our finding of unspliced long antisense intronic RNAs is in agreement with the recent report of 0.4–1.3 kbp unspliced antisense totally intronic RNAs in chromosome 22 (Rinn *et al.*, 2003).

We simultaneously detected in a poly(A)⁺ RNA sample the presence of antisense and sense pairs of overlapping intronic transcripts that are correlated to the degree of tumor differentiation and map to 6 different genomic loci. Totally intronic antisense transcripts would have to target their complementary intronic sense message (in pre-mRNA) to play a regulatory role in processing protein-encoding RNA messages. Similarly, totally intronic antisense messages have been detected by a genome-wide mapping analysis in mice (Kiyosawa *et al.*, 2003). Antisense RNA transcripts have been previously implicated with imprinting and X-chromosome inactivation (Lee *et al.*, 1999; Sleutels *et al.*, 2000). Antisense RNA has also been shown to mediate gene silencing and methylation in alpha-thalassemia, a genetic autosomal disease (Tufarelli *et al.*, 2003). Moreover, some antisense intronic transcripts have been shown to act as negative effectors of the cancer-related genes cMyc (Nepveu and Marcu, 1986) and p53 (Khochbin and Lawrence, 1989) and as positive effector of *RPS14* gene transcription (Tasheva and Roufa, 1995). We speculate that sense/antisense intronic transcript pairs may be part of a general regulatory mechanism of gene expression not yet fully appreciated in eukaryotes. In fact, a recent unbiased mapping of transcription factor binding sites along chromosomes 21 and 22 revealed that 36% of the binding sites for Sp1, cMyc and p53 are significantly correlated with non-coding RNAs, and overlapping pairs of protein-coding and non-coding RNAs are often co-regulated (Cawley *et al.*, 2004).

Both sense and antisense intronic transcript levels of *RASSF1* were diminished in high-GS tumors. In fact, by analyzing the fraction of tumor-correlated genes with

respect to all genes expressed in prostate, we observed a considerably higher fraction of intronic transcripts (6.6%, [23/346]) when compared to transcripts mapping to unannotated genomic regions (1%, [6/539]) or to exons of RefSeq genes (2%, [27/1,369]). As evidence is continuing to accumulate that the transcriptional output of the human genome is far more complex than anticipated, so are the questions as to the biological relevance of these novel and apparently non-coding transcripts. Correlation with a stage of tumor progression argues that such transcripts are not just random biological noise. This finding argues for inclusion of probes to such non-coding transcripts, both intra- and inter-genic, into the arsenal of tools used for molecular diagnostics, so far almost exclusively populated by assays of protein-coding transcripts.

A paradigm in oncogenesis is the accumulation of mutations in the open reading frames of protein-encoding oncogenes and tumor suppressors (McManus, 2003). The implication in cancer of various classes of non-coding RNAs such as microRNAs (McManus, 2003; Calin *et al.*, 2004) is becoming evident. We suggest that identification of antisense intronic RNAs correlated to cancer, such as those described here, will likely increase the complexity of the current paradigm.

Materials and methods

Microarray construction

A bioinformatics approach was used to select partial transcripts that map to intronic segments of known genes. The initial sequence dataset contained 1 187 342 ORESTES cDNAs that were generated by the Human Cancer Genome Project with the ORESTES technique (Dias-Neto *et al.*, 2000). 956 456 filtered ORESTES and 20 793 RefSeq sequences were aligned to the genome sequence (UCSC draft HG12 Jun2002) using BLAT. Genomic coordinates of each ORESTES were compared to the exon/intron genomic coordinates of all RefSeq sequences and three categories were defined: 1) ORESTES that were collinear to a RefSeq and had at least 5 bases overlapping a known exon were annotated as known exons of RefSeq genes; 2) those that were collinear to a RefSeq and did not overlap any known exon were annotated as mapping to introns of RefSeq genes; 3) ORESTES non-collinear to RefSeq and with no match to full-length or partial mRNAs from GenBank were categorized as mapping to unannotated genomic regions. Within each of these categories we determined the number of unique genomic regions to which the ORESTES were mapped by merging overlapping sequences. Glass-slides were constructed containing 3686 cDNA fragments in duplicate, in a 9216 elements microarray, using Amersham Biosciences Generation III microarray system. Detailed information including cDNA amplification is available in the Supplementary Materials and methods available online.

Patient samples and target labeling

Samples of prostate tumor were obtained from freshly-frozen tissue collections with informed consent from patients submitted to radical prostatectomy and were snap-frozen in liquid nitrogen within 10 min from resection. Each sample was reviewed by a pathologist for determination of the degree of tumor differentiation, which is measured by the Gleason score

(GS) of the tumor (Gleason, 1977). Prior to RNA extraction, samples were re-examined and hematoxylin/eosin stained micro-sections obtained from each side of the frozen blocks were used to grossly delimit the spatial distribution of the tumor mass. If necessary, tissue blocks were further dissected to warrant that at least 70% of the section used for RNA extraction was composed of malignant cells. Macro-dissected tumor samples were returned to liquid nitrogen until use. Poly(A)⁺ RNA was isolated using μ MACS mRNA isolation kit (Miltényi Biotec). Each sample was labeled in parallel with either Cy3- or Cy5-labelled nucleotides (CyScribe first-strand cDNA labeling kit, Amersham Biosciences) using 1 μ g mRNA and a mixture of oligo-dT and 9-mer random oligonucleotides. Control experiments with DNase-treated samples performed in parallel showed that some contaminant genomic DNA was labeled at a minor extent, as analyzed by the ratio of exon/(intron + unannotated) fragments that are detected in the microarray (see Supplementary Materials and Methods). The fraction of intronic or unannotated transcripts expressed in prostate might be overestimated by a maximum of 11% due to this contaminant. If present, genomic DNA contamination is expected to be randomly distributed across all genes in all samples. Therefore no particular gene is expected to be identified as statistically correlated to tumor differentiation in the analyses due to this contamination. Correlated intronic messages detected in the microarray experiments were further validated by strand-specific RT-PCR, RACE-PCR and/or strand-specific Northern blot (see below). Microarray data acquisition and processing is described under Supplementary Materials and Methods available online.

Statistical analyses

Unsupervised clustering of the data was performed either with Self-organizing map using the GeneCluster2 analysis package (Tamayo *et al.*, 1999) or with Hierarchical clustering (Sanchez-Carbayo *et al.*, 2002), using Euclidean distance metric and Ward linkage. We assessed the robustness of the unsupervised clustering analysis by a bootstrap resampling technique (Sanchez-Carbayo *et al.*, 2002). This was used to generate 10 000 copies of the data set by adding Gaussian noise to the original data, as described in detail in the Supplementary Materials and Methods available online. A consensus tree was constructed by finding for each node the pairing that occurred most often in the 10 000 separate trials. A graph was constructed that displays this count at each node of the tree. Nodes with values closer to 10 000 are more robust than with lower ones. A supervised selection of genes whose expression is significantly correlated to tumor GS score classes was performed as previously described by Singh *et al.* (2002) using Pearson correlation analysis. Statistical significance was ascertained by re-calculating the correlation following 50 000 random permutations of GS sample labels and computing the frequency at which the correlation measured in the original set was observed in the randomly permuted data. The amount of significantly correlated genes (with P -value ≤ 0.001) exceeds by eight-fold the number of genes expected to be found by chance alone in the dataset using the same confidence interval (7 genes in a two-tailed distribution). Robustness of the low-/high-GS gene classifier was evaluated by patient leave-one-out cross-validation (Singh *et al.*, 2002). Essentially one patient is removed and a new set of significantly correlated genes (P -value ≤ 0.001) is determined using the remainder samples. This procedure was repeated for each one of the 11 patient samples and correlated genes were ordered by the frequency at which they appear in the various sets thus identifying the most robust classifiers.

Orientation-specific RT-PCR and RACE-PCR

For orientation-specific RT-PCR 0.5 µg. poly(A)⁺ RNA was treated with 1 U Promega RQ1 RNase-free DNase for 30–60 min at 37°C in a 15 µl volume in the presence of 20 U RNasin ribonuclease inhibitor (Promega). DNase was inactivated at 65°C for 10 min and the sample was kept at 50°C. cDNA was generated at 50°C for 45–60 min in a 50 µl volume using 100 ng of mRNA, 200 units of SuperscriptII reverse transcriptase (Invitrogen) plus 0.36–0.9 µM of a specific primer for either the sense or the antisense strand of each intronic region (Cawley *et al.*, 2004; Shendure and Church, 2002). The reverse transcription primer mapped within 30 to 60 bp upstream from the target intronic segment of the RNA to be reverse transcribed. Controls for the absence of self-priming were done with reverse transcriptase in the absence of primers, and controls for the absence of DNA were done by incubation with primers and no reverse transcriptase. A control with a primer complementary to the coding region of beta-tubulin was done in parallel. The sample was heated for 30 min at 96°C to denature the reverse transcriptase and one pair of primers for PCR amplification of either the beta-tubulin gene or the desired target intronic segment under study were added; reactions were carried out for 40 cycles in a volume of 20 µl and according to the manufacturer's instructions for Sybr-Green PCR Core Reagent using the GenAmp 5700 Sequence Detector (Applied Biosystems). For RACE-PCR we used commercial cDNA libraries prepared from lung (cat. 7445-1) and colon (cat. HL3014B) poly(A)⁺ RNA respectively (BD Biosciences – Palo Alto, CA). The colon cDNA library is

directionally ligated into a lambda gt11 vector, which allows the determination of the orientation of cloned transcripts. RACE-PCR was performed with Taq DNA Polymerase (Bioline – Canton, MA). Additional details are available as Supplementary Materials and Methods available online.

Accession numbers

TLE3, AY566265; antisense *RASSF1*, AY545527 and AY545528; antisense *PPHLN1*, AY545526; antisense *AP1GBPI*, AY545524 and AY545525.

Abbreviations

GS, Gleason score.

Acknowledgements

The authors are greatly indebted to H.A. Armelin for supporting this work, which is part of the Cooperation for Analysis of Gene Expression (CAGE) inter-departmental project that he coordinates. We thank E. Dias-Neto, R. R. Brentani and R. DeMarco for valuable suggestions and for critically reading the manuscript. E.M.R. is the recipient of a grant award from Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior (CAPES). This work was financed by fellowships from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and by a grant and fellowships from Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP).

References

- Best CJ, Leiva IM, Chuaqui RF, Gillespie JW, Duray PH, Murgai M, Zhao Y, Simon R, Kang JJ, Green JE, Bostwick DG, Linehan WM and Emmert-Buck MR. (2003). *Diagn. Mol. Pathol.*, **12**, 63–70.
- Brentani H, Caballero OL, Camargo AA, Da Silva AM, Da Silva Jr WA, Neto ED, Grivet M, Gruber A, Guimaraes PE, Hide W, Iseli C, Jongeneel CV, Kelso J, Nagai MA, Ojopi EP, Osorio EC, Reis EM, Riggins GJ, Simpson AJ, De Souza S, Stevenson BJ, Strausberg RL, Tajara EH and Verjovski-Almeida S. (2003). *Proc. Natl. Acad. Sci. USA*, **100**, 13418–13423.
- Brunschwig EB, Wilson K, Mack D, Dawson D, Lawrence E, Willson JK, Lu S, Nosrati A, Rerko RM, Swinler S, Beard L, Lutterbaugh JD, Willis J, Platzer P and Markowitz S. (2003). *Cancer Res.*, **63**, 1568–1575.
- Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M and Croce CM. (2004). *Proc. Natl. Acad. Sci. USA*, **101**, 2999–3004.
- Camargo AA, Samaia HP, Dias-Neto E, Simao DF, Migotto IA, Briones MR, Costa FF, Nagai MA, Verjovski-Almeida S, Zago MA, Andrade LE, Carrer H, El-Dorry HF, Espreafico EM, Habr-Gama A, Giannella-Neto D, Goldman GH, Gruber A, Hackel C, Kimura ET, Maciel RM, Marie SK, Martins EA, Nobrega MP, Paco-Larson ML, Pardini MI, Pereira GG, Pesquero JB, Rodrigues V, Rogatto SR, da Silva ID, Sogayar MC, Sonati MF, Tajara EH, Valentini SR, Alberto FL, Amaral ME, Aneas I, Arnaldi LA, de Assis AM, Bengtson MH, Bergamo NA, Bombonato V, de Camargo ME, Canevari RA, Carraro DM, Cerutti JM, Correa ML, Correa RF, Costa MC, Curcio C, Hokama PO, Ferreira AJ, Furuzawa GK, Gushiken T, Ho PL, Kimura E, Krieger JE, Leite LC, Majumder P, Marins M, Marques ER, Melo AS, Melo M, Mestriner CA, Miracca EC, Miranda DC, Nascimento AL, Nobrega FG, Ojopi EP, Pandolfi JR, Pessoa LG, Prevedel AC, Rahal P, Rainho CA, Reis EM, Ribeiro ML, da Ros N, da Sa RG, Slaes MM, Santa'anna SC, dos Santos ML, da Silva AM, da Silva NP, Silva Jr WA, da Silveira RA, Sousa JF, Stecconi D, Tsukumo F, Valente V, Soares F, Moreira ES, Nunes DN, Correa RG, Zalberg H, Carvalho AF, Reis LF, Brentani RR, Simpson AJ and de Souza SJ. (2001). *Proc. Natl. Acad. Sci. USA*, **98**, 12103–12108.
- Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K and Gingeras TR. (2004). *Cell*, **116**, 499–509.
- Chen G and Courey AJ. (2000). *Gene*, **249**, 1–16.
- Dias-Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva Jr W, Zago MA, Bordin S, Costa FF, Goldman GH, Carvalho AF, Matsukuma A, Baia GS, Simpson DH, Brunstein A, de Oliveira PS, Bucher P, Jongeneel CV, O'Hare MJ, Soares F, Brentani RR, Reis LF, de Souza SJ and Simpson AJ. (2000). *Proc. Natl. Acad. Sci. USA*, **97**, 3491–3496.
- Epstein JI, Partin AW, Sauvageot J and Walsh PC. (1996). *Am. J. Surg. Pathol.*, **20**, 286–292.
- Gemmill RM, West JD, Boldog F, Tanaka N, Robinson LJ, Smith DI, Li F and Drabkin HA. (1998). *Proc. Natl. Acad. Sci. USA*, **95**, 9572–9577.
- Gleason D. (1977). *Urologic Pathology* M Tannenbaum (ed). Lea & Febiger: Philadelphia, pp. 171–197.
- Iseli C, Jongeneel CV and Bucher P. (1999). *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, 138–148.
- Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, Drenkow J, Piccolboni A, Bekiranov S, Helt G,

- Tammana H and Gingeras TR. (2004). *Genome Res.*, **14**, 331–342.
- Kapranov P, Cawley SE, Drenkow J, Bekiranov S, Strausberg RL, Fodor SP and Gingeras TR. (2002). *Science*, **296**, 916–919.
- Khochbin S and Lawrence JJ. (1989). *EMBO J.*, **8**, 4107–4114.
- Kiyosawa H, Yamanaka I, Osato N, Kondo S and Haya-shizaki Y. (2003). *Genome Res.*, **13**, 1324–1334.
- Lander ES, Linton LM, Birren B, Nussbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczy J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chisoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, GU J, Hood L and Rowen L et al. (2001). *Nature*, **409**, 860–921.
- Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K and Ferrari M. (2004). *Proc. Natl. Acad. Sci. USA*, **101**, 811–816.
- Lee JT, Davidow LS and Warshawsky D. (1999). *Nat. Genet.*, **21**, 400–404.
- Liu L, Yoon JH, Dammann R and Pfeifer GP. (2002). *Oncogene*, **21**, 6835–6840.
- Liu Y, Dehni G, Purcell KJ, Sokolow J, Carcangiu ML, Artavanis-Tsakonas S and Stifani S. (1996). *Genomics*, **31**, 58–64.
- McManus MT. (2003). *Semin. Cancer Biol.*, **13**, 253–258.
- Nepveu A and Marcu KB. (1986). *EMBO J.*, **5**, 2859–2865.
- Ramaswamy S, Ross KN, Lander ES and Golub TR. (2003). *Nat. Genet.*, **33**, 49–54.
- Rinn JL, Euskirchen G, Bertone P, Martone R, Luscombe NM, Hartman S, Harrison PM, Nelson FK, Miller P, Gerstein M, Weissman S and Snyder M. (2003). *Genes Dev.*, **17**, 529–540.
- Sanchez-Carbayo M, Socci ND, Charytonowicz E, Lu M, Prystowsky M, Childs G and Cordon-Cardo C. (2002). *Cancer Res.*, **62**, 6973–6980.
- Shendure J and Church GM. (2002). *Genome Biol.*, **3**, research0044.1–research0044.14.
- Shoemaker DD, Schadt EE, Armour CD, He YD, Garrett-Engel P, McDonagh PD, Loerch PM, Leonardson A, Lum PY, Cavet G, Wu LF, Altschuler SJ, Edwards S, King J, Tsang JS, Schimmack G, Schelter JM, Koch J, Ziman M, Marton MJ, Li B, Cundiff P, Ward T, Castle J, Krolewski M, Meyer MR, Mao M, Burchard J, Kidd MJ, Dai H, Phillips JW, Linsley PS, Stoughton R, Scherer S and Boguski MS. (2001). *Nature*, **409**, 922–927.
- Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C, Tamayo P, Renshaw AA, D’Amico AV, Richie JP, Lander ES, Loda M, Kantoff PW, Golub TR and Sellers WR. (2002). *Cancer Cell*, **1**, 203–209.
- Seutels F, Barlow DP and Lyle R. (2000). *Curr. Opin. Genet. Dev.*, **10**, 229–233.
- Sorek R and Safer HM. (2003). *Nucleic Acids Res.*, **31**, 1067–1074.
- Tamayo P, Slonim D, Mesirov J, Zhu Q, Kitareewan S, Dmitrovsky E, Lander ES and Golub TR. (1999). *Proc. Natl. Acad. Sci. USA*, **96**, 2907–2912.
- Tasheva ES and Roufa DJ. (1995). *Genes Dev.*, **9**, 304–316.
- Tufarelli C, Stanley JA, Garrick D, Sharpe JA, Ayyub H, Wood WG and Higgs DR. (2003). *Nat. Genet.*, **34**, 157–165.
- van ‘t Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R and Friend SH. (2002). *Nature*, **415**, 530–536.
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Hsion DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J and Kalush F et al. (2001). *Science*, **291**, 1304–1351.
- Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A, Diber A, Biton S, Tamir Y, Khosravi R, Nemzer S, Pinner E, Walach S, Bernstein J, Savitsky K and Rotman G. (2003). *Nat. Biotechnol.*, **21**, 379–386.

Supplementary Information accompanies the paper on Oncogene website (<http://www.nature.com/onc>).

***RASL11A*, member of a novel small monomeric GTPase gene family, is down-regulated in prostate tumors[☆]**

Rodrigo Louro,^a Helder I. Nakaya,^a Apuã C.M. Paquola,^a Elizabeth A.L. Martins,^b
Aline M. da Silva,^a Sergio Verjovski-Almeida,^a and Eduardo M. Reis^{a,*}

^a Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-900 São Paulo, SP, Brazil

^b Centro de Biotecnologia, Instituto Butantan, 05503-900 São Paulo, SP, Brazil

Received 7 January 2004

Abstract

We performed a genome-wide search for novel *loci* encoding for Ras-related proteins based on the genome mapping coordinates of the cancer-derived EST dataset at GenBank. Partial sequences from two novel human genes were identified and subsequently used for full length transcript cloning. *RASL11A* and *ARL9* belong to two novel subfamilies coding for small GTPases that we found to be highly conserved among eukaryotes. The Arl9/Arl10 subfamily displays a conserved interswitch toggle that places it evolutionarily closer to the Arf family. Rasl11 proteins are more closely related to the Ras branch of GTPases. All orthologues newly identified here exhibit an Asn residue in place of the highly conserved Thr35 of the G domain, suggesting that the universal switch mechanism of small GTPases may be structurally different in this subfamily. We determined by Northern blot that *RASL11A* is transcribed in several human tissues and that it is down-regulated in prostate tumors as measured by quantitative real-time PCR. These results highlight a previously uncharacterized subfamily of Ras-related genes that may have a tumor suppressor role in prostate cancer.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Tumor suppressor; Ras-related GTPases; Prostate cancer

Proteins from the Ras superfamily of small GTPases modulate a variety of cellular effects ranging from growth control, cytoskeletal rearrangements, cell survival, and apoptosis, to intracellular traffic regulation [1,2]. Over 100 members of the Ras superfamily have been described in eukaryotic cells and tentatively classified into six families (Ras, Rho, Rab, Arf, Ran, and Rad) based on their sequence similarity to some well-characterized members [3,4]. In spite of the wide range of cellular functions, all small GTPases share a structurally conserved G domain that responds for their common molecular mechanism of switching between an active GTP-bound state and an inactive GDP-bound state [5].

Several proteins from the Ras superfamily have been associated to cancer. It has been documented that

mutations in prototypic *RAS* genes (*H-RAS*, *N-RAS* or *K-RAS*) occur in approximately 30% of all human cancers [6]. Rho GTPases were also implicated in processes leading to cancer progression (reviewed in [7]). While some members of the families Ras, Ral, Rit [8], and Rho [9] may present an oncoprotein behavior, others like Noey2/Arh1 [10,11], Rap1a [12], Rerg [13], and Rig [14] have been described as having a tumor suppressor activity.

The involvement of Ras-related GTPases in tumor development places these proteins as potential targets for cancer therapeutics. The main strategy under clinical evaluation targets the association of Ras-related proteins with the plasma membrane through the inhibition of their farnesylation [15]. Other approaches aim to interrupt down-stream signaling (Raf and MEK protein kinase inhibitors), autocrine growth factor signaling (EGF receptor inhibitors), or gene expression of Ras-related proteins such as H-Ras and c-Raf-1 (reviewed in [16]). Notwithstanding, the existence of alternate

[☆] Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.bbrc.2004.02.091](https://doi.org/10.1016/j.bbrc.2004.02.091).

* Corresponding author. Fax: +55-11-3091-2186.

E-mail address: emreis@iq.usp.br (E.M. Reis).

biochemical pathways that eventually turn cancer patients refractory to conventional drug therapies is still a major limitation of current chemotherapeutic strategies. Therefore, the characterization of novel Ras-like proteins is relevant to the development of effective treatment protocols. One way to unveil the full molecular repertoire of cancer-related genes is to explore public databases containing partial transcript sequences generated from human tumors [17]. We have developed a bioinformatics routine to screen the nearly 3 million EST datasets generated by the Human Cancer Genome project (HCGP—[18]) and the Cancer Genome Anatomy Project (CGAP—[19]), and to search for novel members of possible paralogue gene families related to the Ras superfamily. As a result, we identified ESTs that permitted cloning of the complete coding-region of two novel genes, *RASL11A* and *ARL9*, which are novel members of the Ras-superfamily. We found that the *RASL11A* family is extremely conserved among eukaryotes. We also found that the *RASL11A* transcript is ubiquitously expressed in human tissues and its abundance is diminished in prostate tumors in comparison to normal adjacent prostate tissue, suggesting that it might have a tumor suppressor role in prostate.

Materials and methods

Informatics identification of novel Ras-related small GTPases. The amino acid sequences of 36 proteins encoded by genes of the superfamily of Ras-related small GTPases were used to query a local copy of the nearly 3 million Expressed Sequence Tag database from HCGP [18] and CGAP [19] using Tblastn (NCBI) running on a BlastMachine (Paracel, Pasadena, CA). The Tblastn output was parsed using a lower-limit cut-off of 50% identity over 50 amino acids and the results were merged, thus identifying 2061 ESTs as putative new family member candidates. These cut-off parameters proved to be adequate to identify paralogues of cancer-related genes as part of the annotation effort of the Human Cancer Genome Project [20]. Nucleotide sequences from the selected ESTs, from the 36 Ras-related GTPases, as well as from a set of 113,179 full-length and partial human transcripts compiled from GenBank were mapped to a local copy of the full set of human genomic contigs (NCBI) using Blastn. The genomic coordinates of the alignments were parsed [21] and incorporated into a PostgreSQL relational database. Next, the database was queried to identify possible novel Ras-related ESTs by picking those that mapped to genomic regions not covered by the genomic coordinates of the known Ras-related GTPase genes under analysis, nor by any known full-length or partial human transcript. ESTs selected in the previous step were assembled with Cap3 using default parameters. The assembled EST dataset was compared to the GenBank non-redundant protein dataset (NR) using Blastx (cut-off *E*-value $\leq 10^{-10}$). The output was manually inspected, and two candidates for novel Ras-related proteins were selected for further biochemical experiments. These sequences were aligned to the human genome sequence using the BLAT tool (<http://genome.ucsc.edu>) in order to identify putative model full-length transcripts for the novel genes (GenBank Accession Nos. XM_172905 and XM_059617) that were predicted by automated computational analyses and aligned in the same genomic region.

Cloning of *RASL11A* and *ARL9*. Complete coding sequences of *RASL11A* and *ARL9* were obtained using either RACE (Marathon ready cDNA, Clontech) or PCR amplification with specific primers

designed to the 3' and 5' extremities of full-length transcript in silico models of the novel genes. High-quality commercial cDNA libraries derived from a human prostatic adenocarcinoma cell lineage or from human testis (Marathon ready cDNA, Clontech) were used as templates for gene cloning. Cloned sequences from *ARL9* (Accession No. AY439003) and *RASL11A* (Accession No. AY439004) were deposited in GenBank.

Patient samples, mRNA isolation, and quantitative RT-PCR. Fresh tissue samples from tumor nodules and normal adjacent prostate were collected with informed consent from patients with prostate cancer submitted to radical prostatectomy, and were stored in liquid nitrogen until use. Poly(A)⁺ messenger RNA was isolated using μ MACS mRNA isolation kit (Miltenyi Biotec, Auburn, CA). cDNA was generated using SuperscriptII reverse transcriptase (Invitrogen) and oligo(dT) primers. Each cDNA sample was quantified in duplicate reactions using the GenAmp 5700 Sequence Detector (Applied Biosystems). PCR amplifications were carried out in a volume of 20 μ l and according to manufacturer's instructions for SybrGreen PCR Core Reagent (Applied Biosystems). The fold-change of each gene was calculated by using the $2^{-\Delta CT}$ formula, where CT = fluorescence at the defined detection threshold; ΔCT = CT of the target gene – CT of the reference gene, β -tubulin. The expression level of each gene in each sample is represented as fold-change relative to the average ΔCT across all tested samples. Statistical significance for the differential expression between normal and tumor samples (*p* values) was determined using a Student's *t* test, assuming a two-tail distribution and a heteroscedastic variance model.

Northern blot analysis. The analysis of *RASL11A* expression was performed by hybridization of a nylon membrane containing 2 μ g of poly(A)⁺ isolated from several normal human tissues (Multiple Tissue Northern blot II, Clontech) with ³²P-labeled cDNA probes generated from a 3'-end RACE-PCR 710 bp fragment, following manufacturer's protocol (Rediprime II, Amersham Biosciences). Equal amounts of mRNA in each lane were confirmed by hybridization of the same blotting membrane with a ³²P-labeled α -actin cDNA. Radioactivity was detected using a Storm phosphor screen (Molecular Dynamics) and images were processed using the ImageQuant program (Molecular Dynamics).

Protein alignments and construction of phylogenetic trees. All protein alignments were performed using the ClustalX program (v 1.83) with default parameters [22]. Unrooted phylogenetic trees, with confidence of branches assigned by bootstrapping (1000 samplings), were calculated by the same software using the Neighbor Joining method excluding positions with gaps, and were drawn using TreeView program v. 1.6.6 [23]. Prior to construction of phylogenetic trees including Rasl11a or Arl9, protein sequences in the multiple alignment were manually edited to trim off unaligned residues at the amino and carboxy terminal ends to ensure that only conserved core domains were used. GenBank accession numbers to protein sequences used to build alignments and phylogenetic trees are the following: Noey2 (AAD03164.1), N-Ras (NP_002515.1), Tc21 (NP_036382.2), K-Ras (AAB41942.1), H-Ras (NP_005334.1), Rap1a (NP_002875.1), Rap1b (NP_056461.1), Rap2a (NP_066361.1), Rap2b (NP_002877.2), RalA (NP_005393.2), RalB (NP_002872.1), Rheb (NP_005605.1), Rheb11 (NP_653194.1), Rin (NP_002921.1), Rit (NP_008843.1), Rig (NP_660156.1), Cdc42 (NP_426359.1), RhoA (NP_001655.1), RhoG (NP_001656.1), Rab5c (NP_004574.1), Rab6a (NP_002860.2), Rab35 (NP_006852.1), RasD1 (NP_057168.1), Rrp22 (NP_006468.1), Ran (AAH04272.2), Ris (NP_057647.1), Rerg (NP_116307.1), Kir (AAC50067.1), Gem (NP_005252.1), Arf1 (NP_001649.1), Arf3 (NP_001650.1), Arf4 (NP_001651.1), Arf5 (NP_001653.1), Arf6 (NP_001654.1), Arf7 (AAH41803.1), Arl1 (NP_001168.1), Arl2 (NP_001658.1), Arl3 (NP_004302.1), Arl4 (NP_005729.1), Arl5 (NP_036229.1), Arl6 (NP_115522.1), Arl7 (NP_005728.2), Arl8 (NP_083742.2), Arl10c (NP_060654.1), Sar1a (NP_064535.1), and Arp1 (NP_003215.1). The following gene products were previously annotated as hypothetical proteins with unknown function and we

submitted GenBank Third-Party Annotations: Rasl11b (NP_076429.1; TPA BK001672), Arl10 (NP_775935.1; TPA BK001673), ANOGA Rasl11a (XP_317248.1; TPA BK001703), DROME Rasl11a (NP_609112.1; TPA BK001702), RAT Rasl11a (XP_221886.1; TPA BK001714), MOUSE Rasl11a (XP_132509.1; TPA BK001706), BRARE Rasl11b (AAH44471.1; TPA BK001704), XENLA Rasl11b (AAH41294.1; TPA BK001713), RAT Rasl11b (XP_223353.1; TPA BK001710), MOUSE Rasl11b (AAH08101.1; TPA BK001707), RAT Arl9 (XP_214020.1; TPA BK001708), MOUSE Arl9 (XP_144363.1; TPA BK001705), XENLA Arl10 (AAH56114.1; TPA BK001712), RAT Arl10 (XP_225175.1; TPA BK001709), and MOUSE Arl10 (NP_064352.1; TPA BK001711).

Results

Cloning of novel genes related to the Ras superfamily of small GTPases

In order to search for novel members of paralogue gene families related to the Ras superfamily, we developed a dedicated bioinformatics routine described under Materials and methods to automatically map all known members of the superfamily to the draft sequence of the human genome. Subsequently, we identified all EST partial transcripts having similarity to members of the family that mapped to novel genomic regions not yet identified as containing Ras-related genes. Essentially, protein sequences derived from 36 cancer-related representative members of this family were aligned to six-frame translations of approximately three million ESTs from the HCGP-CGAP dataset. A set of 2061 ESTs was identified as paralogue candidates using a cut-off of 50% identity over 50 amino acids. By comparing their genome mapping coordinates a smaller set of 74 candidate ESTs was identified for further manual inspection. Two partial transcript sequences, which were deemed to potentially

encode for novel Ras-related proteins, were selected for further biochemical characterization. The first candidate gene showed 89% coverage with 55% similarity at the amino acid level to Rap2a and was therefore named *RASL11A* (RAS-like, family 11, member A). The second one displayed 60% coverage with 51% amino acid similarity to Arl3 (ADP-ribosylation factor-like 3) and was therefore named *ARL9* (ADP-ribosylation factor-like 9). Unigene (UG) clusters [24] harboring partial transcript sequences of either *RASL11A* or *ARL9* (Hs.192131 and Hs.444449, respectively) were identified, and suitable cDNA sources for full-length cloning of the novel genes were selected based on the tissue composition of the respective EST libraries.

A full-length transcript from *RASL11A* was obtained by 5'- and 3'-RACE using specific internal primers based on the partial EST sequence and cDNA from prostate tissue. *RASL11A* transcript maps to chromosome 13q12.2 and is defined by four exons and three introns displaying typical splice donor and acceptor sites (Fig. 1A). The longest open-reading frame (ORF) in the cloned *RASL11A* transcript encodes for a protein with 242 aa residues that contains the conserved G domain found in all Ras-related GTPases (Fig. 2A) and extends the N-terminal end by 5 residues with respect to the in silico predicted protein derived from automatic genomic annotation (GenBank Accession No. XP_172905).

ARL9 was amplified by PCR from a testis cDNA pool with a specific primer pair derived from a virtual *ARL9* transcript sequence. This virtual *ARL9* sequence was generated by assembly of 22 spliced ESTs and an ab initio gene prediction (GenBank Accession No. XM_059617) that have overlapping genomic coordinates with the *ARL9* partial transcript identified in the informatics pipeline. The cloned fragment comprising

Fig. 1. *RASL11A* and *ARL9* map to unannotated regions on 13q12.2 and 4q12 chromosomes. Exons of *RASL11A* (A) and *ARL9* (B) are shown as filled boxes. Protein coding regions are in grey; untranslated regions (UTRs) are in black. Canonical acceptor (GT) and donor (AG) splice sites are indicated. Predicted protein lengths are indicated at the top of each panel. Exon and intron sizes are in base pairs (bp).

Fig. 2. Rasl11a and Arl9 are members of novel subfamilies of small GTPases. Phylogenetic trees were reconstructed separately for Ras/Rho/Rab protein subfamilies, including Rasl11a (A) and for Arf/Arl/Sar protein subfamilies, including Arl9 (B). The numbers represent the confidence of the branches assigned by bootstrap (in 1000 samplings). Bootstrap values lower than 500 are not shown. Conserved paralogues defining the novel Rasl11 and Arl9 subfamilies are boxed.

ARL9 coding sequence has 836 bp and maps to chromosome 4q12 (Fig. 1B). The longest ORF found in the cloned *ARL9* transcript, if an ATG codon is considered as the translation initiation site, encodes for a truncated 123-residue protein that spans only partially the conserved Arf domain. On the other hand, usage of a less frequent CTG alternative start codon extends the N-terminal extremity of the predicted protein in the same reading frame, generating a 187 aa polypeptide that covers the complete Arf/Ras domain.

Identification of human paralogues of *RASL11A* and *ARL9*

To find potential paralogues of *RASL11A* and *ARL9*, their predicted protein products were used to interrogate the NCBI human protein database. The best 10 matches of Rasl11a are representative members of Ras/Rho/Rab subfamilies (Supplementary Fig. 1A), whereas the top 10 matches of Arl9 are members of the Arf/Arl subfamilies (Supplementary Fig. 1B). Top 50 protein matches from either Rasl11a or Arl9 were manually inspected to reduce redundancy and to include more distantly related representative members of known subfamilies of the Ras superfamily that were missing. Next, these curated lists were used to generate multiple protein alignments and to reconstruct phylogenetic trees of the human Ras superfamily. For clarity and to emphasize that proteins from the Arf subfamily differ from other small GTPases by the presence of a unique structural device, the so-called “interswitch toggle” [2], phylogenetic analyses were performed separately on Ras/Rho/Rab (Fig. 2A) and Arf/Arl/Sar (Fig. 2B) family members.

It can be seen in Fig. 2A that Rasl11a was clearly placed on a separate branch of the tree, along with an unannotated hypothetical protein (GenBank Accession No. NP_076429.1) that shares 68% similarity to it, which will be referred to as Rasl11b hereafter. The Rasl11 tree branch was significantly related to a second branch containing two recently described Ras-related GTPases, Rerg [13] and Ris [25] (Fig. 2A). In Fig. 2B it is shown that Arl9 was placed on a separate branch, in the vicinity of other Arl proteins, along with a hypothetical protein (GenBank Accession No. NP_775935.1) that shows 62% similarity to it, and was therefore named Arl10.

Rasl11 proteins share most of the essential residues found in the conserved G domain of Ras-related small GTPases (Fig. 3). Noteworthy, the conserved residue Thr35 (relative to H-Ras) in the switch 1 region is replaced by Asn in Rasl11a and Rasl11b (Fig. 3). Rasl11 proteins are probably not prenylated as they lack the carboxyl terminus CaaX motif [8], but display stretches of poly-basic residues at the C-terminal (Fig. 3). The latter may warrant their interaction with membranes as previously suggested for Rin and Ris proteins [26].

In addition to the G domain of small GTPases, Arl9, as well as Arl10, show sequence conservation in the interswitch toggle, which is located between switch 1 and switch 2 and is known to interact with the N-terminal hasp in the GTP-bound form of Arf/Arl proteins [2] (Fig. 3). The few substitutions observed in the canonical interswitch toggle signature (wDvGGqxxxRxxW) [2] are conservative (Fig. 3), and occur in residues that are also substituted in Sar1a and Arp1, which have been shown to be functionally active in vivo [27,28].

Fig. 3. Alignment of the novel small GTPases to selected members of the human Ras superfamily. Selected members of the Ras/Rho/Rab and the Arf/Arl/Sar protein subfamilies were aligned along with the novel Ras111a/Ras111b and Arl9/Arl10 protein sequences. The five conserved regions that contribute to the G-domain of small GTPases are boxed. Invariant residues are shown in red. The Asn residue that replaces Thr in the switch 1 region is in bold. Residues of the interswitch toggle of the Arf family are in blue. Prenylation signal boxes (Cys-a-a-X) are shaded in grey and carboxyl-terminal poly-basic residues are in green. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this paper.)

Fig. 4. Rasl11 and ARL9/ARL10 are highly conserved in metazoans. (A) and (C) Phylogenetic tree of the Rasl11 (A) or ARL9/ARL10 (C) subfamilies, including all orthologues identified in GenBank. The numbers represent the confidence of the branches assigned by bootstrap (in 1000 samplings). (B) and (D) Alignment of Rasl11 (B) or ARL9/ARL10 (D) orthologues. Conserved regions that contribute to the G-domain of small GTPases are boxed. Invariant residues are shown in red. Residues of the interswitch toggle of the Arf family are in blue. The Asn residue that replaces Thr in the switch 1 region of Rasl11 orthologues is in bold. Carboxyl-terminal poly-basic residues are in green. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this paper.)

RASL11A and ARL9 are conserved in eukaryotes

Sequence similarity searches against GenBank NR protein database were performed aiming to identify members of orthologue families related to *RASL11* and *ARL9/ARL10* in other species. Significantly, we only identified hypothetical proteins in this step. Rasl11a/

Rasl11b and ARL9/ARL10 orthologues were aligned and phylogenetic trees representing evolutionary distances within each subfamily were constructed. As shown in Fig. 4A, human Rasl11a and Rasl11b are placed in two separate branches, each one containing closely related protein orthologues. While we only found orthologues for Rasl11a in rat and mouse, we also identified new

Fig. 4. (continued)

Rasl11b orthologues in zebra fish and *Xenopus* (Fig. 4A). In addition, we found more distantly related Rasl1 orthologues in *Drosophila* and *Anopheles* (Fig. 4A). Multiple alignment shows that Rasl1 proteins are well conserved among metazoans (Fig. 4B). All identified Rasl1 orthologues have the Asn residue in the switch 1 region (Fig. 4B) replacing the canonical Thr residue generally found in other Ras-related GTPase families (Fig. 3). We found hypothetical proteins highly similar to Arl9 and Arl10 in mouse, rat, and *Xenopus*, which further corroborates the existence of a novel, uncharacterized subfamily of proteins related to the Arf/Arl subfamily of Ras-like small GTPases (Figs. 4C and D).

Differential expression of *RASL11A* in prostate tumors

Besides the newly identified paralogue Rasl11b (Figs. 2A and 3), Rasl11a protein best match in GenBank is the product of *REG* (58% similarity, 36% identity) (Supplementary Fig. 1), a gene shown to be down-regulated in primary breast tumors with poor clinical prognosis [13]. This observation, along with the well-documented involvement of small GTPases in the processes leading to cancer progression [6–13], and the possibility to access samples from prostate tumors and normal prostate tissues, prompted us to measure the transcript levels of both *RASL11A* and *ARL9* in prostate. Quantitative RT-PCR analyses were performed using mRNAs derived from four samples of prostate tumor nodules and from five samples of normal prostate tissue. Fig. 5 shows that a significant excess ($p = 0.00065$) of *RASL11A* transcript

was detected in normal prostate as compared to prostate tumor samples (Fig. 5A). On average, *RASL11A* transcript level was 2.3-fold higher in normal prostate in relation to tumor. No difference in the expression of *ARL9* was observed in parallel experiments using the same tissue samples (Fig. 5B).

The differential expression of *RASL11A* observed between normal and neoplastic prostate led us to investigate its transcription in other human tissues by Northern blot. Two isoforms of the *RASL11A* transcript with different sizes (1.6 and 1.2 kb) were clearly detected in bladder, prostate, testis, and colon (Fig. 6). *RASL11A* expression was very weak in ovary and absent in thymus and leukocytes (Fig. 6). While the ~1.6 kb band observed in the Northern blot (Fig. 6) is compatible with the size of the cloned full-length *RASL11A* transcript (Fig. 1A), the smaller band, with approximately 1.2 kb, is likely to represent a transcript isoform that undergoes either alternative splicing and/or has an additional, internal alternative transcription initiation site.

Discussion

Functional characterization of novel paralogue families of cancer-related genes is relevant in view of the acquisition of tolerance that is frequently observed in cancer patients undergoing chemotherapy. A conceivable explanation for the acquisition of drug resistance is that compensatory alternate pathways provided by

Fig. 5. *RASL11A* is down-regulated in prostate tumors. Quantitative RT-PCR using mRNA isolated from tumor (black bars) or normal adjacent prostate samples (grey bars). Transcript levels of *RASL11A* (A) and *ARL9* (B) were normalized to β -tubulin mRNA and are represented as fold-changes relative to the average transcript level across all samples. Standard error bars from replicate experiments are shown. Student's *t* test *p* values for differential expression between normal and tumor samples were determined for *RASL11A* ($p = 0.00065$) and *ARL9* ($p = 0.52$).

Fig. 6. Two *RASL11A* isoforms are transcribed in human tissues. (A) *RASL11A* transcription in several human tissues was evaluated by Northern blot. (B) The presence of comparable amounts of mRNA in each lane was confirmed by subsequent hybridization of the membrane with a β -actin probe.

proteins with similar function to that of chemotherapeutic targets may allow cancer cells to evade and proliferate. In this work we identified and cloned two new genes *RASL11A* and *ARL9* that map to previously

unannotated human genomic regions. Open reading frames without frame-shifts or truncations, as well as conserved exon/intron structures, were identified for both *RASL11A* and *ARL9*. In spite of low similarity at the nucleotide level to known human transcripts, predicted proteins encoded by *RASL11A* and *ARL9* are very similar (68% and 62%, respectively) to the protein translations (GenBank Accession Nos. NP_076429.1 and NP_775935.1, respectively) of two complete cDNA transcripts that map on different genomic loci. These findings, supported by the evidence that both Rasl11a/Rasl11b and Arl9/Arl10 protein subfamilies are highly conserved among eukaryotes, strongly suggest that these hypothetical proteins are members of evolutionary conserved gene families.

Sequence analyses of the newly described human *ARL9* gene product showed that the G domain, including the typical Arf interswitch, is only fully represented if an alternative non-AUG initiation codon is used. Usage of in-frame CUG start codons has been

already documented for functional human genes, including proto-oncogenes, transcription factors, and growth factors [29,30]. In all these cases, non-AUG start codons appear in a similar sequence context, which is characterized by a purine at position –3 and a guanosine (or less frequently a cytosine) at position +4 [29]. These sequence features are also present in the ORF starting at the alternate codon CUG in the *ARL9* transcript. Further experimentation will aid to explain the biological significance of the use of an alternate initiation start codon in *ARL9*.

In Rasl11a, Asn59 replaces the conserved Thr residue in the switch 1 region of the G domain. Asn59 is observed in all Rasl11 orthologues and is not restricted to the Rasl11 family since Ris, a recently described family of Ras-related proteins, displays the same residue substitution [25]. In the prototype superfamily member H-Ras, residue Thr35 in switch 1, along with Gly60 in switch 2, contributes to stabilize the GTP-bound form through a hydrogen bond between its main chain NH and a γ -phosphate oxygen of GTP [5]. It is possible that the Thr → Asn substitution in Rasl11, which preserves an uncharged polar side chain, does not interfere significantly with the stabilization of GTP and ultimately with the GTPase function. Mutational studies have shown that replacement of Thr35 by a serine residue in Ras proteins reduces the intrinsic GTPase activity, but high-affinity binding of effector proteins is sustained [31]. Detailed biochemical characterization, including the identification of down-stream effectors, will be required to determine functional consequences of the presence of Asn59 in switch 1 on Rasl11 proteins.

Ras-related small GTPases generally possess sequence determinants that mediate their association to the plasma membrane and this interaction is often essential for their biological function. Rasl11 proteins lack any known recognition signal for C-terminal lipidation found in Ras proteins, such as farnesylation or palmitoylation [32,33]. On the other hand, a conserved polybasic consensus signature is found at C-terminal of all Rasl11 orthologues from vertebrates suggesting these proteins may rely, at least in part, on a membrane binding mechanism based on charged residues similar to the postulated for Rin and Rit, members of the human Ras subfamily [26,33].

Real-time quantitative PCR analyses revealed that *RASL11A* transcript levels are significantly lower in prostate tumors as compared to normal adjacent prostate tissue, suggesting a tumor suppressor role. There are several reports in the literature showing that small GTPases may act as tumor suppressors; loss of heterozygosity in the *NOEY2/AHR1* gene was detected in 41% of ovarian and breast cancers [10]. In the same work, re-expression of Noey2 through transfection suppressed the clonogenic growth of breast and ovarian cancer cells, indicating that this member of the Ras family has a

tumor suppressor activity. Over-expression of Rig inhibited Ras-mediated cellular transformation and activation of down-stream signaling in NIH 3T3 cells [14]. Also, coexpression of Rap1a was shown to revert the genomic instability induced by oncogenic H-Ras [12]. Recently, it was demonstrated that the expression of Rerg protein in breast carcinoma cells inhibited their growth in vitro and also tumor formation in nude mice [13]. We found that Rasl11a shows significant similarity to Rerg, further corroborating the hypothesis that the *RASL11A* gene might have a tumor suppressor role in prostate cancer. In view of the importance of Ras-related proteins in many aspects of the cell cycle, and their interaction with a broad range of signaling proteins, further studies to elucidate Rasl11a intracellular functional mechanisms are warranted.

Acknowledgments

This work was funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). E.M.R. is the recipient of a grant award from Coordenação de Aperfeiçoamento do Pessoal do Ensino Superior (CAPES). R.L., H.I.N., and A.C.M.P. received fellowships from FAPESP. S.V.A. and A.M.d.S. were partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

- [1] J.L. Bos, Ras-like GTPases, *Biochim. Biophys. Acta* 1333 (1997) M19–M31.
- [2] S. Pasqualato, L. Renault, J. Cherfils, Arf, Arl, Arp and Sar proteins: a family of GTP-binding proteins with a structural device for 'front-back' communication, *EMBO Rep.* 3 (2002) 1035–1041.
- [3] M. Paduch, F. Jelen, J. Otlewski, Structure of small G proteins and their regulators, *Acta Biochim. Pol.* 48 (2001) 829–850.
- [4] G. Oxford, D. Theodorescu, Ras superfamily monomeric G proteins in carcinoma cell motility, *Cancer Lett.* 189 (2003) 117–128.
- [5] I.R. Vetter, A. Wittinghofer, The guanine nucleotide-binding switch in three dimensions, *Science* 294 (2001) 1299–1304.
- [6] A.A. Adjei, Blocking oncogenic Ras signaling for cancer therapy, *J. Natl. Cancer Inst.* 93 (2001) 1062–1074.
- [7] B. Boettner, L. Van Aelst, The role of Rho GTPases in disease development, *Gene* 286 (2002) 155–174.
- [8] G.W. Reuther, C.J. Der, The Ras branch of small GTPases: Ras family members don't fall far from the tree, *Curr. Opin. Cell Biol.* 12 (2000) 157–165.
- [9] I.M. Zohn, S.L. Campbell, R. Khosravi-Far, K.L. Rossman, C.J. Der, Rho family proteins and Ras transformation: the RHOad less traveled gets congested, *Oncogene* 17 (1998) 1415–1438.
- [10] Y. Yu, F. Xu, H. Peng, X. Fang, S. Zhao, Y. Li, et al., NOEY2 (ARHI), an imprinted putative tumor suppressor gene in ovarian and breast carcinomas, *Proc. Natl. Acad. Sci. USA* 96 (1999) 214–219.
- [11] R.Z. Luo, X. Fang, R. Marquez, S.Y. Liu, G.B. Mills, W.S. Liao, Y. Yu, R.C. Bast, ARHI is a Ras-related small G-protein with a novel N-terminal extension that inhibits growth of ovarian and breast cancers, *Oncogene* 22 (2003) 2897–2909.

- [12] M.A. Wani, N.C. Denko, P.J. Stambrook, Expression of Rap 1 suppresses genomic instability of H-ras transformed mouse fibroblasts, *Somat. Cell Mol. Genet.* 23 (1997) 123–133.
- [13] B.S. Finlin, C.L. Gau, G.A. Murphy, H. Shao, T. Kimel, R.S. Seitz, et al., RERG is a novel ras-related, estrogen-regulated and growth-inhibitory gene in breast cancer, *J. Biol. Chem.* 276 (2001) 42259–42267.
- [14] C.A. Ellis, M.D. Vos, H. Howell, T. Vallecorsa, D.W. Fufts, G.J. Clark, Rig is a novel Ras-related protein and potential neural tumor suppressor, *Proc. Natl. Acad. Sci. USA* 99 (2002) 9876–9881.
- [15] A.A. Adjei, A. Mauer, L. Bruzek, R.S. Marks, S. Hillman, S. Geyer, et al., Phase II study of the farnesyl transferase inhibitor R115777 in patients with advanced non-small-cell lung cancer, *J. Clin. Oncol.* 21 (2003) 1760–1766.
- [16] A.D. Cox, C.J. Der, Ras family signaling: therapeutic targeting, *Cancer Biol. Ther.* 1 (2002) 599–606.
- [17] R.L. Strausberg, A.J. Simpson, R. Wooster, Sequence-based cancer genomics: progress, lessons and opportunities, *Nat. Rev. Genet.* 4 (2003) 409–418.
- [18] A.A. Camargo, H.P. Samaia, E. Dias-Neto, D.F. Simao, I.A. Migotto, M.R. Briones, et al., The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome, *Proc. Natl. Acad. Sci. USA* 98 (2001) 12103–12108.
- [19] R.L. Strausberg, The Cancer Genome Anatomy Project: new resources for reading the molecular signatures of cancer, *J. Pathol.* 195 (2001) 31–40.
- [20] H. Brentani, O.L. Caballero, A.A. Camargo, A.M. da Silva, W.A. da Silva Jr., E. Dias Neto, et al., The generation and utilization of a cancer-oriented representation of the human transcriptome by using expressed sequence tags, *Proc. Natl. Acad. Sci. USA* 100 (2003) 13418–13423.
- [21] A.C.M. Paquola, A.A. Machado, E.M. Reis, A.M. da Silva, S. Verjovski-Almeida, Zerg: a very fast BLAST parser library, *Bioinformatics* 19 (2003) 1035–1036.
- [22] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, *Nucleic Acids Res.* 25 (1997) 4876–4882.
- [23] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, *Comput. Appl. Biosci.* 12 (1996) 357–358.
- [24] D.L. Wheeler, D.M. Church, S. Federhen, A.E. Lash, T.L. Madden, J.U. Pontius, et al., Database resources of the national center for biotechnology, *Nucleic Acids Res.* 31 (2003) 28–33.
- [25] G. Wistow, S.L. Bernstein, S. Ray, M.K. Wyatt, A. Behal, J.W. Touchman, G. Bouffard, D. Smith, K. Peterson, Expressed sequence tag analysis of adult human iris for the NEIBank Project: steroid-response factors and similarities with retinal pigment epithelium, *Mol. Vis.* 8 (2002) 185–195.
- [26] C.H. Lee, N.G. Della, C.E. Chew, D.J. Zack, Rin, a neuron-specific and calmodulin-binding small G-protein, and Rit define a novel subfamily of ras proteins, *J. Neurosci.* 16 (1996) 6784–6794.
- [27] O. Kuge, C. Dascher, L. Orci, T. Rowe, M. Amherdt, H. Plutner, et al., Sar1 promotes vesicle budding from the endoplasmic reticulum but not Golgi compartments, *J. Cell Biol.* 125 (1994) 51–65.
- [28] D.M. Eckley, T.A. Schroer, Interactions between the evolutionarily conserved, actin-related protein, Arp11, actin, and Arp1, *Mol. Biol. Cell* 14 (2003) 2645–2654.
- [29] R. Boeck, D. Kolakofsky, Positions +5 and +6 can be major determinants of the efficiency of non-AUG initiation codons for protein synthesis, *EMBO J.* 13 (1994) 3017–3608.
- [30] M.K. Tee, R.B. Jaffe, A precursor form of vascular endothelial growth factor arises by initiation from an upstream in-frame CUG codon, *Biochem. J.* 359 (2001) 219–226.
- [31] M. Spoerner, C. Herrmann, I.R. Vetter, H.R. Kalbitzer, A. Wittinghofer, Dynamic properties of the Ras switch I region and its importance for binding to effectors, *Proc. Natl. Acad. Sci. USA* 98 (2001) 4944–4949.
- [32] Q. Dai, E. Choy, V. Chiu, J. Romano, S.R. Slivka, S.A. Steitz, S. Michaelis, M.R. Philips, Mammalian prenylcysteine carboxyl methyltransferase is in the endoplasmic reticulum, *J. Biol. Chem.* 273 (1998) 15030–15034.
- [33] M. Sinensky, Recent advances in the study of prenylated proteins, *Biochim. Biophys. Acta* 1484 (2000) 93–106.

Tabela Suplementar 1 - Transcritos exônicos regulados por andrógeno.

Transcritos com alterações significativas na expressão em pelo menos 3 tempos consecutivos indentificado por SAM, usando como parâmetros: respostas de duas classes (dados pareados), 1000 permutações *K-Nearest Neighbors Imputer*, *fold-change* $\geq 1,5$ e taxa de falsa descoberta (FDR) $< 5\%$. Coluna de efeito do andrógeno é baseada no *fold change* ($\log_2$) observado após 24 horas de tratamento.

N.D. - Não disponível

Spot ID	GenBank Accession (EST depositada)	Mapeamento	Nome do Locus	Anotação do Locus	GenBank Accession (da anotação)	Efeito do andrógeno		Clustering Hierárquico (ordem)
3299	BG876589	Exonic	UGT2B17	Homo sapiens UDP glycosyltransferase 2 family, polypeptide B17 (UGT2B17), mRNA	NM_001077	Downregulated	-2.54	1
3636	BF989967	Exonic	ARHGAP29	Homo sapiens Rho GTPase activating protein 29 (ARHGAP29), mRNA	NM_004815	Downregulated	-0.95	2
1225	BF805750	Exonic	COLEC12	Homo sapiens collectin sub-family member 12 (COLEC12), transcript variant I, mRNA	NM_130386	Downregulated	-0.92	3
1674	BF090400	Exonic	N.A.	Homo sapiens mRNA; cDNA DKFZp686K0548 (from clone DKFZp686K0548)	BX649128	Downregulated	-1.10	4
3668	AW995234	Exonic	C6orf55	Homo sapiens chromosome 6 open reading frame 55 (C6orf55), mRNA	NM_016485	Downregulated	-0.86	5
3321	BQ377372	Exonic	CPNE3	Homo sapiens copine III (CPNE3), mRNA	NM_003909	Downregulated	-0.70	6
3276	BQ366895	Exonic	SLAC2-B	Homo sapiens SLAC2-B (SLAC2-B), mRNA	NM_015065	Downregulated	-1.26	7
552	BF817791	Exonic	ADD3	Homo sapiens adducin 3 (gamma) (ADD3), transcript variant 1, mRNA	NM_016824	Downregulated	-1.70	8
565	BF945476	Exonic	MXI1	Homo sapiens MAX interactor 1 (MXI1), transcript variant 1, mRNA	NM_005962	Downregulated	-1.34	9
396	AW375604	Exonic	N.A.	Homo sapiens hypothetical protein (LOC387882), mRNA	NM_207376	Downregulated	-0.36	10
427	AW901818	Exonic	TMEM65	Homo sapiens transmembrane protein 65 (TMEM65), mRNA	NM_194291	Downregulated	-0.06	11
577	BF155848	Exonic	SYNE2	Homo sapiens spectrin repeat containing, nuclear envelope 2 (SYNE2), transcript variant 2, mRNA	NM_182910	Downregulated	-0.65	12
49	BE840007	Exonic	AMOTL1	Homo sapiens angiomin like 1 (AMOTL1), mRNA	NM_130847	Downregulated	-0.99	13
245	BE165360	Exonic	ROCK1	Homo sapiens Rho-associated, coiled-coil containing protein kinase 1 (ROCK1), mRNA	NM_005406	Downregulated	-1.18	14
297	BF805983	Exonic	ADK	Homo sapiens adenosine kinase (ADK), transcript variant ADK-short, mRNA	NM_001123	Downregulated	-0.93	15
518	BE162922	Exonic	N.A.	Homo sapiens cDNA FLJ37408 fis, clone BRAMY2028491	AK094727	Downregulated	-0.99	16
642	BF883122	Exonic	CD63	Homo sapiens CD63 antigen (melanoma 1 antigen) (CD63), mRNA	NM_001780	Downregulated	-1.17	17
239	BE926612	Exonic	MNAT1	Homo sapiens menage a trois 1 (CAK assembly factor) (MNAT1), mRNA	NM_002431	Downregulated	-0.81	18
612	AW370601	Exonic	CNOT8	Homo sapiens CCR4-NOT transcription complex, subunit 8 (CNOT8), mRNA	NM_004779	Downregulated	-0.71	19
589	AW860776	Exonic	TERF1	Homo sapiens telomeric repeat binding factor (NIMA-interacting) 1 (TERF1), transcript variant 2, mRNA	NM_003218	Downregulated	-0.79	20
645	BF328037	Exonic	UBAP2L	Homo sapiens NICE-4 protein (NICE-4), mRNA	NM_014847	Downregulated	-0.92	21
1610	BG009879	Exonic	N.A.	Homo sapiens mRNA; cDNA DKFZp779K2051 (from clone DKFZp779K2051)	BX537506	Downregulated	-0.79	22

695	BE768777	Exonic	ST7	Homo sapiens suppression of tumorigenicity 7 (ST7), transcript variant a, mRNA	NM_018412	Downregulated	-1.40	23
2546	BF087381	Exonic	TP53	Homo sapiens tumor protein p53 (Li-Fraumeni syndrome) (TP53), mRNA	NM_000546	Downregulated	-0.92	24
1346	AW880649	Exonic	TMSB4X	Homo sapiens thymosin-like 3 (TMSL3), mRNA	NM_183049	Downregulated	-0.77	25
1410	BG012653	Exonic	RALB	Homo sapiens v-ral simian leukemia viral oncogene homolog B (ras related; GTP binding protein) (RALB), mRNA	NM_002881	Downregulated	-1.13	26
2569	AW937823	Exonic	MYC	Homo sapiens v-myc myelocytomatosis viral oncogene homolog (avian) (MYC), mRNA	NM_002467	Downregulated	-0.95	27
1031	BE074224	Exonic	CD151	Homo sapiens CD151 antigen (CD151), transcript variant 2, mRNA	NM_139030	Downregulated	-0.95	28
613	BF951773	Exonic	RAN	Homo sapiens RAN, member RAS oncogene family (RAN), mRNA	NM_006325	Downregulated	-0.80	29
791	BE815309	Exonic	ARRDC4	Homo sapiens arrestin domain containing 4 (ARRDC4), mRNA	NM_183376	Downregulated	-0.88	30
392	BF809098	Exonic	ZFP36L2	Homo sapiens zinc finger protein 36, C3H type-like 2 (ZFP36L2), mRNA	NM_006887	Downregulated	-0.82	31
1049	BE168033	Exonic	NFKB1	Homo sapiens nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105) (NFKB1), mRNA	NM_003998	Downregulated	-0.73	32
479	AW859979	Exonic	N.A.	Homo sapiens cDNA FLJ38048 fis, clone CTONG2014264	AK095367	Downregulated	-0.97	33
475	BF761020	Exonic	PALM2-AKAP2	Homo sapiens PALM2-AKAP2 protein (PALM2-AKAP2), transcript variant 1, mRNA	NM_007203	Downregulated	-0.77	34
214	BF754489	Exonic	FAU	Homo sapiens Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitously expressed (fox derived); ribosomal protein S30 (FAU), mRNA	NM_001997	Downregulated	-1.12	35
231	BE165509	Exonic	DEFB1	Homo sapiens defensin, beta 1 (DEFB1), mRNA.	NM_005218	Downregulated	-1.08	36
3	AW994959	Exonic	ROD1	Homo sapiens ROD1 regulator of differentiation 1 (S. pombe) (ROD1), mRNA	NM_005156	Upregulated	0.08	37
273	AW373943	Exonic	CAPNS1	Homo sapiens calpain, small subunit 1 (CAPNS1), mRNA	NM_001749	Downregulated	-1.24	38
470	BF845798	Exonic	USP54	Homo sapiens ubiquitin specific peptidase 54, mRNA (cDNA clone IMAGE:6503621), with apparent retained intron	BC110845	Downregulated	-1.34	39
168	AW842748	Exonic	SIN3B	PREDICTED: Homo sapiens SIN3 homolog B, transcriptional regulator (yeast) (SIN3B), mRNA	XM_050561	Downregulated	-1.40	40
674	BF890754	Exonic	UQCRH	Homo sapiens ubiquinol-cytochrome c reductase hinge protein (UQCRH), mRNA	NM_006004	Downregulated	-1.56	41
682	AW603395	Exonic	YEATS4	Homo sapiens glioma-amplified sequence-41 (GAS41), mRNA	NM_006530	Downregulated	-1.55	42
873	CK327135	Exonic	N.A.	Homo sapiens cDNA FLJ41973 fis, clone SKNMC2004651	AK123967	Downregulated	-1.41	43
621	BF898656	Exonic	TRIP3	Homo sapiens thyroid hormone receptor interactor 3 (TRIP3), mRNA	NM_004773	Downregulated	-1.36	44
644	BQ314686	Exonic	MSH3	Homo sapiens mutS homolog 3 (E. coli) (MSH3), mRNA	NM_002439	Downregulated	-1.35	45
692	BF751539	Exonic	FOXK2	Homo sapiens forkhead box K2 (FOXK2), transcript variant 3, mRNA	NM_181431	Downregulated	-1.06	46
1002	BQ334482	Exonic	CDA	Homo sapiens cytidine deaminase (CDA), mRNA	NM_001785	Downregulated	-1.05	47
941	BE172353	Exonic	FUBP1	Homo sapiens far upstream element (FUSE) binding protein 1 (FUBP1), mRNA	NM_003902	Downregulated	-1.25	48
910	BE833259	Exonic	STN2	Homo sapiens mRNA for Stonin 2 variant protein	AB208948	Downregulated	-1.11	49
622	BE708661	Exonic	RXRG	Homo sapiens retinoid X receptor, gamma (RXRG), mRNA	NM_006917	Downregulated	-0.76	50
1809	BQ302418	Exonic	RAC1	Homo sapiens ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1) (RAC1), transcript variant Rac1, mRNA	NM_006908	Downregulated	-0.76	51
1302	BE163603	Exonic	TMEM141	Homo sapiens transmembrane protein 141 (TMEM141), mRNA	NM_032928	Downregulated	-0.60	52
2699	BF330679	Exonic	TMC4	Homo sapiens transmembrane channel-like 4 (TMC4), mRNA	NM_144686	Downregulated	-0.89	53

1460	BF856925	Exonic	WDR23	Homo sapiens WD repeat domain 23 (WDR23), transcript variant 2, mRNA	NM_181357	Downregulated	-0.78	54
1388	AW883439	Exonic	PRKAG1	Homo sapiens protein kinase, AMP-activated, gamma 1 non-catalytic subunit (PRKAG1), transcript variant 2, mRNA	NM_212461	Downregulated	-0.98	55
1402	AW883175	Exonic	CASP7	Homo sapiens caspase 7, apoptosis-related cysteine protease (CASP7), transcript variant delta, mRNA	NM_033338	Downregulated	-0.85	56
1469	AW365027	Exonic	CSNK1G2	Homo sapiens casein kinase 1, gamma 2 (CSNK1G2), mRNA	NM_001319	Downregulated	-0.91	57
1326	BF922962	Exonic	CTSB	Homo sapiens cathepsin B (CTSB), transcript variant 1, mRNA	NM_001908	Downregulated	-0.93	58
1458	AW880788	Exonic	LIMK1	Homo sapiens LIM domain kinase 1 (LIMK1), transcript variant dLIMK, mRNA	NM_016735	Downregulated	-0.90	59
1784	BF892704	Exonic	GSN	Homo sapiens gelsolin (amyloidosis, Finnish type) (GSN), transcript variant 2, mRNA	NM_198252	Downregulated	-0.61	60
704	BF089987	Exonic	NAG8	Homo sapiens nasopharyngeal carcinoma associated gene protein-8 (NAG8), mRNA	NM_014411	Downregulated	-0.90	61
2240	BG001664	Exonic	CD86	Homo sapiens CD86 antigen (CD28 antigen ligand 2, B7-2 antigen) (CD86), transcript variant 2, mRNA	NM_006889	Downregulated	-0.60	62
3327	BF082343	Exonic	LGALS8	Homo sapiens lectin, galactoside-binding, soluble, 8 (galectin 8) (LGALS8), transcript variant 3, mRNA	NM_201544	Downregulated	-0.56	63
2468	BF930663	Exonic	S100P	Homo sapiens S100 calcium binding protein P (S100P), mRNA	NM_005980	Upregulated	0.26	64
2003	BE087107	Exonic	N.A.	Homo sapiens cDNA: FLJ21288 fis, clone COL01927	AK024941	Upregulated	0.66	65
1770	BF922194	Exonic	N.A.	Homo sapiens KIAA0657 protein (KIAA0657), mRNA	XM_051017	Upregulated	0.55	66
1627	BE080712	Exonic	ZCCHC6	Homo sapiens zinc finger, CCHC domain containing 6 (ZCCHC6), mRNA	NM_024617	Upregulated	0.85	67
2226	AW852323	Exonic	N.A.	Homo sapiens cDNA FLJ45532 fis, clone BRTHA2030036	AK127440	Upregulated	0.64	68
1752	BF918989	Exonic	DNAJC10	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 10 (DNAJC10), mRNA	NM_018981	Upregulated	0.72	69
1703	AW890141	Exonic	ITGAV	Homo sapiens integrin, alpha V (vitronectin receptor, alpha polypeptide, antigen CD51) (ITGAV), mRNA	NM_002210	Upregulated	0.81	70
3349	AW749339	Exonic	DACT1	Homo sapiens dapper homolog 1, antagonist of beta-catenin (xenopus) (DACT1), mRNA	NM_016651	Upregulated	0.90	71
3646	BF921541	Exonic	TERF2IP	Homo sapiens telomeric repeat binding factor 2, interacting protein (TERF2IP), mRNA	NM_018975	Upregulated	0.95	72
4021	BE927584	Exonic	GSPT1	Homo sapiens G1 to S phase transition 1 (GSPT1), mRNA	NM_002094	Upregulated	0.76	73
4462	BE711811	Exonic	TIAM1	Homo sapiens T-cell lymphoma invasion and metastasis 1 (TIAM1), mRNA	NM_003253	Upregulated	1.05	74
3635	BE708528	Exonic	PACSIN2	Homo sapiens protein kinase C and casein kinase substrate in neurons 2 (PACSIN2), mRNA	NM_007229	Upregulated	0.73	75
4143	BQ302049	Exonic	TRA1	Homo sapiens tumor rejection antigen (gp96) 1 (TRA1), mRNA	NM_003299	Upregulated	0.80	76
3351	BE163747	Exonic	KDEL2	Homo sapiens KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2 (KDEL2), mRNA	NM_006854	Upregulated	1.15	77
3323	BE827259	Exonic	IDI1	Homo sapiens isopentenyl-diphosphate delta isomerase (IDI1), mRNA	NM_004508	Upregulated	0.83	78
3593	BF857286	Exonic	PICALM	Homo sapiens phosphatidylinositol binding clathrin assembly protein (PICALM), transcript variant 2, mRNA	NM_001008660	Upregulated	1.08	79
3525	BE087783	Exonic	HLCDGP1	Homo sapiens HLCDGP1 mRNA, complete cds	AF447582	Upregulated	1.10	80

1405	BG009544	Exonic	EGFR	Homo sapiens epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian) (EGFR), transcript variant 1, mRNA	NM_005228	Upregulated	0.45	81
1799	BE074283	Exonic	PPM1D	Homo sapiens protein phosphatase 1D magnesium-dependent, delta isoform (PPM1D), mRNA	NM_003620	Upregulated	0.88	82
2202	BE831984	Exonic	SP140	Homo sapiens SP140 nuclear body protein (SP140), mRNA	NM_007237	Upregulated	0.91	83
2227	BQ292121	Exonic	RAB3B	Homo sapiens RAB3B, member RAS oncogene family (RAB3B), mRNA	NM_002867	Upregulated	0.80	84
4525	BE818647	Exonic	EFNA5	Homo sapiens ephrin-A5 (EFNA5), mRNA	NM_001962	Upregulated	0.81	85
2852	BF361046	Exonic	PECI	Homo sapiens peroxisomal D3,D2-enoyl-CoA isomerase (PECI), transcript variant 1, mRNA	NM_006117	Upregulated	1.09	86
2900	AW995233	Exonic	BMPR1A	Homo sapiens bone morphogenetic protein receptor, type IA (BMPR1A), mRNA	NM_004329	Upregulated	1.15	87
3741	BF736661	Exonic	FAT	Homo sapiens FAT tumor suppressor homolog 1 (Drosophila) (FAT), mRNA	NM_005245	Upregulated	1.34	88
594	AW994870	Exonic	PPP1CB	Homo sapiens protein phosphatase 1, catalytic subunit, beta isoform (PPP1CB), transcript variant 1, mRNA	NM_002709	Upregulated	0.33	89
287	BE833289	Exonic	VCL	Homo sapiens vinculin (VCL), transcript variant VCL, mRNA	NM_003373	Upregulated	0.30	90
954	BF757813	Exonic	PBEF1	Homo sapiens pre-B-cell colony enhancing factor 1 (PBEF1)	NM_005746	Upregulated	0.44	91
1314	BE703218	Exonic	MYOM2	Homo sapiens myomesin (M-protein) 2, 165kDa (MYOM2), mRNA	NM_003970	Upregulated	0.29	92
4484	AW604799	Exonic	S100A11	Homo sapiens S100 calcium binding protein A11 (calgizzarin) (S100A11), mRNA	NM_005620	Upregulated	0.77	93
2099	BE708316	Exonic	WIPI49	Homo sapiens WD40 repeat protein Interacting with phospholipids of 49kDa (WIPI49), mRNA	NM_017983	Upregulated	1.09	94
1566	BF848731	Exonic	MRPL14	Homo sapiens mitochondrial ribosomal protein L14 (MRPL14), nuclear gene encoding mitochondrial protein, mRNA	NM_032111	Upregulated	1.11	95
3137	BE939455	Exonic	ANKRD37	Homo sapiens low density lipoprotein receptor-related protein binding protein (Lrp2bp), mRNA	NM_181726	Upregulated	1.30	96
3755	BE768971	Exonic	MTAP	Homo sapiens methylthioadenosine phosphorylase (MTAP)	NM_002451	Upregulated	1.00	97
3703	BQ334342	Exonic	LCP1	Homo sapiens lymphocyte cytosolic protein 1 (L-plastin) (LCP1), mRNA	NM_002298	Upregulated	1.48	98
3986	BF893393	Exonic	RNF149	Homo sapiens LOC389011 (LOC389011), mRNA.	XM_373998	Upregulated	0.90	99
3162	AW606115	Exonic	N.A.	Homo sapiens cDNA clone IMAGE:4752183	BC106022	Upregulated	0.69	100
3717	BE932430	Exonic	FOSL2	Homo sapiens FOS-like antigen 2 (FOSL2), mRNA	NM_005253	Upregulated	1.03	101
393	AW747998	Exonic	ZFYVE19	Homo sapiens zinc finger, FYVE domain containing 19 (ZFYVE19), mRNA	NM_032850	Upregulated	0.72	102
4338	AW805672	Exonic	N.A.	Homo sapiens cDNA FLJ38048 fis, clone CTONG2014264.	AK095367	Upregulated	0.61	103
4242	BF745581	Exonic	UNQ473	Homo sapiens DMC (UNQ473), mRNA	NM_198477	Upregulated	1.36	104
4239	BF332449	Exonic	MGEA5	Homo sapiens meningioma expressed antigen 5 (hyaluronidase) (MGEA5), mRNA	NM_012215	Upregulated	1.20	105
2897	BF743126	Exonic	CLK2	Homo sapiens CDC-like kinase 2 (CLK2), transcript variant 2, mRNA	NM_001291	Upregulated	0.99	106
3586	BF997724	Exonic	PAPPA	Homo sapiens pregnancy-associated plasma protein A, pappalysin 1 (PAPPA), mRNA.	NM_002581	Upregulated	0.63	107
1352	BE089538	Exonic	N.A.	Homo sapiens hypothetical protein MGC33365 (MGC33365), mRNA	NM_173552	Upregulated	0.85	108
2870	BF920836	Exonic	TFE3	Homo sapiens mRNA; cDNA DKFZp761J1810 (from clone DKFZp761J1810)	AL161985	Upregulated	0.96	109

3581	BF373987	Exonic	PTBP1	Homo sapiens polypyrimidine tract binding protein 1 (PTBP1), transcript variant 4, mRNA	NM_175847	Upregulated	1.11	110
3858	AW862460	Exonic	A2ML1	Homo sapiens alpha-2-macroglobulin-like 1 (A2ML1), mRNA	NM_144670	Upregulated	1.13	111
2605	BG002636	Exonic	HOXB6	Homo sapiens homeo box B6 (HOXB6), transcript variant 2, mRNA	NM_018952	Upregulated	1.24	112
4155	BG999796	Exonic	FES	Homo sapiens feline sarcoma oncogene (FES), mRNA	NM_002005	Upregulated	1.00	113
4237	BE089334	Exonic	N.A.	Homo sapiens mRNA; cDNA DKFZp762K067 (from clone DKFZp762K067)	CR627373	Upregulated	1.06	114
3674	BE694434	Exonic	AKAP11	Homo sapiens A kinase (PRKA) anchor protein 11 (AKAP11), transcript variant 1, mRNA	NM_016248	Upregulated	1.08	115
3705	BF895259	Exonic	CARD8	Homo sapiens caspase recruitment domain family, member 8 (CARD8), mRNA	NM_014959	Upregulated	1.14	116
3003	AW391808	Exonic	SIL	Homo sapiens TAL1 (SCL) interrupting locus (SIL), mRNA	NM_003035	Upregulated	1.30	117
4147	BE148898	Exonic	FANCE	Homo sapiens Fanconi anemia, complementation group E (FANCE), mRNA	NM_021922	Upregulated	1.56	118
3673	AW606595	Exonic	BTG1	Homo sapiens B-cell translocation gene 1, anti-proliferative (BTG1), mRNA	NM_001731	Upregulated	0.55	119
4099	BQ314772	Exonic	RAB1A	Homo sapiens RAB1A, member RAS oncogene family (RAB1A), mRNA	NM_004161	Upregulated	0.98	120
3339	BQ310838	Exonic	TRA1	Homo sapiens tumor rejection antigen (gp96) 1 (TRA1), mRNA	NM_003299	Upregulated	0.81	121
4358	BE163446	Exonic	SLC25A23	Homo sapiens solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 23 (SLC25A23), mRNA	NM_024103	Upregulated	0.49	122
1924	BF742190	Exonic	N.A.	Homo sapiens cDNA FLJ14201 fis, clone NT2RP3002955	AK024263	Upregulated	0.49	123
2764	BF881091	Exonic	HDAC4	Homo sapiens histone deacetylase 4 (HDAC4), mRNA	NM_006037	Upregulated	0.54	124
2935	BG000772	Exonic	CD14	Homo sapiens CD14 antigen (CD14), mRNA	NM_000591	Upregulated	0.72	125
1833	BF919972	Exonic	N.A.	Homo sapiens neurotrophic tyrosine kinase, receptor, type 3 (NTRK3), mRNA	NM_002530	Upregulated	0.85	126
2007	BF805191	Exonic	NUDT16L1	Homo sapiens hypothetical protein MGC11275 (SDOS), mRNA	NM_032349	Upregulated	0.99	127
2956	BF371928	Exonic	TMEFF2	Homo sapiens transmembrane protein with EGF-like and two follistatin-like domains 2 (TMEFF2), mRNA	NM_016192	Upregulated	0.52	128
4402	BF853544	Exonic	S100A13	Homo sapiens S100 calcium binding protein A13 (S100A13), mRNA	XM_371380	Upregulated	0.79	129
2987	BF332165	Exonic	ALDH1A2	Homo sapiens aldehyde dehydrogenase 1 family, member A2 (ALDH1A2), transcript variant 2, mRNA	NM_170696	Upregulated	1.16	130
4513	BQ339716	Exonic	ERBB4	Homo sapiens v-erb-a erythroblastic leukemia viral oncogene homolog 4 (avian) (ERBB4), mRNA	NM_005235	Upregulated	0.89	131
983	AW352153	Exonic	PAWR	Homo sapiens PRKC, apoptosis, WT1, regulator (PAWR), mRNA	NM_002583	Upregulated	1.24	132
3908	BF996163	Exonic	N.A.	Homo sapiens hypothetical protein FLJ14490 (FLJ14490), mRNA	NM_032793	Upregulated	1.61	133
4495	BF926477	Exonic	IHPK2	Homo sapiens inositol hexaphosphate kinase 2 (IHPK2), mRNA	NM_016291	Upregulated	1.25	134
1835	CK326987	Exonic	ABCF1	Homo sapiens ATP-binding cassette, sub-family F (GCN20), member 1 (ABCF1), mRNA	NM_001090	Upregulated	0.99	135
1284	BG012009	Exonic	PHLDB2	Homo sapiens pleckstrin homology-like domain, family B, member 2 (PHLDB2), mRNA	NM_145753	Upregulated	1.13	136
3165	BF331880	Exonic	ARPC2	Homo sapiens actin related protein 2/3 complex, subunit 2, 34kDa (ARPC2), transcript variant 2, mRNA	NM_005731	Upregulated	1.51	137
2580	BF956989	Exonic	PALM2-AKAP2	Homo sapiens PALM2-AKAP2 protein (PALM2-AKAP2), transcript variant 1, mRNA	NM_007203	Upregulated	1.27	138

2098	CK326979	Exonic	RFC4	Homo sapiens replication factor C (activator 1) 4, 37kDa (RFC4), transcript variant 2, mRNA	NM_181573	Upregulated	1.64	139
4479	BF851788	Exonic	MTA1	Homo sapiens metastasis associated 1 (MTA1), mRNA	NM_004689	Upregulated	1.94	140
4009	BE815077	Exonic	CPA4	Homo sapiens carboxypeptidase A4 (CPA4), mRNA	NM_016352	Upregulated	1.77	141
3750	BE699157	Exonic	N.A.	Homo sapiens hypothetical protein DKFZp761G2113 (DKFZp761G2113), mRNA	XM_375456	Upregulated	1.72	142
2597	BF372924	Exonic	BITE	Homo sapiens p10-binding protein (BITE), mRNA	NM_024491	Upregulated	1.57	143
3117	BE156258	Exonic	EEF1A1	Homo sapiens eukaryotic translation elongation factor 1 alpha 1 (EEF1A1), mRNA	NM_001402	Upregulated	1.86	144
2960	BF365078	Exonic	ANXA6	Homo sapiens annexin A6 (ANXA6), transcript variant 2, mRNA	NM_004033	Upregulated	1.71	145
4435	AW369683	Exonic	ARHGAP26	Homo sapiens Rho GTPase activating protein 26 (ARHGAP26), mRNA	NM_015071	Upregulated	2.11	146
252	BF747659	Exonic	TCRG	Human T-cell receptor gamma chain VJCI-3(CII)-CIII region mRNA, 3' end.	M16804	Upregulated	1.39	147
928	BE841633	Exonic	OACT2	Homo sapiens O-acyltransferase (membrane bound) domain containing 2 (OACT2), mRNA	NM_138799	Upregulated	1.00	148
1028	BG002368	Exonic	LIFR	Homo sapiens leukemia inhibitory factor receptor (LIFR), mRNA	NM_002310	Upregulated	1.28	149
3773	BG008641	Exonic	HPGD	Homo sapiens hydroxyprostaglandin dehydrogenase 15-(NAD) (HPGD), mRNA	NM_000860	Upregulated	1.90	150
4431	BE166669	Exonic	B2M	Homo sapiens beta-2-microglobulin (B2M), mRNA	NM_004048	Upregulated	2.00	151
1558	AW860935	Exonic	LONRF1	Homo sapiens LON peptidase N-terminal domain and ring finger 1 (LONRF1), mRNA	NM_152271	Upregulated	1.71	152
2136	AW948647	Exonic	SOCS2	Homo sapiens suppressor of cytokine signaling 2 (SOCS2), mRNA	NM_003877	Upregulated	1.89	153
3730	AW373660	Exonic	SEPP1	Homo sapiens selenoprotein P, plasma, 1 (SEPP1), mRNA	NM_005410	Upregulated	1.94	154
2711	AW369662	Exonic	TMEPAI	Homo sapiens transmembrane, prostate androgen induced RNA (TMEPAI), transcript variant 4, mRNA	NM_199171	Upregulated	1.58	155
1718	BQ365807	Exonic	ELL2	Homo sapiens elongation factor, RNA polymerase II, 2 (ELL2), mRNA	NM_012081	Upregulated	1.68	156
2388	BF996314	Exonic	TMEPAI	Homo sapiens transmembrane, prostate androgen induced RNA (TMEPAI), transcript variant 4, mRNA	NM_199171	Upregulated	1.48	157
4497	BE769452	Exonic	ABCC4	Homo sapiens ATP-binding cassette, sub-family C (CFTR/MRP), member 4 (ABCC4), mRNA	NM_005845	Upregulated	1.68	158
4469	BF856737	Exonic	NKX3NO1	Homo sapiens NK3 transcription factor related, locus 1 (Drosophila) (NKX3-1), mRNA	NM_006167	Upregulated	1.94	159
2871	BE146184	Exonic	NDRG1	Homo sapiens N-myc downstream regulated gene 1 (NDRG1), mRNA.	NM_006096	Upregulated	2.45	160
2155	BF894880	Exonic	N.A.	Homo sapiens KIAA0056 protein (KIAA0056), mRNA	NM_015261	Upregulated	2.33	161
655	BF736834	Exonic	TMPRSS2	Homo sapiens transmembrane protease, serine 2 (TMPRSS2), mRNA	NM_005656	Upregulated	1.69	162
4403	BE814543	Exonic	EMP1	Homo sapiens epithelial membrane protein 1 (EMP1), mRNA	NM_001423	Upregulated	1.00	163
4262	BE702744	Exonic	N.A.	Homo sapiens hypothetical protein MGC27345, mRNA (cDNA clone MGC:27345 IMAGE:4670552), complete cds.	BC024231	Upregulated	2.33	164
2194	BQ367470	Exonic	CSF3R	Homo sapiens colony stimulating factor 3 receptor (granulocyte) (CSF3R), transcript variant 2, mRNA	NM_156038	Upregulated	2.65	165
2814	AW996919	Exonic	N.A.	Homo sapiens mRNA; cDNA DKFZp686L193 (from clone DKFZp686L193)	BX648778	Upregulated	2.47	166
1839	CK326989	Exonic	ZNF273	Homo sapiens zinc finger protein 273 (ZNF273), mRNA	NM_021148	Upregulated	2.39	167
2215	BE160926	Exonic	BFAR	Homo sapiens bifunctional apoptosis regulator (BFAR), mRNA	NM_016561	Upregulated	3.56	168

Tabela Suplementar 2 - Transcritos intrônicos regulados por andrógeno.

Transcritos com alterações significativas na expressão em pelo menos 3 tempos consecutivos indentificado por SAM, usando como parâmetros: respostas de duas classes (dados pareados), 1000 permutações *K-Nearest Neighbors Imputer*, *fold-change* $\geq 1,5$ e taxa de falsa descoberta (FDR) $< 5\%$. Coluna de efeito do andrógeno é baseada no *fold change* (Log_2) observado após 24 horas de tratamento.

N.D. - Não disponível

Spot ID	GenBank Accession (EST depositada)	Mapeamento	Nome do Locus	Anotação do Locus	GenBank Accession (da anotação)	Efeito do andrógeno	Clustering Hierárquico (ordem)
517	BF879099	Intronic	GAS6	Homo sapiens growth arrest-specific 6 (GAS6), mRNA	NM_000820	Downregulated -1.70	1
156	BF360792	Intronic	RAB25	Homo sapiens RAB25, member RAS oncogene family (RAB25), mRNA	NM_020387	Downregulated -1.27	2
70	CK327123	Intronic	SPIRE2	Homo sapiens spire homolog 2 (Drosophila) (SPIRE2), mRNA	XM_047462	Downregulated -1.06	3
482	AW805354	Intronic	N.D.	Homo sapiens cDNA FLJ45018 fis, clone BRAWH3015610	AK126965	Downregulated -0.50	4
559	BF854186	Intronic	PHC2	Homo sapiens partial mRNA for polyhomeotic 2 protein (PH2 gene)	AJ242730	Downregulated -0.65	5
402	BF333281	Intronic	NIBP	Homo sapiens NIK and IKK{beta} binding protein (NIBP), mRNA	NM_031466	Downregulated -1.24	6
790	AW860958	Intronic	DST	Homo sapiens dystonin (DST), transcript variant 1eA, mRNA	NM_015548	Downregulated -1.03	7
1294	BF768459	Intronic	ACTN4	Homo sapiens actinin, alpha 4 (ACTN4), mRNA	NM_004924	Downregulated -1.07	8
1293	BF368584	Intronic	DNAJC3	Homo sapiens DnaJ (Hsp40) homolog, subfamily C, member 3 (DNAJC3), mRNA	NM_006260	Downregulated -1.13	9
1301	BF882783	Intronic	ACTN4	Homo sapiens actinin, alpha 4 (ACTN4), mRNA	NM_004924	Downregulated -1.15	10
1592	AW815632	Intronic	CFLAR	Homo sapiens CASP8 and FADD-like apoptosis regulator (CFLAR), mRNA	NM_003879	Upregulated 0.48	11
105	CK327134	Intronic	RSNL2	Homo sapiens hypothetical protein FLJ21069 (FLJ21069), mRNA	NM_024692	Upregulated 0.19	12
2441	CK327184	Intronic	ITGA6	Homo sapiens integrin, alpha 6 (ITGA6), mRNA	NM_000210	Upregulated 0.22	13
3131	BE156190	Intronic	PALLD	Homo sapiens palladin, cytoskeletal associated protein (PALLD), mRNA	NM_016081	Upregulated 1.02	14
4255	BF332494	Intronic	SAP18	Homo sapiens sin3-associated polypeptide, 18kDa (SAP18), mRNA	NM_005870	Upregulated 1.13	15
2005	BE087892	Intronic	STARD13	Homo sapiens STARD domain containing 13, mRNA	BX647695	Upregulated 1.15	16
3479	BE061111	Intronic	TBN	Homo sapiens taube nuss homolog (mouse) (TBN), mRNA	NM_138572	Upregulated 1.36	17
2802	AW805635	Intronic	ERG	Homo sapiens v-ets erythroblastosis virus E26 oncogene like (avian) (ERG), transcript variant 2, mRNA	NM_004449	Upregulated 0.95	18
3912	BG004239	Intronic	NBPF9	Homo sapiens clone IIIa12 NBPF9 isoform 1 mRNA, complete cds, alternatively spliced.	AY894565	Upregulated 1.42	19

3115	BF350758	Intronic	KDEL2	Homo sapiens KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 2 (KDEL2), mRNA	NM_006854	Upregulated	0.97	20
2761	BF805741	Intronic	PMF1	Homo sapiens polyamine-modulated factor 1 (PMF1), mRNA	NM_007221	Upregulated	0.95	21
2821	CK327191	Intronic	PPP3CB	Homo sapiens protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta) (PPP3CB), mRNA	NM_021132	Upregulated	0.99	22
3179	BF894911	Intronic	P2RX1	Homo sapiens purinergic receptor P2X, ligand-gated ion channel, 1 (P2RX1), mRNA	NM_002558	Upregulated	1.30	23
1200	AW880165	Intronic	TGFBR2	Homo sapiens transforming growth factor, beta receptor II (70/80kDa) (TGFBR2), mRNA	NM_003242	Upregulated	0.81	24
2786	AW805625	Intronic	N.D.	Homo sapiens hypothetical protein FLJ10154 (FLJ10154), mRNA	NM_018011	Upregulated	0.80	25
4357	CK327190	Intronic	ZNF644	Homo sapiens zinc finger motif enhancer binding protein 2 (Zep-2), transcript variant 3, mRNA	NM_032186	Upregulated	0.86	26
4360	BE168353	Intronic	PASK	Homo sapiens PAS domain containing serine/threonine kinase (PASK), mRNA	NM_015148	Upregulated	1.21	27
2758	BF848956	Intronic	MYO5A	Homo sapiens myosin VA (heavy polypeptide 12, myoxin) (MYO5A), mRNA	NM_000259	Upregulated	2.36	28
3483	BE062809	Intronic	UBE2V1	Homo sapiens ubiquitin-conjugating enzyme E2 variant 1 (Kua-UEV), transcript variant 1, mRNA	NM_199203	Upregulated	1.94	29
1296	BF768642	Intronic	N.D.	AGENCOURT_6594963 NIH_MGC_41 Homo sapiens cDNA clone IMAGE:5466041 5', mRNA sequence	BM811380	Upregulated	1.73	30
1283	CK327189	Intronic	P2RY14	Homo sapiens purinergic receptor P2Y, G-protein coupled, 14 (P2RY14), mRNA	NM_014879	Upregulated	1.26	31
814	AW819948	Intronic	PPP2R2A	Homo sapiens protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform (PPP2R2A), mRNA	NM_002717	Upregulated	1.46	32
3900	AW819863	Intronic	PECAM1	Homo sapiens platelet/endothelial cell adhesion molecule (CD31 antigen) (PECAM1), mRNA	NM_000442	Upregulated	2.02	33
1690	BF926453	Intronic	XRCC1	Homo sapiens X-ray repair complementing defective repair in Chinese hamster cells 1 (XRCC1), mRNA	NM_006297	Upregulated	2.07	34
2458	BF926454	Intronic	ADD3	Homo sapiens adducin 3 (gamma) (ADD3), transcript variant 1, mRNA	NM_016824	Upregulated	2.26	35
1230	BF881469	Intronic	ALMS1	Homo sapiens Alstrom syndrome 1 (ALMS1), mRNA	NM_015120	Upregulated	1.82	36
4258	BF364131	Intronic	C16orf46	Homo sapiens hypothetical protein FLJ32702 (FLJ32702), mRNA	NM_152337	Upregulated	2.16	37
1629	CK327094	Intronic	ATF2	Homo sapiens activating transcription factor 2 (ATF2), mRNA	NM_001880	Upregulated	2.48	38
3503	BF364068	Intronic	MEG3	Homo sapiens hypothetical gene supported by BX161452 (LOC440199), mRNA	XR_000167	Upregulated	2.53	39

Androgen responsive intronic non-coding RNAs

Rodrigo Louro, Helder I. Nakaya, Paulo P. Amaral, Fernanda Festa, Mari C. Sogayar, Aline M. da Silva, Sergio Verjovski-Almeida and Eduardo M. Reis*

Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, 05508-900
São Paulo, Brazil.

*Corresponding author:
Dr. Eduardo M. Reis
Tel. +55-11-3091-2173
Fax: +55-11-3091-2186

Email addresses:
R.L.: rolouro@iq.usp.br
H.I.N.: hnakaya@iq.usp.br
P.P.A.: ppamaral@iq.usp.br
F.F.: fernanda@cecm.usp.br
M.C.S.: mcsoga@iq.usp.br
A.M.D.S.: almsilva@iq.usp.br
S.V.A.: verjo@iq.usp.br
E.M.R.: emreis@iq.usp.br

Abstract

Background

Transcription of large numbers of non-coding RNAs originating from intronic regions of human genes has been recently reported, but mechanisms governing their biosynthesis and biological functions are largely unknown. In this work, we evaluated the existence of a common mechanism of regulation of transcription shared by protein-coding mRNAs and intronic RNAs by measuring the effect of androgen on the transcriptional profile of a prostate cancer cell line.

Results

Using a custom cDNA microarray enriched in intronic transcribed sequences, we found 39 intronic non-coding RNAs whose abundance levels were significantly regulated upon androgen exposure. Orientation-specific RT-PCR indicated that 10 out of 13 were transcribed in the antisense direction. These transcripts are long (0.5–5 kb), unspliced and apparently do not code for proteins. Interestingly, we found that the relative abundance of androgen-regulated intronic transcripts could be correlated either to the abundance of the corresponding protein-coding gene (*asGAS6* and *asDNAJC3*) or to the alternative usage of exons (*asKDELR2* and *asITGA6*) in the corresponding protein-coding transcripts. Binding of androgen receptor to a putative regulatory region upstream from *asMYO5A*, an androgen-regulated antisense intronic transcript was confirmed by chromatin immunoprecipitation.

Conclusions

Altogether, these results indicate that at least a fraction of naturally transcribed intronic non-coding RNAs may be regulated by common physiological signals such as hormones, and further corroborate the notion that the intronic complement of the transcriptome play functional roles in the human gene expression program.

Background

Non-coding RNAs (ncRNAs), particularly those transcribed from genomic regions spanning introns of spliced genes have been proposed as a fundamental advance in the genetic operating system of higher organisms by influencing the genomic program of differentiation and development among individuals and species [1, 2]. Many functions such as transcriptional/translational regulation, RNA splicing, gene silencing, imprinting and dosage compensation have been assigned to mammalian ncRNAs [3]. Most functionally characterized ncRNAs belong to classes of small RNAs such as microRNAs [4] and small nucleolar RNAs [5]. While several novel long ncRNAs have been described in mammalian organisms [6, 7] few of them have been characterized in more detail, being the ~15 kb Xist RNA involved in inactivation of chromosome X an exception [8].

Recent studies based on computational analysis and experimental validations have revealed that introns, as well as intergenic regions constitute major sources of ncRNAs [3, 9, 10]. Extensive analyses of the human EST database particularly focused on spliced mRNAs and ESTs that overlap at least one exon have identified an abundant number of sense-antisense transcript pairs in human [7, 11, 12], mice [13] and plants [14] pointing to a well-conserved mechanism of post-transcriptional regulation of gene expression in eukaryotes.

Recently, microRNAs were shown to be consistently altered in normal and tumor cells, allowing the identification of tissue-specific expression signatures [15]. Using a custom cDNA microarray enriched in ESTs corresponding to fragments of intronic transcripts, we have previously detected transcription of a set of totally intronic non-coding RNAs, which are long (0.6 to 1.1 Kb) unspliced and oriented in antisense direction relative to the corresponding protein-coding transcript [16]. Moreover, expression levels of 23 intronic ncRNAs were shown to correlate to the degree of tumor differentiation in prostate cancer [16]. The potential role of long non-coding RNAs in human cancer is also exemplified by *MALAT-1*, a 7 kb RNA

that was reported as being associated to metastatic non-small cell lung cancer [17]. Other reports have shown that totally or partially intronic long antisense transcripts may have regulatory functions, such as modulating the methylation status of promoters [18] or the alternative splicing pattern of the corresponding protein-coding gene [19]. However, little attention has been given to the mechanism used by the cell to control the expression of these long intronic RNAs.

Androgens have been implicated in central events governing the regulation of distinct and diverse physiological processes in normal and neoplastic prostate cells. These hormones are known to promote cell division, proliferation of epithelial cells, modulate programmed cell death, cell quiescence [20, 21] and regulate cellular metabolism [22]. More recently, several studies have characterized the temporal transcriptional program reflecting the cellular response to androgens, which led to the identification of novel androgen regulated protein-coding genes [23-25].

In the present study, we used a custom spotted cDNA microarray enriched in intronic transcribed sequences and a prostate cancer cell line (LNCaP) cultured in the presence of a synthetic androgen to identify androgen-responsive non-coding intronic RNAs. Using orientation-specific reverse-transcription followed by PCR we determined the sense/antisense direction of a sub-set of intronic transcripts. We found that the abundance of these intronic RNAs in androgen-treated and control cells may correlate to the abundance of message or to the alternative usage of exons in the corresponding protein-coding transcript. Additional approaches such as RACE and chromatin immunoprecipitation assays demonstrated that these intronic messages represent long and unspliced RNA transcripts that may be directly regulated by androgen receptor *in vivo*. These findings indicate that transcription of long intronic non-coding RNAs may be controlled by mechanisms common to protein-coding transcripts, such as those involving hormonal control of gene promoter activation. In addition, our results

indicate that a fraction of this class of long unspliced intronic RNAs may have a role in post-transcriptional regulation of gene expression by modulating transcript stability and alternative splicing.

Results and Discussion

Identification of intronic RNAs regulated by androgen

We used a spotted cDNA microarray platform enriched in intronic transcripts [16] to identify intronic RNAs as well as exonic protein-coding messages that have their transcription levels significantly changed when hormone-responsive LNCaP prostate cancer cells were treated with the synthetic androgen R1881. Expression levels of *KLK3* (PSA, Prostate Specific Antigen) [24] and *TMEPAI* [26], two genes previously shown to be up-regulated by androgen in prostate epithelium were determined by Real-Time PCR, confirming the effectiveness of androgen to induce responsive genes in our experimental conditions (Supplementary Fig. 1). Next, the significance analysis of microarrays (SAM) approach [27] was used to identify transcripts that showed statistically significant changes in their abundance levels following exposure to androgen for 6 h up to 48 h. Only transcripts showing significant expression changes (fold-change ≥ 1.5 and FDR $< 5\%$) in at least 3 consecutive time points were selected for further analysis. We found 168 protein-coding exonic messages (Supplementary Fig. 2) and 39 intronic RNAs (Fig. 1A) that have statistically different patterns of temporal expression in androgen-treated cells when compared to untreated controls. The complete list of androgen-regulated exonic transcripts coding for proteins is shown in Supplementary Table 1.

Androgen-regulated intronic RNAs could be grouped into three clusters based on their different temporal patterns of expression following androgen exposure (Fig. 1B). Cluster I contained RNAs whose levels were 2 to 3-fold down-regulated within 24 h after androgen exposure and were restored at 48 h. Clusters II and III included intronic RNAs that were up-regulated by androgen within 24 h and were restored thereafter. Cluster II grouped RNAs that were 2 to 3-fold increased, whereas Cluster III grouped RNAs that were up-regulated by 3 to

6-fold. All of these intronic RNAs are unspliced messages, and apparently have no coding potential as determined by ESTScan analysis [28]. A complete list of androgen-regulated intronic RNAs is presented as Supplementary Table 2.

Most mammalian snoRNAs [5] and a large fraction of microRNAs [29] are derived from intronic sequences of protein-coding and noncoding genes. To investigate the possibility that some of the androgen-regulated intronic transcripts might be primary transcripts that after processing could generate known small RNAs of those classes, we compared the sequence of the 39 androgen-regulated intronic RNAs with those of 346 snoRNAs (<http://www-snoRNA.biotoul.fr/>) and 383 microRNAs [30], and no similarity was found.

We sought to investigate the enrichment of specific Gene Ontology (GO) categories among androgen-regulated exonic and intronic transcripts using BiNGO, a gene ontology comparison tool [31]. We postulated that intronic transcripts represent a new class of RNAs which might act as *cis*-regulatory factors [32]. Thus, for each androgen-regulated intronic transcript we used the GO annotation assigned to the corresponding protein-coding mRNA mapping to the same genomic *locus*. We found no significant enrichment of specific GO categories ($p < 0.05$) among the selected sets of exonic or intronic androgen-regulated transcripts, when a correction for multiple testing was applied. However, we noted that 13 out of the set of 39 androgen-regulated intronic transcripts belong to the *signal transduction* GO category (GO: 0007165). Indeed, some of these *loci* have already been implicated in cellular events related to prostate cell growth and differentiation. For example, the protein encoded by *GAS6* mRNA is a ligand of Axl receptor tyrosine kinase and the Gas6/Axl complex has been demonstrated to present a mitogenic activity in human prostatic cancer cell lines, by modulating PI3K/AKT and MEK signal transduction pathways [33]. Similarly, the transcription factor encoded by *ERG*, a member of the ETS family of central genes involved in integrating signals that regulate cell growth and differentiation, stress responses, and

tumorigenesis has been previously identified as the most frequently overexpressed proto-oncogene in malignant prostate epithelial cells [34].

Characterization of androgen-regulated intronic RNAs

Orientation-specific reverse transcription followed by PCR was performed (see *Materials and Methods* for details) for a selected set of androgen-regulated intronic RNAs (Fig. 1C). Sense/antisense orientation of the intronic RNA relative to the corresponding protein-coding gene could be unambiguously assigned in most cases (13 out of 17). For most of these (10 out of 13) an intronic antisense RNA could be detected (Fig. 1D). For six of these antisense RNAs (*DNAJC3*, *SAP18*, *KDEL2*, *FLJ10154*, *MYO5A* and *ATF2*) we detected an overlapping intronic sense message transcribed in the opposite strand. Four RNAs (*GAS6*, *RAB25*, *ITGA6* and *ADD3*) were only detected in antisense orientation. Three intronic messages (*ACTN4*, *STARD13* and *ERG*) were only detected in sense orientation (Fig. 1D). The intronic sense transcripts of *ACTN4*, *STARD13* and *ERG* did not show coding potential as determined by ESTScan. Without further experimentation, we can not determine if these intronic sense transcripts are novel non-coding exons present in alternative splicing forms, or if they represent ncRNAs originated from independent intronic transcriptional units located in the sense strand. The latter is in line with sense intronic transcriptional units that have been observed elsewhere [9, 35].

Next, RACE-PCR experiments with cDNA from prostate adenocarcinoma cells were performed for a selected set of intronic RNAs to obtain their full-length sequences. This selected set comprised intronic RNAs transcribed in antisense (*GAS6* and *ITGA6*), sense (*STARD3*) or both antisense/sense orientation (*MYO5A* and *DNAJC3*) as determined by orientation-specific RT-PCR. RACE experiments revealed that antisense intronic RNAs from *DNAJC3*, *MYO5A* and *ITGA6* were long (0.5 to 0.8 kb), unspliced transcripts that mapped

totally inside the intronic regions of the corresponding protein-coding gene (Fig. 2A–C). An unspliced, totally intronic sense transcript from *STARD3* with approximately 1.4 kb was obtained (Fig. 2D). Further analysis of the coding-potential by ESTScan indicates that extended intronic *STARD3* could generate a 74-residue peptide, but subsequent BLASTp searches [36] did not show any similarity of this small predicted product to known proteins in GenBank or to conserved protein domains at CDD [37]. The intronic antisense RNA mapping to the *GAS6* locus had already been confirmed and extended by an mRNA with 3.2 kb cloned from a brain cDNA library available in GenBank [GenBank:AK123478]. We performed PCR with primers complementary to the cDNA probe mapping to an intron of *GAS6* deposited in the microarray and to the AK123478 mRNA. We obtained a 4.2 kb amplicon, demonstrating that the *GAS6* antisense intronic RNA transcribed in prostate cells was even longer (at least 4.2 kb) than the previously known RNA (Fig. 2E, yellow bar). *In silico* analysis using ESTScan [28] failed to identify ORFs for extended intronic sequences transcribed from *DNAJC3*, *ITGA6*, *MYO5A* and *GAS6* loci.

A strand-specific Northern blot confirmed the expression of a long *GAS6* antisense transcript with approximately 5.0 kb in different human tissues (Fig. 2F). Observation that androgen-regulated intronic RNAs could be detected in tissues not particularly responsive to this hormone (Fig. 2 panel F) suggests that regulation of the steady-state levels of these transcripts may also proceed through an androgen-independent pathway.

Effect of androgen in the levels of intronic RNAs and corresponding protein-coding transcripts

Among the androgen-regulated transcripts identified in our analysis, only *KDEL2* and *ADD3* showed detectable changes in expression levels of both intronic and exonic transcripts. The intronic *ADD3* transcript increased 4.9-fold ($\log_2$ ratio = 2.3) in androgen-

treated cells, while the corresponding protein-coding transcript decreased 3.3-fold (log2 ratio = -1.7) in the same experiment. Intronic *KDELR2* transcript levels increased 2-fold (log2 ratio = 1), while the corresponding exonic transcript increased 2.3-fold (log2 ratio = 1.2) in cells exposed to androgen relative to control cells (Supplementary tables 1 and 2). These results showed that for a given intronic/exonic pair transcribed in the same *locus*, the intronic message may be directly or inversely modulated with respect to the exonic one, following androgen treatment.

Orientation-specific reverse transcription followed by quantitative Real-Time PCR using RNA from treated and control untreated cells was used to confirm the change in abundance of intronic transcripts following androgen exposure. We tested two intronic RNAs that were detected only in antisense orientation (*GAS6* and *ITGA6*), and two intronic messages in which both antisense/sense transcription were measured (*DNAJC3* and *KDELR2*) (Fig. 3A–D, panels “intronic”). In each case, we quantified in parallel the expression levels of exons of the corresponding protein-coding gene (Fig. 3A–D, panels “exonic”). We confirmed a significant ($p < 0.005$) down-regulation of the intronic antisense transcript to *GAS6* following androgen stimulation (Fig. 3A, panel “intronic”). Interestingly, protein-coding messages for *GAS6* that contain exons 4 and 5 are significantly up-regulated ($p < 0.025$) under identical experimental conditions (Fig. 3A, panel “exonic”). We observed that the abundance of sense and antisense intronic RNAs at the *DNAJC3* *locus* are modulated in opposite directions following androgen treatment (Fig. 3B, panel “intronic”). The up-regulation observed for the sense intronic RNA from *DNAJC3* after androgen exposure is in agreement with the results from the microarray experiments (Fig. 3B, panel “intronic”), and was also observed for the protein-coding *DNAJC3* transcript (Fig. 3B, panel “exonic”). Modification in the relative abundance of intronic RNAs from *KDELR2* and *ITGA6* following androgen treatment was confirmed (Fig. 3C and D, panels “intronic”). For *KDELR2*, both antisense and

sense intronic RNA levels were up-regulated by androgen ($p < 0.0035$ and $p < 0.02$, respectively) (Fig. 3C, panel “intronic”). Likewise, the antisense intronic RNA at the *ITGA6* locus was significantly increased ($p < 0.0012$) in the presence of androgen (Fig. 3D, intronic panel). However, in these cases we could not detect significant changes in the relative abundance of the respective protein-coding transcripts when primers designed to measure constitutive exons were used (Fig. 3C and D, panels “exonic”).

To accumulate further evidence of functional involvement of antisense intronic non-coding transcripts on the splicing mechanism [19, 32], we investigated the existence of correlation between their relative abundance and the levels of the corresponding protein-coding mRNAs, using primers designed to probe regulated exons, e.g., exons that are absent in at least one documented transcript isoform. At the *KDELR2* locus we found that exons 2/3 were markedly up-regulated ($p < 0.015$) following androgen treatment (Fig. 3C, panel “exonic”). Noteworthy, exon 2 is absent in one *KDELR2* transcript isoform [GenBank: BP269541]. Therefore, it is conceivable that the observed increase in the level of intronic antisense RNA for *KDELR2* following androgen treatment may induce increased retention of exon 2. We could not detect any significant androgen-driven change in the abundance of transcripts containing exons 1/3 or exons 1/4 (Fig. 3C, panel “exonic”). Notwithstanding, we observed that exon 1 for *KDELR2* was down-regulated ($p < 0.026$) following androgen treatment when a pair of primers designed to probe only this exon was used (Fig. 3C, panel “exonic”). This result can only be accounted for by the accumulation upon androgen treatment of a yet uncharacterized alternative *KDELR2* transcript with a shorter 3' end.

At the *ITGA6* locus we found that the abundance of a protein-coding transcript isoform containing a longer terminal exon (exon 22D) was down-regulated ($p < 0.04$) upon androgen-treatment (Fig. 3D, panel “exonic”) and thus, inversely correlated to the significant ($p < 0.0012$) increase observed in the antisense intronic RNA under the same conditions (Fig. 3D,

panel “intronic”). However, it is not possible to determine whether this result is actually due to down-regulation of the *ITGA6* mRNA isoform containing exon 22D, since we also detected down-regulation of another intronic transcript naturally transcribed in the *ITGA6* locus [GenBank:BX537483] that spans the 3’ terminal exon of the gene (Fig. 3D).

Non-coding RNAs have been shown to affect splicing of protein-coding genes [19, 38]. *HBII-52*, a small nucleolar RNA (snoRNA) was found to regulate alternative splicing of the serotonin receptor gene *5-HT2CRA* in *trans*, by binding to a silencing element in exon Vb of *5-HT2CRA* [38]. Defective *5-HT2CRA* pre-mRNA processing, induced by the *HBII-52* snoRNA, accounts for the congenital Prader-Willi syndrome [38]. A previous study on the *FAS* locus has identified a partially intronic antisense RNA that is naturally transcribed from the first intron of the gene, which was named *Saf* [19]. *In vitro* over-expression of *Saf* induced skipping of several 3’ exons of *Fas* mRNA, resulting in a Fas protein that was no longer able to anchor to the cell membrane or to induce Fas mediated apoptosis [19]. Our results identify novel intronic antisense RNAs that might be involved in regulatory mechanisms of alternative splicing of protein-coding genes, in analogy with those observed for *HBII-52* and *Fas/Saf*.

Increased binding of AR to an ARE motif upstream of *asMYO5A*

Androgen-mediated transcriptional regulation involves direct hormone interaction with the androgen receptor (AR) protein and the translocation of this complex to the nucleus promoting specific interactions with DNA motifs (i.e. androgen response elements, ARE) located near or within the target sequences [39]. To support a mechanism of direct transcriptional control of androgen-regulated intronic RNAs we used a bioinformatics approach [24] to look for putative ARE motifs located upstream of the identified androgen-regulated intronic RNAs. At least one putative ARE motif was identified for 24 androgen-

regulated intronic transcript. A complete list of identified putative ARE motifs is presented as Supplementary information (Supplementary Figure 3).

Chromatin immunoprecipitation (ChIP) assays were performed to verify functional binding of AR to a putative regulatory ARE motif located upstream of the intronic antisense RNA mapping to the *MYO5 locus*. ChIP assays revealed that following androgen treatment there was a 2 to 3-fold increase in *asMYO5A* ARE motifs bound to AR (Fig. 6B and C, *asMYO5A* ARE). The *asMYO5A* ARE motif is located at a 5' genomic region 812 bp upstream of the region where the antisense *MYO5A* intronic transcript maps. Binding of AR to a previously described functional ARE that regulates transcription of the *PSA* gene was tested in parallel, as a positive control for the ChIP assay (Fig. 4B, *PSA* ARE). A genomic region upstream from *TUBA6*, a gene that does not contain putative ARE motifs in its promoter region, was used as a negative control for non-specific immunoprecipitation of DNA fragments (Fig. 6B, *TUBA6*).

The presence of a weak band for *asMYO5A* ARE in control cell lysates immunoprecipitated with anti-AR (Fig. 6B, *asMYO5A* ARE) indicates that binding of AR to *asMYO5A* ARE may occur to some extent in untreated cells. This may be explained by the presence of trace amounts of androgen remaining in charcoal-stripped fetal calf serum added to the cell culture medium. An alternative explanation for the background signal of *asMYO5A* ARE measured in control cells may be the existence of small amounts of AR constitutively bound to the chromatin, as described for another transcription factor, the cyclic AMP (cAMP)-responsive element-binding protein (CREB) [40].

The increased binding of AR to the *asMYO5A* ARE motif is in agreement with the androgen-induced increased expression of the antisense intronic *MYO5A* RNA (*asMYO5A*) detected by microarray (Fig. 1), and confirmed by orientation-specific reverse transcription followed by Real-Time PCR (Fig. 4A). *MYO5A* encodes a protein which is an actin-based

processive motor involved in intracellular trafficking and exocytosis [41]. Tissue-specific Myo5a isoforms are originated from alternative exon usage of the primary transcript, which in turn is thought to modulate the protein specificity for different cargoes within the cell [42]. It is conceivable that the increase of *asMYO5A* levels in prostate cells in response to androgen will affect the relative abundance of the alternatively spliced forms of *MYO5A*, thus modulating protein activity/specificity. Together, these results suggest that functional binding of androgen-AR activated complexes to regulatory ARE regions is required to modulate the abundance of intronic antisense RNAs following androgen treatment of prostate cells.

Conclusions

Although there is still no conclusive evidence about the molecular mechanisms, we have postulated that intronic transcripts exert their regulatory functions by acting in *cis* through Watson-Crick base pairing to complementary sense pre-mRNA [32]. The results presented here corroborate the notion that at least a fraction of these intronic transcripts may directly affect the abundance and/or splicing of the protein-coding transcript originated in the opposite strand, and are in line with previous reports in the literature [43, 44].

It is not yet clear if the intronic transcripts identified in this work would exert their potential regulatory functions as long (0.5 – 5 kb) RNAs, or alternatively, as shorter RNAs resulting from processing of the primary transcript. Processing of these androgen-regulated ncRNAs into shorter RNAs of defined sizes would also raise the possibility that some of them may modulate transcriptional or post-transcriptional events in messages from different *loci*, in a manner similar to microRNAs [4]. The identification of possible short RNAs modulated by androgen will require further work, such as 15% denaturing PAGE/Northern blots of prostate cells exposed to androgen in time-course experiments to document the accumulation of the

long intronic RNAs, the time course of their conversion to short RNAs, and the accumulation of possible short RNA intermediates.

Previous reports have shown that a large number of transcription factor binding sites (TFBSs) for well-characterized transcription factors such as Sp1, c-Myc, p53 and Creb lie upstream to transcriptionally active intronic regions [40, 45], some of which generate transcripts oriented antisense relative to the known gene [45]. Noteworthy, Cawley et al. [45] found that comparable fractions of protein-coding genes and non-coding RNAs having an experimentally determined upstream TFBS region were regulated following exposure to retinoic acid. The results presented here provide further evidence that ncRNAs transcribed from intronic segments of the human genome may be regulated by physiological signals that commonly act on protein coding-transcripts such as hormones. Similar results were observed in mice when the effect of lipopolysaccharide stimulation on the abundance of several long ncRNAs was measured [46]. Together, these results strongly suggest that at least in mammalian cells intronic ncRNAs are transcribed in a regulated fashion and therefore must exert physiological roles, which corroborates the existence of a yet under-appreciated complex RNA-based regulatory system involving non-coding messages in higher organisms [2, 3, 47].

Materials and Methods

Cell Culture and RNA isolation

The experimental design was based on recent published work [24]. In summary, prostate carcinoma cell lines LNCaP, DU145 and PC3 were obtained from ATCC and maintained using the suggested medium supplemented with 10% (vol/vol) fetal calf serum (FCS), 3 mM L-glutamine, 100 µg/ml streptomycin and 100 U/ml penicillin. LNCaP cells were grown in RPMI 1640 whereas DU145 and PC3 cells were cultured in DMEM. For the androgen-

response experiments, LNCaP cells were cultured for 24 hours in androgen-deprived medium, RPMI 1640 with 5% (wt/vol) charcoal-stripped FCS (Life Technologies). After 24h, cells were placed in fresh RPMI 1640 medium, with 10% (wt/vol) charcoal-treated FCS and either 1 nM of the synthetic androgen R1881 (NEN - Life Sciences Products) or equivalent amount of vehicle (ethanol) was added. Cells were harvested after 0, 6, 9, 12, 18, 24, and 48 hours for RNA isolation. Two independent cultures of androgen-treated cells, as well as of control cells (independent biological replicates), were used for RNA isolation.

In parallel, cells were harvested after 4 h of hormone treatment for isolation of DNA that was subsequently used in chromatin immunoprecipitation (ChIP) assays, as described below. Total RNA was purified from experimental and control cells by CsCl cushion [48], after cell lysis (4M guanidine isothiocyanate; 0.1M β -mercaptoethanol; 25 mM sodium citrate pH 7.0). Total RNA (100 μ g sample) was treated with RNase-free DNase (RNeasy – Qiagen) for 15 min in the medium recommended by the manufacturer to minimize genomic DNA contaminants and kept at -80 °C until use.

Microarray experiments

The microarray platform used in this study was constructed using PCR amplified cDNA fragments derived from ORESTES clones [49], and printed on Type 7 STAR glass slides (Amersham Biosciences Inc.) as previously described [16]. It consists of a 9,216-element microarray (3,355 unique cDNA probes spotted in duplicate, plus positive and negative controls) enriched in probes representing intronic (822 probes) and intergenic (241 probes) ncRNAs. For each time point, 15 μ g of DNase-treated total RNA from androgen-treated or untreated control LNCaP cells were used to generate fluorescent targets by incorporation of Cy5-dUTP in reverse transcription reactions using the CyScribe First strand cDNA labeling kit (Amersham Biosciences) and the protocol recommended by the manufacturer. Fluorescent

targets labeled with Cy3-dUTP were generated in parallel from a reference RNA pool consisting of equimolar quantities of DNase-treated total RNA from LNCaP, DU145 and PC3 cells.

For each independent biological replicate hybridization, fluorescently labeled targets from each experimental time point and the reference pool were combined and hybridized on an Automated Slide Processor (Amersham Biosciences) using protocols recommended by the manufacturer. For each time point, two independent biological replicate hybridizations were performed; each hybridization used one RNA, isolated from a different biological replicate of androgen-treated or untreated control cells. Processed slides were scanned in a GenePix 4000B microarray scanner (Axon Instruments Inc.). More detailed information about microarray experiments and analysis is provided following MIAME guidelines [50] as supplementary material.

Data acquisition and statistical analysis

Background-subtracted artifact-removed median intensities of both Cy3 and Cy5 emissions were extracted for each spot from raw images using ArrayVision V.8.0 software (Imaging Research Inc.). Only spots with signal intensity above the average plus 3 standard deviations from a set of negative controls (plant and bacterial DNA) were used in subsequent analyses. As each microarray contains two replicates of each spotted cDNA, a total of four replicate measurements were obtained for each spot at each time point. We have evaluated the average coefficient of variation between replicates of expression ratios between treated and control cells calculated in two different ways; for each time point, expression ratios between treated and control cells were reconstructed either from (i) ratios relative to the reference pool or from (ii) Cy5-intensities obtained from androgen-treated and control cells using a one-color approach [51]. A smaller average coefficient of variation between replicates was obtained

with the latter approach, and therefore we opted to use only the Cy5-labelled sample measurements to calculate the treatment to control sample ratios. We used the mean intensity signal (40% trimmed) of each dataset for normalization between experiments. Raw and normalized microarray intensities are deposited at the Gene Expression Omnibus (GEO – <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE5345.

Transcripts with statistically significant changes in expression in response to androgen stimulation were identified using the SAM approach [27]. Only transcripts showing significant expression changes in at least 3 consecutive time points were selected for further analysis, using as parameters: two-class response (paired data), 1000 permutations, K-Nearest Neighbors Imputer, fold-change ≥ 1.5 and false discovery rate (FDR) $< 5\%$. Expression profiles of SAM selected transcripts were grouped using hierarchical clustering (UPGMA with Euclidean distance) and visualized using Spotfire Decision software Site (Spotfire Inc.).

Orientation-specific RT-PCR and orientation-specific Real-time quantitative RT-PCR

Orientation-specific RT-PCR was performed as previously described [16]. Briefly, 4 μg of DNase-treated total RNA from androgen-treated or control LNCaP cells were used to generate strand-specific cDNA. RNA samples were selected from experimental time points that showed higher fold-change relative to control cells as identified in microarray hybridizations. The same RNA samples were used in microarray hybridizations and Real-Time PCR reactions. To determine the abundance of sense/antisense intronic transcripts in androgen-treated and control cells, we performed one orientation-specific reverse transcription reaction for each strand using a single RNA sample, and gene specific (forward or reverse) primers. RNA samples plus 0.9 μM of specific primer for either antisense or sense strand of each intronic segment were heated to 75 °C for 10 minutes in order to minimize artifacts derived

from the formation of secondary structure. Reverse transcription was performed at 50 °C for 60 min in a 50 µl volume, using 200 units of SuperscriptII reverse transcriptase (Invitrogen). Controls for the absence of self-priming were obtained by performing reverse transcription in the absence of primers, and controls for the absence of genomic DNA contamination were obtained by incubation with primers in the absence of reverse transcriptase enzyme. A reaction containing a primer complementary to the coding region of *TUBA6* (reverse primer: AGTGCCAGTGCGAACTTCATC) was carried out in parallel. Reverse transcriptase was denatured for 30 min at 95 °C. One pair of internal primers for PCR amplification of either the tubulin gene or the desired target intronic segment under study was added. For each intronic transcript, cDNAs with sense or antisense orientation relative to the corresponding gene were analyzed by three replicate measurements in quantitative Real-Time PCR reactions. For exonic transcripts, cDNA was obtained by reverse transcription of RNA samples using oligo-dT primers, and the abundance of each message was also determined by three replicate measurements in Real-Time PCR reactions with gene specific primers. All Real-Time PCR reactions were carried out in a volume of 20 µl for 40 cycles, and according to the manufacturer's instructions for Sybr-Green PCR Core Reagent (Applied Biosystems) using the GenAmp 5700 Sequence Detector (Applied Biosystems). For quantitative results, the levels of each transcript detected by Real-time quantitative PCR was normalized to the level of tubulin, and represented as fold-change using the delta Ct method [52]. Quantitative results for a transcript in androgen treated cells are expressed as relative fold-change with respect to the value of the target transcript in the control (ethanol vehicle), which was set at unity. For each measured transcript, statistical significance for the differential expression between treated and control cells (p-values) were determined using a Student's t-test, assuming a two-tail distribution and a heteroscedastic variance model.

RACE-PCR and Northern Blot

For RACE-PCR we used a commercial cDNA library prepared from poly(A)+ RNA isolated from a human prostatic adenocarcinoma cell line (Human XG Marathon-Ready cDNA, cat. 7498-1 – BD Biosciences – Palo Alto, CA). Two rounds of PCR reactions were performed for each intronic transcript using two pairs of gene specific primers, in nested PCR reactions. To obtain a long (~5 kb) transcript corresponding to *asGAS6*, a gene-specific PCR was performed using a forward primer complementary to the partial *asGAS6* cDNA sequence deposited in the microarray and a reverse primer complementary to the antisense transcript deposited in GeneBank under accession number AK123478. The PCR product was sequenced to confirm the identity. To confirm the size and orientation of the full-length *asGAS6* obtained by PCR, a Northern Blot was performed using a commercial membrane containing 2 µg poly(A)+ RNA from ten human tissues (Ambion First Choice Human Blot 4), using a probe complementary to the antisense strand of *GAS6* intronic transcript spotted on the microarray. Sense α -[32 P]-CTP-labeled RNA probes were obtained by *in vitro* transcription using T7 RNA polymerase (Riboprobe – Promega). Hybridization and wash procedures were carried out as described in the human blot manufacturer's protocol (Ambion, Inc.). The probed blot was processed in a phosphorimager (Storm, Molecular Dynamics).

Putative Androgen Responsive Elements (AREs) identification and Chromatin Immunoprecipitation (ChIP) assay

Using a bioinformatics approach, we searched for regions with similarity to the palindromic ARE consensus AGAACAnnnTGTTCT [39], obtained from the TRANSFAC database of eukaryotic *cis*-acting regulatory DNA elements [53]. The search was conducted within a 3 kb region of genomic sequence, upstream of the putative transcriptional start sites of the 39 androgen-regulated intronic transcripts identified in the microarray analysis. When available,

RACE-extended fragments were used. Chromatin Immunoprecipitation (ChIP) assay was performed as described in the literature [54-56], and specific PCR primers were designed to detect binding of the androgen receptor to the putative ARE promoter sequence of *asMYO5A*, identified by the *in silico* analysis. The Chromatin Immunoprecipitation Assay kit (Upstate Biotechnology Inc.) was used, with the protocols recommended by the manufacturer. Briefly, LNCaP cells treated with androgen or vehicle as described above were fixed with 1 % formaldehyde for 10 min at room temperature and lysed with SDS Lysis Buffer. Cell lysates were sonicated to shear the DNA between 200-1000bp using Sonic Dismembrator Model 500 (Fischer Scientific). 100 μ l lysate aliquots containing approximately 2×10^6 cells were diluted in 900 μ l dilution buffer and 10 μ l from each aliquot (1% of total lysate) were separated and used subsequently as loading controls for input DNA. Parallel chromatin immunoprecipitation reactions were performed with either 3 μ g polyclonal antibody anti-AR (H-280, Santa-Cruz) or 1 μ g Normal Mouse IgG (negative control, Upstate). Immunoprecipitated chromatin fractions and respective loading control fractions were treated for cross-link reversal, treated with RNase A and proteinase K, and DNA fragments were purified using Spin Columns. 2 μ l from each eluted DNA were used as template in PCR reactions with the following primers flanking the putative *asMYO5A* ARE: forward, TTTGCTGGATAAGGATTCCCA; reverse, AGTTCCTTAACTTTAGCTTAGAAAAAGGA. Similar reactions were performed with control primers designed to the known AREIII motif of *PSA* (Prostatic Specific Antigen) enhancer (positive control) (primer forward: GCTCAGCCTTTGTCTCTGATGA, reverse: TGCAAGATGATATCTCTCTCAGATCC) and to *TUBA6* non-related upstream genomic sequence (negative control) (primer forward: GACTACAGGTGTGCGCCATCAT, reverse: TGCCGTGTTCCAGGCAGTAG). For each DNA target, semi-quantitative PCR was performed by running parallel reactions with increasing numbers of PCR cycles (25, 30 or 40 cycles). 25 and 30 PCR cycles resulted in non-saturated amplification of each DNA target in

immunoprecipitated and input DNA fractions from androgen-treated or control cells (data not shown). Equal volumes from each PCR reaction were loaded and separated in a 3% agarose gel. The intensity of PCR-amplified bands was determined by gel densitometry (Image Master, Pharmacia Inc.).

Accession numbers

We have deposited in GenBank the following sequence data under the indicated accession numbers: antisense intronic *ITGA6* [GenBank:DQ866759]; antisense intronic *DNAJC3* [GenBank:DQ866760]; antisense intronic *MYO5A* [GenBank:DQ866761]; intronic *STARD13* [GenBank:DQ866762]; antisense intronic *GAS6* [GenBank:DQ866763]; antisense intronic *RAB25* [GenBank:DQ866751]; antisense intronic *ADD3* [GenBank:DQ866752]; intronic *ACTN4* [GenBank:DQ866753]; intronic *SAP18* [GenBank:DQ866754]; intronic *ERG* [GenBank:DQ866755]; intronic *KDELR2* [GenBank:DQ866756]; intronic *FLJ10154* [GenBank:DQ866757]; and intronic *ATF2* [GenBank:DQ866758].

Authors' contributions

RL, HIN, PPA, AMDS, SVA and EMR conceived the experiments. RL, HIN, PPA and FF performed the experiments. RL, HIN, PPA, SVA and EMR analyzed the data. RL and EMR wrote the manuscript. MCS, AMDS, SVA and EMR contributed reagents/materials/analysis tools. All authors read and approved the final manuscript.

Acknowledgments

Authors thank Junio Cota Silva and Vinícius Coutinho for technical help with RACE experiments. This work was mainly funded by grant 02/13283-6 from Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) to SVA and AMDS, by grant 01/10707-7 to MCS from FAPESP and by fellowships to RL and HIN from FAPESP. Fellowships from Conselho

Nacional de Desenvolvimento Científico e Tecnológico (CNPq) supported PPA and FF. MCS, AMDS and SVA were partially supported by CNPq.

References

1. Mattick JS: **RNA regulation: a new genetics?** *Nat Rev Genet* 2004, **5**:316-323.
2. Mattick JS, Makunin IV: **Non-coding RNA.** *Hum Mol Genet* 2006, **15 Suppl 1**:R17-29.
3. Willingham AT, Gingeras TR: **TUF love for "junk" DNA.** *Cell* 2006, **125**:1215-1220.
4. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.
5. Kiss T: **Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions.** *Cell* 2002, **109**:145-148.
6. Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusic V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, Kawaji H, Kawasaki Y, Kedzierski RM, King BL, Konagaya A, Kurochkin IV, Lee Y, Lenhard B, Lyons PA, Maglott DR, Maltais L, Marchionni L, McKenzie L, Miki H, Nagashima T, Numata K, Okido T, Pavan WJ, Pertea G, Pesole G, Petrovsky N, Pillai R, Pontius JU, Qi D, Ramachandran S, Ravasi T, Reed JC, Reed DJ, Reid J, Ring BZ, Ringwald M, Sandelin A, Schneider C, Semple CA, Setou M, Shimada K, Sultana R, Takenaka Y, Taylor MS, Teasdale RD, Tomita M, Verardo R, Wagner L, Wahlestedt C, Wang Y, Watanabe Y, Wells C, Wilming LG, Wynshaw-Boris A, Yanagisawa M, Yang I, Yang L, Yuan Z, Zavolan M, Zhu Y, Zimmer A, Carninci P, Hayatsu N, Hirozane-Kishikawa T, Konno H, Nakamura M, Sakazume N, Sato K, Shiraki T, Waki K, Kawai J, Aizawa K, Arakawa T, Fukuda S, Hara A, Hashizume W, Imotani K, Ishii Y, Itoh M, Kagawa I, Miyazaki A, Sakai K, Sasaki D, Shibata K, Shinagawa A, Yasunishi A, Yoshino M, Waterston R, Lander ES, Rogers J, Birney E, Hayashizaki Y: **Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.** *Nature* 2002, **420**:563-573.
7. Yelin R, Dahary D, Sorek R, Levanon EY, Goldstein O, Shoshan A, Diber A, Biton S, Tamir Y, Khosravi R, Nemzer S, Pinner E, Walach S, Bernstein J, Savitsky K, Rotman G: **Widespread occurrence of antisense transcription in the human genome.** *Nat Biotechnol* 2003, **21**:379-386.
8. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N: **Requirement for Xist in X chromosome inactivation.** *Nature* 1996, **379**:131-137.
9. Bertone P, Stolc V, Royce TE, Rozowsky JS, Urban AE, Zhu X, Rinn JL, Tongprasit W, Samanta M, Weissman S, Gerstein M, Snyder M: **Global identification of human transcribed sequences with genome tiling arrays.** *Science* 2004, **306**:2242-2246.
10. Kampa D, Cheng J, Kapranov P, Yamanaka M, Brubaker S, Cawley S, Drenkow J, Piccolboni A, Bekiranov S, Helt G, Tammana H, Gingeras TR: **Novel RNAs Identified From an In-Depth Analysis of the Transcriptome of Human Chromosomes 21 and 22.** *Genome Res* 2004, **14**:331-342.

11. Shendure J, Church GM: **Computational discovery of sense-antisense transcription in the human and mouse genomes.** *Genome Biol* 2002, **3**:research0044.0041-research0044.0014.
12. Chen J, Sun M, Kent WJ, Huang X, Xie H, Wang W, Zhou G, Shi RZ, Rowley JD: **Over 20% of human transcripts might form sense-antisense pairs.** *Nucleic Acids Res* 2004, **32**:4812-4820.
13. Kiyosawa H, Yamanaka I, Osato N, Kondo S, Hayashizaki Y: **Antisense transcripts with FANTOM2 clone set and their implications for gene regulation.** *Genome Res* 2003, **13**:1324-1334.
14. Ma J, Morrow DJ, Fernandes J, Walbot V: **Comparative profiling of the sense and antisense transcriptome of maize lines.** *Genome Biol* 2006, **7**:R22.
15. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR: **MicroRNA expression profiles classify human cancers.** *Nature* 2005, **435**:834-838.
16. Reis EM, Nakaya HI, Louro R, Canavez FC, Flatschart AV, Almeida GT, Egidio CM, Paquola AC, Machado AA, Festa F, Yamamoto D, Alvarenga R, Da Silva CC, Brito GC, Simon SD, Moreira-Filho CA, Leite KR, Camara-Lopes LH, Campos FS, Gimba E, Vignal GM, El-Dorry H, Sogayar MC, Barcinski MA, Da Silva AM, Verjovski-Almeida S: **Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer.** *Oncogene* 2004, **23**:6684-6692.
17. Ji P, Diederichs S, Wang W, Boing S, Metzger R, Schneider PM, Tidow N, Brandt B, Buerger H, Bulk E, Thomas M, Berdel WE, Serve H, Muller-Tidow C: **MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer.** *Oncogene* 2003, **22**:8031-8041.
18. Imamura T, Yamamoto S, Ohgane J, Hattori N, Tanaka S, Shiota K: **Non-coding RNA directed DNA demethylation of Sphk1 CpG island.** *Biochem Biophys Res Commun* 2004, **322**:593-600.
19. Yan MD, Hong CC, Lai GM, Cheng AL, Lin YW, Chuang SE: **Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas.** *Hum Mol Genet* 2005, **14**:1465-1474.
20. Isaacs JT, Lundmo PI, Berges R, Martikainen P, Kyprianou N, English HF: **Androgen regulation of programmed death of normal and malignant prostatic cells.** *J Androl* 1992, **13**:457-464.
21. Geck P, Szelei J, Jimenez J, Lin TM, Sonnenschein C, Soto AM: **Expression of novel genes linked to the androgen-induced, proliferative shutoff in prostate cancer cells.** *J Steroid Biochem Mol Biol* 1997, **63**:211-218.
22. Swinnen JV, Verhoeven G: **Androgens and the control of lipid metabolism in human prostate cancer cells.** *J Steroid Biochem Mol Biol* 1998, **65**:191-198.
23. Xu LL, Su YP, Labiche R, Segawa T, Shanmugam N, McLeod DG, Moul JW, Srivastava S: **Quantitative expression profile of androgen-regulated genes in prostate cancer cells and identification of prostate-specific genes.** *Int J Cancer* 2001, **92**:322-328.
24. Nelson PS, Clegg N, Arnold H, Ferguson C, Bonham M, White J, Hood L, Lin B: **The program of androgen-responsive genes in neoplastic prostate epithelium.** *Proc Natl Acad Sci U S A* 2002, **99**:11890-11895.
25. DePrimo SE, Diehn M, Nelson JB, Reiter RE, Matese J, Fero M, Tibshirani R, Brown PO, Brooks JD: **Transcriptional programs activated by exposure of human prostate cancer cells to androgen.** *Genome Biol* 2002, **3**:RESEARCH0032.

26. Xu LL, Shanmugam N, Segawa T, Sesterhenn IA, McLeod DG, Moul JW, Srivastava S: **A novel androgen-regulated gene, PMEPA1, located on chromosome 20q13 exhibits high level expression in prostate.** *Genomics* 2000, **66**:257-263.
27. Tusher VG, Tibshirani R, Chu G: **Significance analysis of microarrays applied to the ionizing radiation response.** *Proc Natl Acad Sci USA* 2001, **98**:5116-5121.
28. Iseli C, Jongeneel CV, Bucher P: **ESTScan: a program for detecting, evaluating, and reconstructing potential coding regions in EST sequences.** *Proc Int Conf Intell Syst Mol Biol* 1999:138-148.
29. Rodriguez A, Griffiths-Jones S, Ashurst JL, Bradley A: **Identification of mammalian microRNA host genes and transcription units.** *Genome Res* 2004, **14**:1902-1910.
30. Griffiths-Jones S: **The microRNA Registry.** *Nucleic Acids Res* 2004, **32**:D109-111.
31. Maere S, Heymans K, Kuiper M: **BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks.** *Bioinformatics* 2005, **21**:3448-3449.
32. Reis EM, Louro R, Nakaya HI, Verjovski-Almeida S: **As antisense RNA gets intronic.** *Omics* 2005, **9**:2-12.
33. Sainaghi PP, Castello L, Bergamasco L, Galletti M, Bellosta P, Avanzi GC: **Gas6 induces proliferation in prostate carcinoma cell lines expressing the Axl receptor.** *J Cell Physiol* 2005, **204**:36-44.
34. Petrovics G, Liu A, Shaheduzzaman S, Furasato B, Sun C, Chen Y, Nau M, Ravindranath L, Dobi A, Srikantan V, Sesterhenn IA, McLeod DG, Vahey M, Moul JW, Srivastava S: **Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome.** *Oncogene* 2005, **24**:3847-3852.
35. Kapranov P, Drenkow J, Cheng J, Long J, Helt G, Dike S, Gingeras TR: **Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays.** *Genome Res* 2005, **15**:987-997.
36. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ: **Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.** *Nucleic Acids Res* 1997, **25**:3389-3402.
37. Marchler-Bauer A, Anderson JB, Cherukuri PF, DeWeese-Scott C, Geer LY, Gwadz M, He S, Hurwitz DI, Jackson JD, Ke Z, Lanczycki CJ, Liebert CA, Liu C, Lu F, Marchler GH, Mullokandov M, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Yamashita RA, Yin JJ, Zhang D, Bryant SH: **CDD: a Conserved Domain Database for protein classification.** *Nucleic Acids Res* 2005, **33**:D192-196.
38. Kishore S, Stamm S: **The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C.** *Science* 2006, **311**:230-232.
39. Shaffer PL, Jivan A, Dollins DE, Claessens F, Gewirth DT: **Structural basis of androgen receptor binding to selective androgen response elements.** *Proc Natl Acad Sci U S A* 2004, **101**:4758-4763.
40. Euskirchen G, Royce TE, Bertone P, Martone R, Rinn JL, Nelson FK, Sayward F, Luscombe NM, Miller P, Gerstein M, Weissman S, Snyder M: **CREB binds to multiple loci on human chromosome 22.** *Mol Cell Biol* 2004, **24**:3804-3814.
41. Watanabe M, Nomura K, Ohyama A, Ishikawa R, Komiya Y, Hosaka K, Yamauchi E, Taniguchi H, Sasakawa N, Kumakura K, Ushiki T, Sato O, Ikebe M, Igarashi M: **Myosin-Va regulates exocytosis through the submicromolar Ca²⁺-dependent binding of syntaxin-1A.** *Mol Biol Cell* 2005, **16**:4519-4530.
42. Au JS, Huang JD: **A tissue-specific exon of myosin Va is responsible for selective cargo binding in melanocytes.** *Cell Motil Cytoskeleton* 2002, **53**:89-102.
43. Lee JT, Davidow LS, Warshawsky D: **Tsix, a gene antisense to Xist at the X-inactivation centre.** *Nat Genet* 1999, **21**:400-404.

44. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, Nishida H, Yap CC, Suzuki M, Kawai J, Suzuki H, Carninci P, Hayashizaki Y, Wells C, Frith M, Ravasi T, Pang KC, Hallinan J, Mattick J, Hume DA, Lipovich L, Batalov S, Engstrom PG, Mizuno Y, Faghihi MA, Sandelin A, Chalk AM, Mottagui-Tabar S, Liang Z, Lenhard B, Wahlestedt C: **Antisense transcription in the mammalian transcriptome.** *Science* 2005, **309**:1564-1566.
45. Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K, Gingeras TR: **Unbiased Mapping of Transcription Factor Binding Sites along Human Chromosomes 21 and 22 Points to Widespread Regulation of Noncoding RNAs.** *Cell* 2004, **116**:499-509.
46. Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, Fukuda S, Ru K, Frith MC, Gongora MM, Grimmond SM, Hume DA, Hayashizaki Y, Mattick JS: **Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome.** *Genome Res* 2006, **16**:11-19.
47. Johnson JM, Edwards S, Shoemaker D, Schadt EE: **Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments.** *Trends Genet* 2005, **21**:93-102.
48. Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ: **Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease.** *Biochemistry* 1979, **18**:5294-5299.
49. Dias-Neto E, Correa RG, Verjovski-Almeida S, Briones MR, Nagai MA, da Silva W, Jr., Zago MA, Bordin S, Costa FF, Goldman GH, Carvalho AF, Matsukuma A, Baia GS, Simpson DH, Brunstein A, de Oliveira PS, Bucher P, Jongeneel CV, O'Hare MJ, Soares F, Brentani RR, Reis LF, de Souza SJ, Simpson AJ: **Shotgun sequencing of the human transcriptome with ORF expressed sequence tags.** *Proc Natl Acad Sci USA* 2000, **97**:3491-3496.
50. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC, Gaasterland T, Glenisson P, Holstege FC, Kim IF, Markowitz V, Matese JC, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J, Vingron M: **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data.** *Nat Genet* 2001, **29**:365-371.
51. Peixoto BR, Vencio RZ, Egidio CM, Mota-Vieira L, Verjovski-Almeida S, Reis EM: **Evaluation of reference-based two-color methods for measurement of gene expression ratios using spotted cDNA microarrays.** *BMC Genomics* 2006, **7**:35.
52. Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR.** *Nucleic Acids Res* 2001, **29**:e45.
53. Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R, Hornischer K, Karas D, Kel AE, Kel-Margoulis OV, Kloos DU, Land S, Lewicki-Potapov B, Michael H, Munch R, Reuter I, Rotert S, Saxel H, Scheer M, Thiele S, Wingender E: **TRANSFAC: transcriptional regulation, from patterns to profiles.** *Nucleic Acids Res* 2003, **31**:374-378.
54. Chen H, Toyooka S, Gazdar AF, Hsieh JT: **Epigenetic regulation of a novel tumor suppressor gene (hDAB2IP) in prostate cancer cell lines.** *J Biol Chem* 2003, **278**:3121-3130.
55. Shim M, Eling TE: **Protein kinase C-dependent regulation of NAG-1/placental bone morphogenic protein/MIC-1 expression in LNCaP prostate carcinoma cells.** *J Biol Chem* 2005, **280**:18636-18642.

56. Yang CC, Ku CY, Wei S, Shiau CW, Chen CS, Pinzone JJ, Ringel MD, Chen CS: **Peroxisome proliferator-activated receptor gamma-independent repression of prostate-specific antigen expression by thiazolidinediones in prostate cancer cells.** *Mol Pharmacol* 2006, **69**:1564-1570.

Legends to the Figures

Figure 1. Androgen-responsive intronic RNAs. (A) Hierarchical clustering of 39 intronic transcripts significantly up or down-regulated following androgen exposure (B) Clusters of intronic transcripts (I–III) with similar temporal expression patterns are indicated. Average fold-change within each cluster in Panel B is indicated by a colored line. (C) Microarray expression pattern of 13 selected androgen-regulated intronic messages selected for orientation-specific reverse transcription PCR (OSRT-PCR). (D) Transcriptional orientation of 13 selected androgen-regulated intronic messages relative to the correspondent protein-coding mRNA determined by OSRT-PCR. Antisense (*GAS6*, *RAB25*, *ITGA6* and *ADD3*), sense (*ACTN4*, *STARD3* and *ERG*) or simultaneous antisense/sense (*DNAJC3*, *SAP18*, *KDEL2*, *FLJ10154*, *MYO5A* and *ATF2*) transcription (Right panel; Antisense, Sense +RT, + primer) was observed. Controls to monitor the absence of contamination with genomic DNA (Right panel; Control – RT, + primer) or the absence of RNA self-priming (Right panel; Control + RT, – primer) are indicated. Names between panels C and D refer to protein-coding RefSeq mRNAs that map on the same *loci* and define intronic regions where active transcription was detected.

Figure 2. Characterization of androgen-responsive non-coding intronic RNAs. Selected intronic transcripts extended by RACE-PCR (A–D, red boxes) or gene-specific PCR (E, yellow box) are shown. Extended unspliced antisense intronic transcript fragments were mapped to the genomic sequence (A –E, gray lines) relative to the respective spliced protein-coding transcript expressed in the opposite strand (A –E, green arrows). Double-stranded

cDNA clones spotted on the microarrays are represented as blue boxes (A–E). A previously described antisense unspliced intronic transcript mapping to *GAS6 locus* is shown (E, black arrow). Chromosome coordinates and the length of extended intronic transcript fragments are indicated. (F) Strand-specific multi-tissue Northern blot using probes complementary to the antisense intronic transcript mapped in the *GAS6 locus*

Figure 3. Antisense intronic transcription may regulate the abundance and the post-transcriptional processing of sense protein-coding mRNAs. (A - D) Expression levels of sense, antisense intronic transcripts (Panels “intronic”) or specific exons of protein-coding messages (Panels “exonic”) from *GAS6* (A), *DNAJC3* (B), *KDELR2* (C) or *ITGA6* (D) *loci* were measured in androgen-treated (black bars) or control cells (white bars). Full-length curated mRNAs mapping to each *loci* are shown as black arrows; non-curated mRNAs and ESTs are shown in gray arrows; androgen-regulated intronic unspliced transcripts are shown as hashed arrows. Relative fold-change (Y axis) indicates the increase or decrease in transcript levels after androgen-treatment relative to their abundance in untreated control cells. Error bars refer to standard deviations relative to mean fold-change values. Transcripts showing statistically significant differential expression ($p < 0.05$) are indicated with asterisks.

Figure 4. Increased abundance of antisense intronic MYO5A transcript correlates to binding of androgen receptor to a putative upstream androgen response element. (A) Expression level of antisense intronic transcript of *MYO5A* was measured in androgen-treated (black bars) or untreated control cells (white bars). Relative fold-change (Y axis) showed the increase or decrease in transcript levels after androgen-treatment relative to their abundance in untreated control cells. Error bars refer to standard deviations relative to mean fold-change values. Differential expression of sense and antisense *MYO5A* transcripts was statistically significant ($p < 0.05$) as indicated by asterisks. (B) Cellular lysates from androgen-treated (+)

or control cells (–) were used in chromatin immunoprecipitation (ChIP) assays in the presence of an antibody reactive with the androgen receptor protein (anti-AR). PCR reactions were used for detection of immunoprecipitated DNA containing an androgen response element (ARE) located upstream to the *MYO5A* antisense intronic transcript (*asMYO5A* ARE) in androgen-treated (+) or untreated control cells (–). Prostate Specific Antigen ARE (*PSA* ARE), or non-related genomic sequence upstream of tubulin alpha 6 gene (*TUBA6*) were used as positive and negative PCR controls, respectively. An unrelated antibody (mouse IgG) was used as a control for nonspecific immunoprecipitation. Equal volumes of lysates from androgen-treated or control cells (1/100 of total lysate) were treated in parallel, and used as loading controls of input DNA for each ChIP assay (input). (C) Enrichment ratios of *asMYO5A* ARE bound to AR in androgen-treated cells relative to control cells. Enrichment ratios (Raw) represent mean ratio values calculated from the integrated band intensities obtained at two different PCR cycle numbers (25 and 30 cycles), in reactions similar to those of Figure 4B, as described in *Methods*. Band intensities from input DNA were used to obtain normalized enrichment ratios (Normalized). Error bars show the standard deviation of mean ratios.

Additional files

Additional file 1: Supplementary Figure 1. Induction of androgen-responsive genes in prostate cells.

Additional file 2: Supplementary Figure 2. Temporal expression profile of androgen-responsive exonic RNAs.

Additional file 3: Supplementary Table 1 - Androgen-responsive exonic transcripts.

Additional file 4: Supplementary Table 2 - Androgen-responsive intronic transcripts.

Additional file 5: Supplementary Figure 3. *In silico* identification of ARE motifs in upstream regions of androgen-regulated intronic RNAs.

Additional file 6: MIAME file.

Figure 1

A*DNAJC3* (Chr13q32.1) **B***MYO5A* (Chr15q21.2) **C***ITGA6* (Chr2q31.1) **D***STARD13* (Chr13q13.1) **E***GAS6* (Chr13q34) **F** **Figure 2**

A **B** **C**

Figure 4

Additional files provided with this submission:

Additional file 6 : MIAME-revised.pdf : 204Kb

<http://www.biomedcentral.com/imedia/6108424381188038/sup6.PDF>

Additional file 5 : Supplementary Figure 3.pdf : 191Kb

<http://www.biomedcentral.com/imedia/1038587815118803/sup5.PDF>

Additional file 4 : Supplementary Table 2.pdf : 60Kb

<http://www.biomedcentral.com/imedia/2142089718114036/sup4.PDF>

Additional file 3 : Supplementary Table 1.pdf : 73Kb

<http://www.biomedcentral.com/imedia/6158271311140364/sup3.PDF>

Additional file 2 : Supplementary Figure 2.pdf : 147Kb

<http://www.biomedcentral.com/imedia/7250073211140364/sup2.PDF>

Additional file 1 : Supplementary Figure 1-revised.pdf : 191Kb

<http://www.biomedcentral.com/imedia/1123963286118802/sup1.PDF>

Tabela Suplementar 3 - Assinatura de expressão de transcritos intrônicos conservada em tecidos humanos e de camundongos.

Transcritos intrônicos com alterações significativas na expressão em pelo menos 1 dos tecidos (Próstata, Rim e Fígado) testados, indentificado por SAM, usando como parâmetros: resposta com múltiplas classes, 1.000 permutações, K-Nearest Neighbors Imputer, e taxa de falsa descoberta (FDR) < 0,1, e também por ANOVA usando-se o limite de significância de $p \leq 0,001$, de forma independente em humanos e camundongos. As colunas "Tecido com maior expressão" se refere ao resultado obtido pela análise de SAM.
N.D. - não disponível

Humano				Camundongo		
Nome do Gene	Gene ID	Anotação	Tecido com maior expressão	Nome do Gene	Gene ID	Tecido com maior expressão
LOC147650	147650	hypothetical protein	Fígado	LOC75202	75202	Fígado
<i>ASPH</i>	444	aspartate beta-hydroxylase	Fígado	<i>Asph</i>	65973	Fígado
<i>WWTR1</i>	25937	WW domain containing transcription regulator 1	Fígado	<i>Wwtr1</i>	97064	Fígado
<i>ETF1</i>	2107	eukaryotic translation termination factor 1	Fígado	<i>Etf1</i>	225363	Fígado
<i>RAB11FIP1</i>	80223	RAB11 family interacting protein 1	Fígado	4833414G05Rik	75767	Fígado
<i>PDXK</i>	8566	pyridoxal (pyridoxine, vitamin B6) kinase	Fígado	N.D.	N.D.	Fígado
<i>ITGA6</i>	3655	integrin, alpha 6	Fígado	<i>Itga6</i>	16403	Fígado
<i>CENTG2</i>	116987	centaurin, gamma 2	Rim	<i>Centg2</i>	347722	Rim
<i>ERG</i>	2078	v-ets erythroblastosis virus E26 oncogene like (avian); regulation of transcription, DNA-dependent	Fígado	<i>Erg</i>	13876	Fígado
<i>SAP18</i>	10284	Sin3A-associated protein	Fígado	<i>Sap18</i>	20220	Fígado
AJ563378		-	Rim	<i>Cyp3a13</i>	13113	Rim
C20orf118	140711	chromosome 20 open reading frame 118	Fígado	<i>Samhd1</i>	56045	Fígado
KIAA0319L	79932	a member of the polycystin protein family	Fígado	AU040320	100317	Fígado
BC021117		-	Fígado	AI314180	230249	Fígado
<i>HDAC4</i>	9759	histone deacetylase 4; regulation of transcription, DNA-dependent	Fígado	<i>Hdac4</i>	208727	Fígado
C11orf49 (MGC4707)	79096	chromosome 11 open reading frame 49	Próstata	1110051M20Rik	228356	Fígado
<i>INPP4A</i>	3631	inositol polyphosphate-4-phosphatase, type I	Fígado	N.D.	N.D.	Fígado

Tabela Suplementar 4 -Categorias de GO enriquecidas de transcritos intrônicos imunoprecipitados com RBPs

Transcritos intrônicos que foram imunoprecipitados associados com proteínas ligadoras de RNA (RBPs) foram selecionados e o enriquecimento de categorias de ontologia gênica (GO) em relação ao conjunto completo de mensagens intrônicas (intrônicos total) contidas no microarray foi calculado utilizando o programa eGOn (<http://www.genetools.microarray.ntnu.no/egon/>). Termos de GO com $p < 0.05$ são mostrados.

Intrônicos				
GO	Nome	Total	PABP	P-value
GO:0003674	molecular_function	497	145	
GO:0016491	oxidoreductase activity	15	8	0.046
GO:0004871	signal transducer activity	87	37	0.004
GO:0005070	SH3/SH2 adaptor activity	3	3	0.024
GO:0005211	transporter activity	50	22	0.021
GO:0008150	biological_process	472	139	
GO:0007165	signal transduction	125	46	0.040
GO:0000071	regulation of progression through cell cycle	21	11	0.026

Intrônicos				
GO	Nome	Total	HuD	P-value
GO:0003674	molecular_function	497	92	
GO:0016491	oxidoreductase activity	15	6	0.041
GO:0004871	signal transducer activity	87	25	0.010
GO:0005070	SH3/SH2 adaptor activity	3	3	0.006
GO:0005211	transporter activity	50	16	0.019
GO:0008150	biological_process	472	87	
GO:0000071	regulation of progression through cell cycle	21	8	0.037
GO:0005575	cellular_component	430	76	
GO:0019898	extrinsic to membrane	4	3	0.019
GO:0019899	extrinsic to plasma membrane	2	2	0.031

Intrônicos				
GO	Nome	Total	EIF4E	P-value
GO:0008150	biological_process	472	61	
GO:0006470	protein amino acid dephosphorylation	10	4	0.029
GO:0016311	dephosphorylation	10	4	0.029
GO:0000071	regulation of progression through cell cycle	21	6	0.041
GO:0005575	cellular_component	430	59	
GO:0031224	intrinsic to membrane	112	22	0.039
GO:0016021	integral to membrane	112	22	0.039
GO:0019899	extrinsic to plasma membrane	2	2	0.019

Intrônicos				
GO	Nome	Total	hTRA2	P-value
GO:0005575	cellular_component	430	75	
GO:0019899	extrinsic to plasma membrane	2	2	0.030

Tabela Suplementar 5 - Transcritos exônicos e intrônicos alterados pelo tratamento com alfa-amanitina.

Transcritos com alterações significativas na expressão em resposta ao tratamento com alfa-amanitina, identificados por SAM, usando como parâmetros: resposta de duas classes não pareadas, com 1.000 permutações, *K-Nearest Neighbors Imputer*, e taxa de falsa descoberta (FDR) de 0,002 a 0,035; e de *Signal-to-Noise Ratio* (SNR) com 10.000 permutações ($p < 0,05$).

Spot ID	Mapeamento	Gene ID	Nome do Locus	Anotação do Locus	Efeito alfa-amanitina
3110	Intrônico sTI	84441	MAML2	mastermind-like 2 (Drosophila)	Upregulated
3797	Intrônico sTI	3749	KCNC4	potassium voltage-gated channel, Shaw-related subfamily, member 4	Upregulated
4090	Intrônico sTI	7403	UTX	ubiquitously transcribed tetratricopeptide repeat, X chromosome	Upregulated
4425	Intrônico sTI	989	SEPT7	septin 7	Upregulated
4501	Intrônico sTI	7804	LRP8	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	Upregulated
7820	Intrônico sTI	2034	EPAS1	endothelial PAS domain protein 1	Upregulated
8782	Intrônico sTI	64754	SMYD3	SET and MYND domain containing 3	Upregulated
9298	Intrônico sTI	64097	EPB41L4A	erythrocyte membrane protein band 4.1 like 4A	Upregulated
9391	Intrônico sTI	89910	UBE3B	ubiquitin protein ligase E3B	Upregulated
9898	Intrônico sTI	219771	C10orf9	chromosome 10 open reading frame 9	Upregulated
10151	Intrônico sTI	10927	SPIN	spindlin	Upregulated
11159	Intrônico sTI	55626	-	-	Upregulated
11957	Intrônico sTI	254268	C1orf62	chromosome 1 open reading frame 62	Upregulated
12415	Intrônico sTI	22823	MTF2	metal response element binding transcription factor 2	Upregulated
12479	Intrônico sTI	114876	OSBPL1A	oxysterol binding protein-like 1A	Upregulated
13137	Intrônico sTI	11083	DATF1	death associated transcription factor 1	Upregulated
13312	Intrônico sTI	79720	-	-	Upregulated
15331	Intrônico sTI	4255	MGMT	O-6-methylguanine-DNA methyltransferase	Upregulated
15944	Intrônico sTI	5496	PPM1G	protein phosphatase 1G (formerly 2C), magnesium-dependent, gamma isoform	Upregulated
16847	Intrônico sTI	10608	MXD4	MAX dimerization protein 4	Upregulated
17140	Intrônico sTI	51341	ZBTB7A	zinc finger and BTB domain containing 7A	Upregulated
17219	Intrônico sTI	89801	PPP1R3F	protein phosphatase 1, regulatory (inhibitor) subunit 3F	Upregulated
17726	Intrônico sTI	220032	GDPD4	glycerophosphodiester phosphodiesterase domain containing 4	Upregulated
18380	Intrônico sTI	27032	ATP2C1	ATPase, Ca++ transporting, type 2C, member 1	Upregulated
20472	Intrônico sTI	6403	SELP	selectin P (granule membrane protein 140kDa, antigen CD62)	Upregulated
21508	Intrônico sTI	5095	PCCA	propionyl Coenzyme A carboxylase, alpha polypeptide	Upregulated

21764	Intrônico sTI	57185	-	-	Upregulated
22435	Intrônico sTI	5046	PCSK6	proprotein convertase subtilisin/kexin type 6	Upregulated
22479	Intrônico sTI	79823	C2orf34	chromosome 2 open reading frame 34	Upregulated
23986	Intrônico sTI	3607	FOXK2	forkhead box K2	Upregulated
24068	Intrônico sTI	8635	RNASET2	ribonuclease T2	Upregulated
24586	Intrônico sTI	7936	RDBP	RD RNA binding protein	Upregulated
26912	Intrônico sTI	57706	KIAA1608	KIAA1608	Upregulated
26943	Intrônico sTI	10217	CTDSPL	CTD (carboxy-terminal domain, RNA polymerase II polypeptide A) small phosphatase-like	Upregulated
27073	Intrônico sTI	4292	MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	Upregulated
27454	Intrônico sTI	11177	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	Upregulated
27481	Intrônico sTI	5298	PIK4CB	phosphatidylinositol 4-kinase, catalytic, beta polypeptide	Upregulated
28464	Intrônico sTI	55370	PPP4R1L	protein phosphatase 4, regulatory subunit 1-like	Upregulated
28494	Intrônico sTI	9444	QKI	quaking homolog, KH domain RNA binding (mouse)	Upregulated
30975	Intrônico sTI	7072	TIA1	TIA1 cytotoxic granule-associated RNA binding protein	Upregulated
31464	Intrônico sTI	23512	SUZ12	suppressor of zeste 12 homolog (Drosophila)	Upregulated
31995	Intrônico sTI	9630	GNA14	guanine nucleotide binding protein (G protein), alpha 14	Upregulated
32225	Intrônico sTI	3837	KPNB1	karyopherin (importin) beta 1	Upregulated
33157	Intrônico sTI	2132	EXT2	exostoses (multiple) 2	Upregulated
33974	Intrônico sTI	4149	MAX	MYC associated factor X	Upregulated
34068	Intrônico sTI	6446	SGK	serum/glucocorticoid regulated kinase	Upregulated
34400	Intrônico sTI	444	ASPH	aspartate beta-hydroxylase	Upregulated
36031	Intrônico sTI	4976	OPA1	optic atrophy 1 (autosomal dominant)	Upregulated
36503	Intrônico sTI	51701	NLK	nemo like kinase	Upregulated
36519	Intrônico sTI	11010	GLIPR1	GLI pathogenesis-related 1 (glioma)	Upregulated
37593	Intrônico sTI	636	BICD1	bicaudal D homolog 1 (Drosophila)	Upregulated
39037	Intrônico sTI	164	AP1G1	adaptor-related protein complex 1, gamma 1 subunit	Upregulated
40159	Intrônico sTI	2762	GMDS	GDP-mannose 4,6-dehydratase	Upregulated
40462	Intrônico sTI	4705	NDUFA10	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 10, 42kDa	Upregulated
40491	Intrônico sTI	151162	-	-	Upregulated
41080	Intrônico sTI	4706	NDUFAB1	NADH dehydrogenase (ubiquinone) 1, alpha/beta subcomplex, 1, 8kDa	Upregulated
41523	Intrônico sTI	128869	CDC91L1	CDC91 cell division cycle 91-like 1 (S. cerevisiae)	Upregulated
41883	Intrônico sTI	9644	SH3MD1	SH3 multiple domains 1	Upregulated
44204	Intrônico sTI	91523	-	-	Upregulated
29	Intrônico sTI	58155	PTBP2	polypyrimidine tract binding protein 2	Downregulated

57	Intrônico sTI	10905	MAN1A2	mannosidase, alpha, class 1A, member 2	Downregulated
284	Intrônico sTI	6310	ATXN1	ataxin 1	Downregulated
317	Intrônico sTI	284001	-	-	Downregulated
361	Intrônico sTI	23195	MDN1	MDN1, midasin homolog (yeast)	Downregulated
598	Intrônico sTI	5218	PFTK1	PFTAIR protein kinase 1	Downregulated
662	Intrônico sTI	29880	ALG5	asparagine-linked glycosylation 5 homolog (yeast, dolichyl-phosphate beta-glucosyltransferase)	Downregulated
670	Intrônico sTI	23196	C9orf10	chromosome 9 open reading frame 10	Downregulated
673	Intrônico sTI	22794	CASC3	cancer susceptibility candidate 3	Downregulated
709	Intrônico sTI	3096	HIVEP1	human immunodeficiency virus type I enhancer binding protein 1	Downregulated
773	Intrônico sTI	23195	MDN1	MDN1, midasin homolog (yeast)	Downregulated
796	Intrônico sTI	7024	TFCP2	transcription factor CP2	Downregulated
869	Intrônico sTI	27332	ZNF638	zinc finger protein 638	Downregulated
998	Intrônico sTI	8287	USP9Y	ubiquitin specific protease 9, Y-linked (fat facets-like, Drosophila)	Downregulated
1019	Intrônico sTI	2869	GRK5	G protein-coupled receptor kinase 5	Downregulated
1054	Intrônico sTI	79987	SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	Downregulated
1056	Intrônico sTI	79086	-	-	Downregulated
1061	Intrônico sTI	57325	CSRP2BP	CSRP2 binding protein	Downregulated
1203	Intrônico sTI	8525	DGKZ	diacylglycerol kinase, zeta 104kDa	Downregulated
1256	Intrônico sTI	84441	MAML2	mastermind-like 2 (Drosophila)	Downregulated
1274	Intrônico sTI	54919	-	-	Downregulated
1620	Intrônico sTI	7024	TFCP2	transcription factor CP2	Downregulated
1629	Intrônico sTI	5069	PAPPA	pregnancy-associated plasma protein A, pappalysin 1	Downregulated
1637	Intrônico sTI	2869	GRK5	G protein-coupled receptor kinase 5	Downregulated
1686	Intrônico sTI	54919	-	-	Downregulated
1772	Intrônico sTI	84062	DTNBP1	dystrobrevin binding protein 1	Downregulated
1773	Intrônico sTI	81619	TSPAN14	tetraspanin 14	Downregulated
1819	Intrônico sTI	8869	ST3GAL5	ST3 beta-galactoside alpha 2,3-sialyltransferase 5	Downregulated
1847	Intrônico sTI	1793	DOCK1	dedicator of cytokinesis 1	Downregulated
1865	Intrônico sTI	170959	ZNF431	zinc finger protein 431	Downregulated
1926	Intrônico sTI	7849	PAX8	paired box gene 8	Downregulated
2013	Intrônico sTI	11107	PRDM5	PR domain containing 5	Downregulated
2032	Intrônico sTI	7024	TFCP2	transcription factor CP2	Downregulated
2304	Intrônico sTI	54927	CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing 3	Downregulated
2407	Intrônico sTI	54914	KIAA1797	KIAA1797	Downregulated
2423	Intrônico sTI	22982	KIAA0934	KIAA0934	Downregulated
2489	Intrônico sTI	93145	OLFM2	olfactomedin 2	Downregulated
2491	Intrônico sTI	84909	C9orf3	chromosome 9 open reading frame 3	Downregulated
2594	Intrônico sTI	84898	PLXDC2	plexin domain containing 2	Downregulated

2604	Intrônico sTI	57826	RAP2C	RAP2C, member of RAS oncogene family	Downregulated
2689	Intrônico sTI	170959	ZNF431	zinc finger protein 431	Downregulated
2691	Intrônico sTI	146174	C16orf52	chromosome 16 open reading frame 52	Downregulated
2712	Intrônico sTI	55748	CNDP2	CNDP dipeptidase 2 (metallopeptidase M20 family)	Downregulated
2900	Intrônico sTI	115294	-	-	Downregulated
3010	Intrônico sTI	80262	-	-	Downregulated
3049	Intrônico sTI	9962	SLC23A2	solute carrier family 23 (nucleobase transporters), member 2	Downregulated
3068	Intrônico sTI	5607	MAP2K5	mitogen-activated protein kinase kinase 5	Downregulated
3124	Intrônico sTI	55748	CNDP2	CNDP dipeptidase 2 (metallopeptidase M20 family)	Downregulated
3314	Intrônico sTI	90203	C20orf161	chromosome 20 open reading frame 161	Downregulated
3489	Intrônico sTI	3480	IGF1R	insulin-like growth factor 1 receptor	Downregulated
3723	Intrônico sTI	124460	-	-	Downregulated
3776	Intrônico sTI	8883	APPBP1	amyloid beta precursor protein binding protein 1	Downregulated
3900	Intrônico sTI	3716	JAK1	Janus kinase 1 (a protein tyrosine kinase)	Downregulated
3958	Intrônico sTI	29894	CPSF1	cleavage and polyadenylation specific factor 1, 160kDa	Downregulated
4136	Intrônico sTI	115350	FCRL1	Fc receptor-like 1	Downregulated
4145	Intrônico sTI	79753	SNIP1	Smad nuclear interacting protein 1	Downregulated
4170	Intrônico sTI	23389	THRAP2	thyroid hormone receptor associated protein 2	Downregulated
4448	Intrônico sTI	84894	LRRN6A	leucine rich repeat neuronal 6A	Downregulated
4694	Intrônico sTI	10466	COG5	component of oligomeric golgi complex 5	Downregulated
4915	Intrônico sTI	7170	TPM3	tropomyosin 3	Downregulated
4973	Intrônico sTI	58483	C9orf27	chromosome 9 open reading frame 27	Downregulated
5020	Intrônico sTI	6788	STK3	serine/threonine kinase 3 (STE20 homolog, yeast)	Downregulated
5391	Intrônico sTI	55616	DDEFL1	development and differentiation enhancing factor-like 1	Downregulated
5414	Intrônico sTI	10651	MTX2	metaxin 2	Downregulated
5476	Intrônico sTI	92822	ZFP276	zinc finger protein 276 homolog (mouse)	Downregulated
5595	Intrônico sTI	55909	BIN3	bridging integrator 3	Downregulated
5605	Intrônico sTI	51053	GMNN	geminin, DNA replication inhibitor	Downregulated
5654	Intrônico sTI	2887	GRB10	growth factor receptor-bound protein 10	Downregulated
5740	Intrônico sTI	7009	TEGT	testis enhanced gene transcript (BAX inhibitor 1)	Downregulated
5764	Intrônico sTI	965	CD58	CD58 antigen, (lymphocyte function-associated antigen 3)	Downregulated
5808	Intrônico sTI	54084	C21orf29	chromosome 21 open reading frame 29	Downregulated
6175	Intrônico sTI	1282	COL4A1	collagen, type IV, alpha 1	Downregulated
6267	Intrônico sTI	4232	MEST	mesoderm specific transcript homolog (mouse)	Downregulated
6336	Intrônico sTI	23080	-	-	Downregulated
6345	Intrônico sTI	9961	MVP	major vault protein	Downregulated

6428	Intrônico sTI	51312	SLC25A37	solute carrier family 25, member 37	Downregulated
6454	Intrônico sTI	8887	TAX1BP1	Tax1 (human T-cell leukemia virus type I) binding protein 1	Downregulated
6476	Intrônico sTI	3490	IGFBP7	insulin-like growth factor binding protein 7	Downregulated
6864	Intrônico sTI	9373	PLAA	phospholipase A2- activating protein	Downregulated
7266	Intrônico sTI	11124	FAF1	Fas (TNFRSF6) associated factor 1	Downregulated
7451	Intrônico sTI	55619	DOCK10	dedicator of cytokinesis 10	Downregulated
7466	Intrônico sTI	23392	KIAA0368	KIAA0368	Downregulated
7516	Intrônico sTI	776	CACNA1D	calcium channel, voltage- dependent, L type, alpha 1D subunit	Downregulated
7545	Intrônico sTI	79068	-	-	Downregulated
7558	Intrônico sTI	54733	SLC35F2	solute carrier family 35, member F2	Downregulated
7581	Intrônico sTI	9958	USP15	ubiquitin specific protease 15	Downregulated
7982	Intrônico sTI	23273	KIAA0367	KIAA0367	Downregulated
8053	Intrônico sTI	84910	-	-	Downregulated
8180	Intrônico sTI	50863	-	-	Downregulated
8392	Intrônico sTI	23567	ZNF346	zinc finger protein 346	Downregulated
8436	Intrônico sTI	2562	GABRB3	gamma-aminobutyric acid (GABA) A receptor, beta 3	Downregulated
8439	Intrônico sTI	1785	DNM2	dynamain 2	Downregulated
8492	Intrônico sTI	26974	ZNF285	zinc finger protein 285	Downregulated
8497	Intrônico sTI	23294	ANKS1	ankyrin repeat and sterile alpha motif domain containing 1	Downregulated
8546	Intrônico sTI	776	CACNA1D	calcium channel, voltage- dependent, L type, alpha 1D subunit	Downregulated
8581	Intrônico sTI	56889	-	-	Downregulated
8744	Intrônico sTI	2889	RAPGEF1	Rap guanine nucleotide exchange factor (GEF) 1	Downregulated
8776	Intrônico sTI	83989	-	-	Downregulated
8795	Intrônico sTI	53615	MBD3	methyl-CpG binding domain protein 3	Downregulated
8844	Intrônico sTI	3709	ITPR2	inositol 1,4,5-triphosphate receptor, type 2	Downregulated
8851	Intrônico sTI	1785	DNM2	dynamain 2	Downregulated
9211	Intrônico sTI	29767	TMOD2	tropomodulin 2 (neuronal)	Downregulated
9484	Intrônico sTI	284124	-	-	Downregulated
9610	Intrônico sTI	57214	KIAA1199	KIAA1199	Downregulated
10035	Intrônico sTI	29766	TMOD3	tropomodulin 3 (ubiquitous)	Downregulated
10039	Intrônico sTI	25829	C22orf5	chromosome 22 open reading frame 5	Downregulated
10435	Intrônico sTI	56886	UGCG1	UDP-glucose ceramide glucosyltransferase-like 1	Downregulated
10481	Intrônico sTI	6284	S100A13	S100 calcium binding protein A13	Downregulated
11054	Intrônico sTI	55869	HDAC8	histone deacetylase 8	Downregulated
11203	Intrônico sTI	5998	RGS3	regulator of G-protein signalling 3	Downregulated
11480	Intrônico sTI	26019	UPF2	UPF2 regulator of nonsense transcripts homolog (yeast)	Downregulated
11535	Intrônico sTI	287	ANK2	ankyrin 2, neuronal	Downregulated
11571	Intrônico sTI	55626	-	-	Downregulated
11615	Intrônico sTI	5998	RGS3	regulator of G-protein signalling 3	Downregulated

11630	Intrônico sTI	2222	FDFT1	farnesyl-diphosphate farnesyltransferase 1	Downregulated
11948	Intrônico sTI	55	ACPP	acid phosphatase, prostate	Downregulated
11985	Intrônico sTI	55172	C14orf104	chromosome 14 open reading frame 104	Downregulated
12121	Intrônico sTI	8499	PPFIA2	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 2	Downregulated
12342	Intrônico sTI	4784	NFIX	nuclear factor I/X (CCAAT- binding transcription factor)	Downregulated
12601	Intrônico sTI	55627	-	-	Downregulated
12654	Intrônico sTI	4036	LRP2	low density lipoprotein- related protein 2	Downregulated
13313	Intrônico sTI	79033	PRNP	prion protein interacting protein	Downregulated
13474	Intrônico sTI	5087	PBX1	pre-B-cell leukemia transcription factor 1	Downregulated
13533	Intrônico sTI	53354	PANK1	pantothenate kinase 1	Downregulated
13794	Intrônico sTI	2023	ENO1	enolase 1, (alpha)	Downregulated
14199	Intrônico sTI	3955	LFNG	lunatic fringe homolog (Drosophila)	Downregulated
14204	Intrônico sTI	2548	GAA	glucosidase, alpha; acid (Pompe disease, glycogen storage disease type II)	Downregulated
14281	Intrônico sTI	9159	PCSK7	proprotein convertase subtilisin/kexin type 7	Downregulated
14399	Intrônico sTI	5420	PODXL	podocalyxin-like	Downregulated
14550	Intrônico sTI	64746	ACBD3	acyl-Coenzyme A binding domain containing 3	Downregulated
14655	Intrônico sTI	58508	MLL3	myeloid/lymphoid or mixed- lineage leukemia 3	Downregulated
14710	Intrônico sTI	5087	PBX1	pre-B-cell leukemia transcription factor 1	Downregulated
15066	Intrônico sTI	64131	XYLT1	xylosyltransferase I	Downregulated
15087	Intrônico sTI	23633	KPNA6	karyopherin alpha 6 (importin alpha 7)	Downregulated
15379	Intrônico sTI	56681	SARA1	SAR1a gene homolog 1 (S. cerevisiae)	Downregulated
15413	Intrônico sTI	9320	TRIP12	thyroid hormone receptor interactor 12	Downregulated
15789	Intrônico sTI	57553	MICAL3	microtubule associated monooxygenase, calponin and LIM domain containing 3	Downregulated
15998	Intrônico sTI	55861	C20orf35	chromosome 20 open reading frame 35	Downregulated
16266	Intrônico sTI	2017	CTTN	cortactin	Downregulated
16514	Intrônico sTI	55763	SEC3L1	SEC3-like 1 (S. cerevisiae)	Downregulated
16650	Intrônico sTI	9101	USP8	ubiquitin specific protease 8	Downregulated
16819	Intrônico sTI	57552	AADACL1	arylacetamide deacetylase- like 1	Downregulated
17044	Intrônico sTI	23527	CENTB2	centaurin, beta 2	Downregulated
17108	Intrônico sTI	220002	CYBASC3	cytochrome b, ascorbate dependent 3	Downregulated
17263	Intrônico sTI	9943	OXS1	oxidative-stress responsive 1	Downregulated
17689	Intrônico sTI	6733	SRPK2	SFRS protein kinase 2	Downregulated
17692	Intrônico sTI	5921	RASA1	RAS p21 protein activator (GTPase activating protein) 1	Downregulated

17764	Intrônico sTI	25896	C1orf73	chromosome 1 open reading frame 73	Downregulated
17890	Intrônico sTI	8148	TAF15	TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 68kDa	Downregulated
18071	Intrônico sTI	26297	-	-	Downregulated
18400	Intrônico sTI	9403	-	-	Downregulated
18521	Intrônico sTI	5048	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit 45kDa	Downregulated
18556	Intrônico sTI	116159	CYYR1	cysteine and tyrosine-rich 1	Downregulated
18820	Intrônico sTI	7204	TRIO	triple functional domain (PTPRF interacting) disabled homolog 2,	Downregulated
18843	Intrônico sTI	1601	DAB2	mitogen-responsive phosphoprotein (Drosophila)	Downregulated
18924	Intrônico sTI	6942	TCF20	transcription factor 20 (AR1)	Downregulated
19068	Intrônico sTI	143879	KBTBD3	kelch repeat and BTB (POZ) domain containing 3	Downregulated
19168	Intrônico sTI	220107	DLEU7	deleted in lymphocytic leukemia, 7	Downregulated
19271	Intrônico sTI	201627	-	-	Downregulated
19297	Intrônico sTI	55275	-	-	Downregulated
19336	Intrônico sTI	6942	TCF20	transcription factor 20 (AR1)	Downregulated
19377	Intrônico sTI	147949	ZNF583	zinc finger protein 583	Downregulated
19771	Intrônico sTI	1198	CLK3	CDC-like kinase 3	Downregulated
19930	Intrônico sTI	23263	MCF2L	MCF.2 cell line derived transforming sequence-like	Downregulated
20365	Intrônico sTI	7145	TNS1	tensin 1	Downregulated
20385	Intrônico sTI	2170	FABP3	fatty acid binding protein 3, muscle and heart (mammary-derived growth inhibitor)	Downregulated
20490	Intrônico sTI	1847	DUSP5	dual specificity phosphatase 5	Downregulated
20752	Intrônico sTI	23522	MYST4	MYST histone acetyltransferase (monocytic leukemia) 4	Downregulated
20879	Intrônico sTI	7447	VSNL1	visinin-like 1	Downregulated
20981	Intrônico sTI	7704	ZBTB16	zinc finger and BTB domain containing 16	Downregulated
21008	Intrônico sTI	928	CD9	CD9 antigen (p24)	Downregulated
21059	Intrônico sTI	26074	C20orf26	chromosome 20 open reading frame 26	Downregulated
21256	Intrônico sTI	55003	PAK1IP1	PAK1 interacting protein 1	Downregulated
21257	Intrônico sTI	54805	CNNM2	cyclin M2	Downregulated
21393	Intrônico sTI	7704	ZBTB16	zinc finger and BTB domain containing 16	Downregulated
21401	Intrônico sTI	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	Downregulated
21609	Intrônico sTI	5378	PMS1	PMS1 postmeiotic segregation increased 1 (S. cerevisiae)	Downregulated
21618	Intrônico sTI	3064	HD	huntingtin (Huntington disease)	Downregulated
21622	Intrônico sTI	2011	MARK2	MAP/microtubule affinity-regulating kinase 2	Downregulated
21667	Intrônico sTI	55187	VPS13D	vacuolar protein sorting 13D (yeast)	Downregulated
21674	Intrônico sTI	29960	FTSJ2	FtsJ homolog 2 (E. coli)	Downregulated

21740	Intrônico sTI	337867	PHGDHL1	phosphoglycerate dehydrogenase like 1	Downregulated
21778	Intrônico sTI	27183	VPS4A	vacuolar protein sorting 4A (yeast)	Downregulated
21796	Intrônico sTI	9798	-	-	Downregulated
22020	Intrônico sTI	5581	PRKCE	protein kinase C, epsilon	Downregulated
22138	Intrônico sTI	1859	DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	Downregulated
22267	Intrônico sTI	84961	FBXL20	F-box and leucine-rich repeat protein 20	Downregulated
22445	Intrônico sTI	2162	F13A1	coagulation factor XIII, A1 polypeptide	Downregulated
22593	Intrônico sTI	55269	PSPC1	paraspeckle component 1	Downregulated
22597	Intrônico sTI	51773	HBXAP	hepatitis B virus x associated protein	Downregulated
22600	Intrônico sTI	50807	DDEF1	development and differentiation enhancing factor 1	Downregulated
22604	Intrônico sTI	26002	MOXD1	monooxygenase, DBH-like 1	Downregulated
22659	Intrônico sTI	238	ALK	anaplastic lymphoma kinase (Ki-1)	Downregulated
22809	Intrônico sTI	26276	VPS33B	vacuolar protein sorting 33B (yeast)	Downregulated
22823	Intrônico sTI	10233	-	-	Downregulated
22962	Intrônico sTI	1859	DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	Downregulated
23036	Intrônico sTI	9096	TBX18	T-box 18	Downregulated
23042	Intrônico sTI	7363	UGT2B4	UDP glycosyltransferase 2 family, polypeptide B4	Downregulated
23107	Intrônico sTI	55651	NOLA2	nucleolar protein family A, member 2 (H/ACA small nucleolar RNPs)	Downregulated
23156	Intrônico sTI	5095	PCCA	propionyl Coenzyme A carboxylase, alpha polypeptide	Downregulated
23412	Intrônico sTI	57178	RAI17	retinoic acid induced 17	Downregulated
23423	Intrônico sTI	51222	ZNF219	zinc finger protein 219	Downregulated
23713	Intrônico sTI	80380	PDCD1LG2	programmed cell death 1 ligand 2	Downregulated
23829	Intrônico sTI	55268	ECHDC2	enoyl Coenzyme A hydratase domain containing 2	Downregulated
23831	Intrônico sTI	54875	C9orf39	chromosome 9 open reading frame 39	Downregulated
24071	Intrônico sTI	7700	ZNF141	zinc finger protein 141 (clone pHZ-44)	Downregulated
24465	Intrônico sTI	22920	KIFAP3	kinesin-associated protein 3	Downregulated
24468	Intrônico sTI	10815	CPLX1	complexin 1	Downregulated
24546	Intrônico sTI	56997	CABC1	chaperone, ABC1 activity of bc1 complex like (S. pombe)	Downregulated
24871	Intrônico sTI	25792	CIZ1	CDKN1A interacting zinc finger protein 1	Downregulated
24895	Intrônico sTI	7678	ZNF124	zinc finger protein 124 (HZF 16)	Downregulated
24955	Intrônico sTI	59271	C21orf63	chromosome 21 open reading frame 63	Downregulated
24966	Intrônico sTI	54453	RIN2	Ras and Rab interactor 2	Downregulated
25210	Intrônico sTI	6429	SFRS4	splicing factor, arginine/serine-rich 4	Downregulated
25696	Intrônico sTI	23514	-	-	Downregulated
25980	Intrônico sTI	80036	TRPM3	transient receptor potential cation channel, subfamily M, member 3	Downregulated

26007	Intrônico sTI	23313	C22orf9	chromosome 22 open reading frame 9	Downregulated
26073	Intrônico sTI	130340	AP1S3	adaptor-related protein complex 1, sigma 3 subunit	Downregulated
26074	Intrônico sTI	121260	SLC15A4	solute carrier family 15, member 4	Downregulated
26113	Intrônico sTI	22913	RALY	RNA binding protein (autoantigenic, hnRNP-associated with lethal yellow)	Downregulated
26118	Intrônico sTI	10426	TUBGCP3	tubulin, gamma complex associated protein 3	Downregulated
26185	Intrônico sTI	80705	TSGA10	testis specific, 10	Downregulated
26258	Intrônico sTI	1870	E2F2	E2F transcription factor 2	Downregulated
26273	Intrônico sTI	283349	RASSF3	Ras association (RalGDS/AF-6) domain family 3	Downregulated
26287	Intrônico sTI	81532	-	-	Downregulated
26293	Intrônico sTI	60685	TEX27	testis expressed sequence 27	Downregulated
26295	Intrônico sTI	57521	-	-	Downregulated
26356	Intrônico sTI	3055	HCK	hemopoietic cell kinase	Downregulated
26379	Intrônico sTI	196463	-	-	Downregulated
26402	Intrônico sTI	55784	-	-	Downregulated
26481	Intrônico sTI	201266	SLC39A11	solute carrier family 39 (metal ion transporter), member 11	Downregulated
26500	Intrônico sTI	57706	KIAA1608	KIAA1608	Downregulated
26511	Intrônico sTI	51762	RAB8B	RAB8B, member RAS oncogene family	Downregulated
26546	Intrônico sTI	6934	TCF7L2	transcription factor 7-like 2 (T-cell specific, HMG-box)	Downregulated
26806	Intrônico sTI	79591	C10orf76	chromosome 10 open reading frame 76	Downregulated
26835	Intrônico sTI	22859	LPHN1	latrophilin 1	Downregulated
26891	Intrônico sTI	283349	RASSF3	Ras association (RalGDS/AF-6) domain family 3	Downregulated
26898	Intrônico sTI	121260	SLC15A4	solute carrier family 15, member 4	Downregulated
27030	Intrônico sTI	29992	PILRA	paired immunoglobulin-like type 2 receptor alpha	Downregulated
27032	Intrônico sTI	27044	-	-	Downregulated
27105	Intrônico sTI	114793	FMNL2	formin-like 2	Downregulated
27229	Intrônico sTI	55193	-	-	Downregulated
27344	Intrônico sTI	23514	-	-	Downregulated
27360	Intrônico sTI	9501	RPH3AL	rabphilin 3A-like (without C2 domains)	Downregulated
27438	Intrônico sTI	54454	-	-	Downregulated
27499	Intrônico sTI	607	BCL9	B-cell CLL/lymphoma 9	Downregulated
27527	Intrônico sTI	66005	-	-	Downregulated
27531	Intrônico sTI	57521	-	-	Downregulated
28135	Intrônico sTI	114793	FMNL2	formin-like 2	Downregulated
28150	Intrônico sTI	57157	PHTF2	putative homeodomain transcription factor 2	Downregulated
28249	Intrônico sTI	64855	C9orf88	chromosome 9 open reading frame 88	Downregulated
28252	Intrônico sTI	57620	STIM2	stromal interaction molecule 2	Downregulated
28384	Intrônico sTI	10425	ARIH2	ariadne homolog 2 (Drosophila)	Downregulated
28435	Intrônico sTI	375260	-	-	Downregulated
28452	Intrônico sTI	80036	TRPM3	transient receptor potential cation channel, subfamily M, member 3	Downregulated
28580	Intrônico sTI	23514	-	-	Downregulated

28687	Intrônico sTI	23112	TNRC6B	trinucleotide repeat containing 6B	Downregulated
28772	Intrônico sTI	55556	ENOSF1	enolase superfamily member 1	Downregulated
28773	Intrônico sTI	55254	TMEM39A	transmembrane protein 39A	Downregulated
28886	Intrônico sTI	27044	-	-	Downregulated
29080	Intrônico sTI	55788	C6orf209	chromosome 6 open reading frame 209	Downregulated
29166	Intrônico sTI	92105	-	-	Downregulated
29186	Intrônico sTI	55083	-	-	Downregulated
29260	Intrônico sTI	284996	RNF149	ring finger protein 149	Downregulated
29262	Intrônico sTI	221037	JMJD1C	jumonji domain containing 1C	Downregulated
29298	Intrônico sTI	27044	-	-	Downregulated
29306	Intrônico sTI	23024	PDZRN3	PDZ domain containing RING finger 3	Downregulated
29361	Intrônico sTI	399474	-	-	Downregulated
29521	Intrônico sTI	9887	C1orf16	chromosome 1 open reading frame 16	Downregulated
29686	Intrônico sTI	83605	CCM2	cerebral cavernous malformation 2	Downregulated
29794	Intrônico sTI	64430	C14orf135	chromosome 14 open reading frame 135	Downregulated
29846	Intrônico sTI	5903	RANBP2	RAN binding protein 2	Downregulated
29904	Intrônico sTI	55788	C6orf209	chromosome 6 open reading frame 209	Downregulated
29994	Intrônico sTI	83734	APG10L	APG10 autophagy 10-like (S. cerevisiae)	Downregulated
30012	Intrônico sTI	54536	SEC15L1	SEC15-like 1 (S. cerevisiae)	Downregulated
30094	Intrônico sTI	91074	ANKRD30A	ankyrin repeat domain 30A	Downregulated
30097	Intrônico sTI	84186	ZCCHC7	zinc finger, CCHC domain containing 7	Downregulated
30191	Intrônico sTI	151963	-	-	Downregulated
30302	Intrônico sTI	84629	-	-	Downregulated
30319	Intrônico sTI	55198	-	-	Downregulated
30398	Intrônico sTI	143187	VTI1A	vesicle transport through interaction with t-SNAREs homolog 1A (yeast)	Downregulated
30402	Intrônico sTI	92017	-	-	Downregulated
30425	Intrônico sTI	51742	ARID4B	AT rich interactive domain 4B (RBP1- like)	Downregulated
30445	Intrônico sTI	10207	INADL	InaD-like (Drosophila)	Downregulated
30488	Intrônico sTI	25	ABL1	v-abl Abelson murine leukemia viral oncogene homolog 1	Downregulated
30502	Intrônico sTI	138046	-	-	Downregulated
30719	Intrônico sTI	79848	-	-	Downregulated
30726	Intrônico sTI	57020	-	-	Downregulated
30747	Intrônico sTI	23112	TNRC6B	trinucleotide repeat containing 6B	Downregulated
30828	Intrônico sTI	57154	SMURF1	SMAD specific E3 ubiquitin protein ligase 1	Downregulated
30869	Intrônico sTI	7627	ZNF75A	zinc finger protein 75a	Downregulated
31039	Intrônico sTI	55252	ASXL2	additional sex combs like 2 (Drosophila)	Downregulated
31059	Intrônico sTI	11052	CPSF6	cleavage and polyadenylation specific factor 6, 68kDa	Downregulated
31119	Intrônico sTI	149076	-	-	Downregulated
31120	Intrônico sTI	138046	-	-	Downregulated
31196	Intrônico sTI	3632	INPP5A	inositol polyphosphate-5-phosphatase, 40kDa	Downregulated
31349	Intrônico sTI	55198	-	-	Downregulated
31466	Intrônico sTI	23254	-	-	Downregulated
31529	Intrônico sTI	196527	TMEM16F	transmembrane protein 16F	Downregulated

31634	Intrônico sTI	143187	VTI1A	vesicle transport through interaction with t-SNAREs homolog 1A (yeast)	Downregulated
31744	Intrônico sTI	84629	-	-	Downregulated
31790	Intrônico sTI	9445	ITM2B	integral membrane protein 2B	Downregulated
32040	Intrônico sTI	286053	-	-	Downregulated
32087	Intrônico sTI	22906	-	-	Downregulated
32178	Intrônico sTI	51379	CRLF3	cytokine receptor-like factor 3	Downregulated
32210	Intrônico sTI	7267	TTC3	tetratricopeptide repeat domain 3	Downregulated
32466	Intrônico sTI	83716	CRISPLD2	cysteine-rich secretory protein LCCL domain containing 2	Downregulated
32491	Intrônico sTI	26235	FBXL4	F-box and leucine-rich repeat protein 4	Downregulated
32715	Intrônico sTI	9657	IQCB1	IQ motif containing B1	Downregulated
32786	Intrônico sTI	57038	RARSL	arginyl-tRNA synthetase-like	Downregulated
32977	Intrônico sTI	113174	-	-	Downregulated
32981	Intrônico sTI	84187	-	-	Downregulated
32986	Intrônico sTI	79618	-	-	Downregulated
33182	Intrônico sTI	116987	CENTG2	centaurin, gamma 2	Downregulated
33321	Intrônico sTI	23145	-	-	Downregulated
33379	Intrônico sTI	378805	-	-	Downregulated
33496	Intrônico sTI	83715	ESPN	espin	Downregulated
33602	Intrônico sTI	80108	ZFP2	zinc finger protein 2 homolog (mouse)	Downregulated
33710	Intrônico sTI	57688	ZSWIM6	zinc finger, SWIM-type containing 6	Downregulated
33719	Intrônico sTI	54845	-	-	Downregulated
33721	Intrônico sTI	51741	WWOX	WW domain containing oxidoreductase	Downregulated
33805	Intrônico sTI	84187	-	-	Downregulated
33815	Intrônico sTI	57478	USP31	ubiquitin specific protease 31	Downregulated
33901	Intrônico sTI	129685	TBN	taube nuss homolog (mouse)	Downregulated
33909	Intrônico sTI	80895	ILKAP	integrin-linked kinase-associated serine/threonine phosphatase 2C	Downregulated
33944	Intrônico sTI	10785	WDR4	WD repeat domain 4	Downregulated
34035	Intrônico sTI	28969	BZW2	basic leucine zipper and W2 domains 2	Downregulated
34206	Intrônico sTI	221120	-	-	Downregulated
34305	Intrônico sTI	388815	C21orf34	chromosome 21 open reading frame 34	Downregulated
34326	Intrônico sTI	64417	-	-	Downregulated
34416	Intrônico sTI	138428	C9orf115	chromosome 9 open reading frame 115	Downregulated
34426	Intrônico sTI	80115	BAIAP2L2	BAI1-associated protein 2-like 2	Downregulated
34644	Intrônico sTI	55422	ZNF331	zinc finger protein 331	Downregulated
34826	Intrônico sTI	160335	-	-	Downregulated
34970	Intrônico sTI	23043	TNIK	TRAF2 and NCK interacting kinase	Downregulated
34973	Intrônico sTI	11044	POLS	polymerase (DNA directed) sigma	Downregulated
35049	Intrônico sTI	60412	SEC8L1	SEC8-like 1 (S. cerevisiae)	Downregulated
35073	Intrônico sTI	23118	MAP3K7IP2	mitogen-activated protein kinase kinase kinase 7 interacting protein 2	Downregulated
35140	Intrônico sTI	91833	WDR20	WD repeat domain 20	Downregulated
35354	Intrônico sTI	79684	-	-	Downregulated
35454	Intrônico sTI	83641	C10orf45	chromosome 10 open reading frame 45	Downregulated

35580	Intrônico sTI	27113	BBC3	BCL2 binding component 3	Downregulated
35785	Intrônico sTI	29072	-	-	Downregulated
35881	Intrônico sTI	55207	ARL10C	ADP-ribosylation factor-like 10C	Downregulated
35907	Intrônico sTI	9896	KIAA0274	KIAA0274	Downregulated
35981	Intrônico sTI	55691	FRMD4A	FERM domain containing 4A	Downregulated
35988	Intrônico sTI	51454	GULP1	GULP, engulfment adaptor PTB domain containing 1	Downregulated
36046	Intrônico sTI	862	RUNX1T1	runt-related transcription factor 1; translocated to, 1 (cyclin D-related)	Downregulated
36166	Intrônico sTI	140885	PTPNS1	protein tyrosine phosphatase, non-receptor type substrate 1	Downregulated
36170	Intrônico sTI	91768	CABLES1	Cdk5 and Abl enzyme substrate 1	Downregulated
36287	Intrônico sTI	57484	RNF150	ring finger protein 150	Downregulated
36308	Intrônico sTI	23236	PLCB1	phospholipase C, beta 1 (phosphoinositide-specific)	Downregulated
36605	Intrônico sTI	51741	WWOX	WW domain containing oxidoreductase	Downregulated
36679	Intrônico sTI	196951	-	-	Downregulated
36697	Intrônico sTI	60412	SEC8L1	SEC8-like 1 (S. cerevisiae)	Downregulated
36799	Intrônico sTI	60561	-	-	Downregulated
36899	Intrônico sTI	79873	-	-	Downregulated
37048	Intrônico sTI	8036	SHOC2	soc-2 suppressor of clear homolog (C. elegans)	Downregulated
37135	Intrônico sTI	22872	SEC31L1	SEC31-like 1 (S. cerevisiae)	Downregulated
37344	Intrônico sTI	10745	PHTF1	putative homeodomain transcription factor 1	Downregulated
37501	Intrônico sTI	257194	NEGR1	neuronal growth regulator 1	Downregulated
37521	Intrônico sTI	60412	SEC8L1	SEC8-like 1 (S. cerevisiae)	Downregulated
37538	Intrônico sTI	27068	-	-	Downregulated
37651	Intrônico sTI	11034	DSTN	destrin (actin depolymerizing factor)	Downregulated
37854	Intrônico sTI	23042	-	-	Downregulated
37898	Intrônico sTI	1408	CRY2	cryptochrome 2 (photolyase like)	Downregulated
37939	Intrônico sTI	55677	-	-	Downregulated
37955	Intrônico sTI	23332	CLASP1	cytoplasmic linker associated protein 1	Downregulated
38048	Intrônico sTI	51444	RNF138	ring finger protein 138	Downregulated
38219	Intrônico sTI	388536	-	-	Downregulated
38226	Intrônico sTI	140733	C20orf133	chromosome 20 open reading frame 133	Downregulated
38257	Intrônico sTI	29057	TMEM29	transmembrane protein 29	Downregulated
38277	Intrônico sTI	9649	RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	Downregulated
38326	Intrônico sTI	221395	GPR116	G protein-coupled receptor 116	Downregulated
38600	Intrônico sTI	6456	SH3GL2	SH3-domain GRB2-like 2	Downregulated
38867	Intrônico sTI	55234	SMU1	smu-1 suppressor of mec-8 and unc-52 homolog (C. elegans)	Downregulated
38891	Intrônico sTI	10196	HRMT1L3	HMT1 hnRNP methyltransferase-like 3 (S. cerevisiae)	Downregulated
38894	Intrônico sTI	9778	-	-	Downregulated

39010	Intrônico sTI	6871	TADA2L	transcriptional adaptor 2 (ADA2 homolog, yeast)-like	Downregulated
39497	Intrônico sTI	23788	MTCH2	mitochondrial carrier homolog 2 (C. elegans)	Downregulated
39602	Intrônico sTI	23476	BRD4	bromodomain containing 4	Downregulated
39792	Intrônico sTI	55800	SCN3B	sodium channel, voltage-gated, type III, beta	Downregulated
39930	Intrônico sTI	8620	NPFF	neuropeptide FF-amide peptide precursor	Downregulated
39981	Intrônico sTI	113540	CKLFSF1	chemokine-like factor super family 1	Downregulated
40037	Intrônico sTI	7528	YY1	YY1 transcription factor	Downregulated
40179	Intrônico sTI	257397	-	-	Downregulated
40245	Intrônico sTI	7082	TJP1	tight junction protein 1 (zonula occludens 1)	Downregulated
40440	Intrônico sTI	9759	HDAC4	histone deacetylase 4	Downregulated
40528	Intrônico sTI	23499	MACF1	microtubule-actin crosslinking factor 1	Downregulated
40619	Intrônico sTI	55209	-	-	Downregulated
40740	Intrônico sTI	11216	AKAP10	A kinase (PRKA) anchor protein 10	Downregulated
41031	Intrônico sTI	55209	-	-	Downregulated
41148	Intrônico sTI	23247	-	-	Downregulated
41150	Intrônico sTI	23041	-	-	Downregulated
41227	Intrônico sTI	65056	-	-	Downregulated
41414	Intrônico sTI	285458	-	-	Downregulated
41481	Intrônico sTI	7082	TJP1	tight junction protein 1 (zonula occludens 1)	Downregulated
41725	Intrônico sTI	200008	-	-	Downregulated
42186	Intrônico sTI	10404	-	-	Downregulated
42278	Intrônico sTI	25938	C14orf125	chromosome 14 open reading frame 125	Downregulated
42296	Intrônico sTI	9467	SH3BP5	SH3-domain binding protein 5 (BTK-associated)	Downregulated
42299	Intrônico sTI	8813	DPM1	dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit	Downregulated
42341	Intrônico sTI	259230	TMEM23	transmembrane protein 23	Downregulated
42431	Intrônico sTI	1112	CHES1	checkpoint suppressor 1	Downregulated
42803	Intrônico sTI	10561	IFI44	interferon-induced protein 44	Downregulated
43215	Intrônico sTI	10561	IFI44	interferon-induced protein 44	Downregulated
43255	Intrônico sTI	1112	CHES1	checkpoint suppressor 1	Downregulated
43392	Intrônico sTI	57666	-	-	Downregulated
44183	Intrônico sTI	1112	CHES1	checkpoint suppressor 1	Downregulated
16	Intrônico asTI	115004	C6orf150	chromosome 6 open reading frame 150	Upregulated
5981	Intrônico asTI	347731	LRRTM3	leucine rich repeat transmembrane neuronal 3	Upregulated
6690	Intrônico asTI	1292	COL6A2	collagen, type VI, alpha 2	Upregulated
7415	Intrônico asTI	288	ANK3	ankyrin 3, node of Ranvier (ankyrin G)	Upregulated
9076	Intrônico asTI	157680	VPS13B	vacuolar protein sorting 13B (yeast)	Upregulated
9791	Intrônico asTI	404266	-	-	Upregulated
9916	Intrônico asTI	64097	EPB41L4A	erythrocyte membrane protein band 4.1 like 4A	Upregulated
10956	Intrônico asTI	54971	BANP	BTG3 associated nuclear protein	Upregulated
11794	Intrônico asTI	23213	SULF1	sulfatase 1	Upregulated
12811	Intrônico asTI	54796	BNC2	basonuclein 2	Upregulated
13212	Intrônico asTI	64122	FN3K	fructosamine 3 kinase	Upregulated

13268	Intrônico asTI	5087	PBX1	pre-B-cell leukemia transcription factor 1	Upregulated
14444	Intrônico asTI	80014	-	-	Upregulated
15195	Intrônico asTI	22936	ELL2	elongation factor, RNA polymerase II, 2	Upregulated
17060	Intrônico asTI	9524	GPSN2	glycoprotein, synaptic 2	Upregulated
17724	Intrônico asTI	284352	-	-	Upregulated
17996	Intrônico asTI	7204	TRIO	triple functional domain (PTPRF interacting)	Upregulated
18664	Intrônico asTI	83890	SPATA9	spermatogenesis associated 9	Upregulated
20329	Intrônico asTI	54881	TEX10	testis expressed sequence 10	Upregulated
20504	Intrônico asTI	338557	GPR120	G protein-coupled receptor 120	Upregulated
20871	Intrônico asTI	9611	NCOR1	nuclear receptor co-repressor 1	Upregulated
20953	Intrônico asTI	29116	MYLIP	myosin regulatory light chain interacting protein	Upregulated
21051	Intrônico asTI	54805	CNNM2	cyclin M2	Upregulated
21131	Intrônico asTI	114805	GALNT13	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 13 (GalNAc-T13)	Upregulated
21968	Intrônico asTI	57713	SFMBT2	Scm-like with four mbt domains 2	Upregulated
22929	Intrônico asTI	9877	ZC3H11A	zinc finger CCCH-type containing 11A	Upregulated
23000	Intrônico asTI	57181	SLC39A10	solute carrier family 39 (zinc transporter), member 10	Upregulated
23226	Intrônico asTI	23261	CAMTA1	calmodulin binding transcription activator 1	Upregulated
23755	Intrônico asTI	9612	NCOR2	nuclear receptor co-repressor 2	Upregulated
25266	Intrônico asTI	57178	RAI17	retinoic acid induced 17	Upregulated
26451	Intrônico asTI	5297	PIK4CA	phosphatidylinositol 4-kinase, catalytic, alpha polypeptide	Upregulated
26671	Intrônico asTI	1612	DAPK1	death-associated protein kinase 1	Upregulated
26737	Intrônico asTI	10217	CTDSPL	CTD (carboxy-terminal domain, RNA polymerase II polypeptide A) small phosphatase-like	Upregulated
27667	Intrônico asTI	9886	RHOBTB1	Rho-related BTB domain containing 1	Upregulated
27775	Intrônico asTI	8834	C17orf35	chromosome 17 open reading frame 35	Upregulated
29101	Intrônico asTI	22862	FNDC3A	fibronectin type III domain containing 3A	Upregulated
29111	Intrônico asTI	9620	CELSR1	cadherin, EGF LAG seven-pass G-type receptor 1 (flamingo homolog, Drosophila)	Upregulated
29627	Intrônico asTI	9265	PSCD3	pleckstrin homology, Sec7 and coiled-coil domains 3	Upregulated
30361	Intrônico asTI	6095	RORA	RAR-related orphan receptor A	Upregulated
30666	Intrônico asTI	6920	TCEA3	transcription elongation factor A (SII), 3	Upregulated
31194	Intrônico asTI	4084	MXD1	MAX dimerization protein 1	Upregulated
32941	Intrônico asTI	4988	OPRM1	opioid receptor, mu 1	Upregulated
32996	Intrônico asTI	55384	MEG3	maternally expressed 3	Upregulated
34220	Intrônico asTI	80115	BAIAP2L2	BAI1-associated protein 2-like 2	Upregulated

34776	Intrônico asTI	9482	STX8	syntaxin 8	Upregulated
35164	Intrônico asTI	51455	REV1L	REV1-like (yeast)	Upregulated
35320	Intrônico asTI	2309	FOXO3A	forkhead box O3A	Upregulated
35425	Intrônico asTI	1730	DIAPH2	diaphanous homolog 2 (Drosophila)	Upregulated
35494	Intrônico asTI	10018	BCL2L11	BCL2-like 11 (apoptosis facilitator)	Upregulated
36228	Intrônico asTI	6904	TBCD	tubulin-specific chaperone d	Upregulated
36524	Intrônico asTI	10019	-	-	Upregulated
36744	Intrônico asTI	6868	ADAM17	a disintegrin and metalloproteinase domain 17 (tumor necrosis factor, alpha, converting enzyme)	Upregulated
38596	Intrônico asTI	7294	TXK	TXK tyrosine kinase	Upregulated
38852	Intrônico asTI	83696	-	-	Upregulated
39410	Intrônico asTI	9759	HDAC4	histone deacetylase 4	Upregulated
39490	Intrônico asTI	51441	YTHDF2	YTH domain family, member 2	Upregulated
41771	Intrônico asTI	11030	RBPMS	RNA binding protein with multiple splicing	Upregulated
42331	Intrônico asTI	645	BLVRB	biliverdin reductase B (flavin reductase (NADPH))	Upregulated
42529	Intrônico asTI	2744	GLS	glutaminase	Upregulated
43764	Intrônico asTI	2969	GTF2I	general transcription factor II, i	Upregulated
44068	Intrônico asTI	4129	MAOB	monoamine oxidase B	Upregulated
44275	Intrônico asTI	3899	LAF4	lymphoid nuclear protein related to AF4	Upregulated
64	Intrônico asTI	9698	PUM1	pumilio homolog 1 (Drosophila)	Downregulated
160	Intrônico asTI	10905	MAN1A2	mannosidase, alpha, class 1A, member 2	Downregulated
178	Intrônico asTI	7024	TFCP2	transcription factor CP2	Downregulated
293	Intrônico asTI	4212	MEIS2	Meis1, myeloid ecotropic viral integration site 1 homolog 2 (mouse)	Downregulated
373	Intrônico asTI	9698	PUM1	pumilio homolog 1 (Drosophila)	Downregulated
385	Intrônico asTI	6777	STAT5B	signal transducer and activator of transcription 5B	Downregulated
487	Intrônico asTI	7026	NR2F2	nuclear receptor subfamily 2, group F, member 2	Downregulated
490	Intrônico asTI	6310	ATXN1	ataxin 1	Downregulated
536	Intrônico asTI	84067	-	-	Downregulated
574	Intrônico asTI	10483	SEC23B	Sec23 homolog B (S. cerevisiae)	Downregulated
586	Intrônico asTI	8289	ARID1A	AT rich interactive domain 1A (SWI-like)	Downregulated
903	Intrônico asTI	5980	REV3L	REV3-like, catalytic subunit of DNA polymerase zeta (yeast)	Downregulated
978	Intrônico asTI	23284	LPHN3	latrophilin 3	Downregulated
987	Intrônico asTI	10277	UBE4B	ubiquitination factor E4B (UFD2 homolog, yeast)	Downregulated
993	Intrônico asTI	9356	SLC22A6	solute carrier family 22 (organic anion transporter), member 6	Downregulated
1057	Intrônico asTI	64764	CREB3L2	cAMP responsive element binding protein 3-like 2	Downregulated
1188	Intrônico asTI	11346	SYNPO	synaptopodin	Downregulated
1189	Intrônico asTI	11107	PRDM5	PR domain containing 5	Downregulated
1243	Intrônico asTI	342184	FMN1	formin 1	Downregulated

1260	Intrônico asTI	79987	SVEP1	sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1	Downregulated
1299	Intrônico asTI	9844	ELMO1	engulfment and cell motility 1 (ced-12 homolog, C. elegans)	Downregulated
1654	Intrônico asTI	440574	-	-	Downregulated
1711	Intrônico asTI	9844	ELMO1	engulfment and cell motility 1 (ced-12 homolog, C. elegans)	Downregulated
1774	Intrônico asTI	80267	C1orf22	chromosome 1 open reading frame 22	Downregulated
1924	Intrônico asTI	8531	CSDA	cold shock domain protein A	Downregulated
1979	Intrônico asTI	81619	TSPAN14	tetraspanin 14	Downregulated
2040	Intrônico asTI	5218	PFTK1	PFTAIRE protein kinase 1	Downregulated
2118	Intrônico asTI	10644	-	-	Downregulated
2238	Intrônico asTI	7024	TFCP2	transcription factor CP2	Downregulated
2283	Intrônico asTI	93145	OLFM2	olfactomedin 2	Downregulated
2465	Intrônico asTI	1793	DOCK1	dedicator of cytokinesis 1	Downregulated
2472	Intrônico asTI	81	ACTN4	actinin, alpha 4	Downregulated
2636	Intrônico asTI	10120	ACTR1B	ARP1 actin-related protein 1 homolog B, centractin beta (yeast)	Downregulated
2895	Intrônico asTI	170959	ZNF431	zinc finger protein 431	Downregulated
2933	Intrônico asTI	23597	-	-	Downregulated
2993	Intrônico asTI	440173	-	-	Downregulated
3100	Intrônico asTI	219623	TMEM26	transmembrane protein 26	Downregulated
3129	Intrônico asTI	54765	TRIM44	tripartite motif-containing 44	Downregulated
3224	Intrônico asTI	57223	-	-	Downregulated
3334	Intrônico asTI	54927	CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing 3	Downregulated
3432	Intrônico asTI	55900	ZNF302	zinc finger protein 302	Downregulated
3502	Intrônico asTI	81	ACTN4	actinin, alpha 4	Downregulated
3513	Intrônico asTI	171017	ZNF384	zinc finger protein 384	Downregulated
3567	Intrônico asTI	9577	BRE	brain and reproductive organ-expressed (TNFRSF1A modulator)	Downregulated
3774	Intrônico asTI	9369	NRXN3	neurexin 3	Downregulated
3792	Intrônico asTI	5069	PAPPA	pregnancy-associated plasma protein A, pappalysin 1	Downregulated
4016	Intrônico asTI	309	ANXA6	annexin A6	Downregulated
4219	Intrônico asTI	989	7-Sep	septin 7	Downregulated
4296	Intrônico asTI	7403	UTX	ubiquitously transcribed tetratricopeptide repeat, X chromosome	Downregulated
4378	Intrônico asTI	23199	-	-	Downregulated
4389	Intrônico asTI	9847	-	-	Downregulated
4487	Intrônico asTI	10627	-	-	Downregulated
4523	Intrônico asTI	2187	FANCB	Fanconi anemia, complementation group B	Downregulated
4707	Intrônico asTI	7804	LRP8	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	Downregulated
4757	Intrônico asTI	84909	C9orf3	chromosome 9 open reading frame 3	Downregulated
4790	Intrônico asTI	23199	-	-	Downregulated
5096	Intrônico asTI	23576	DDAH1	dimethylarginine dimethylaminohydrolase 1	Downregulated
5175	Intrônico asTI	79758	-	-	Downregulated

5248	Intrônico asTI	1289	COL5A1	collagen, type V, alpha 1	Downregulated
5272	Intrônico asTI	84893	FBXO18	F-box protein, helicase, 18	Downregulated
5395	Intrônico asTI	54768	HYDIN	hydrocephalus inducing protein tyrosine phosphatase, receptor type, T	Downregulated
5412	Intrônico asTI	11122	PTPRT	LATS, large tumor suppressor, homolog 2 (Drosophila)	Downregulated
5505	Intrônico asTI	26524	LATS2	DNA (cytosine-5)-methyltransferase 3 alpha	Downregulated
5555	Intrônico asTI	1788	DNMT3A	chromosome 21 open reading frame 29	Downregulated
5602	Intrônico asTI	54084	C21orf29	death-associated protein kinase 2	Downregulated
5611	Intrônico asTI	23604	DAPK2	-	Downregulated
5701	Intrônico asTI	55295	-	-	Downregulated
5718	Intrônico asTI	23080	-	-	Downregulated
5721	Intrônico asTI	11101	ATE1	arginyltransferase 1	Downregulated
5729	Intrônico asTI	9690	UBE3C	ubiquitin protein ligase E3C	Downregulated
5762	Intrônico asTI	1508	CTSB	cathepsin B	Downregulated
5795	Intrônico asTI	64770	CCDC14	coiled-coil domain containing 14	Downregulated
5801	Intrônico asTI	55909	BIN3	bridging integrator 3	Downregulated
5984	Intrônico asTI	219736	STOX1	storkhead box 1	Downregulated
6102	Intrônico asTI	79731	-	-	Downregulated
6139	Intrônico asTI	9961	MVP	major vault protein	Downregulated
6615	Intrônico asTI	80314	EPC1	enhancer of polycomb homolog 1 (Drosophila)	Downregulated
6745	Intrônico asTI	23382	-	-	Downregulated
6837	Intrônico asTI	54776	PPP1R12C	protein phosphatase 1, regulatory (inhibitor) subunit 12C	Downregulated
7142	Intrônico asTI	55593	-	-	Downregulated
7548	Intrônico asTI	57799	RAB40C	RAB40C, member RAS oncogene family	Downregulated
7678	Intrônico asTI	11124	FAF1	Fas (TNFRSF6) associated factor 1	Downregulated
8003	Intrônico asTI	7791	ZYX	zyxin	Downregulated
8055	Intrônico asTI	84133	ZNRF3	zinc and ring finger 3	Downregulated
8159	Intrônico asTI	81605	C9orf74	chromosome 9 open reading frame 74	Downregulated
8191	Intrônico asTI	22976	PAXIP1	PAX interacting (with transcription-activation domain) protein 1	Downregulated
8260	Intrônico asTI	84449	ZNF333	zinc finger protein 333	Downregulated
8265	Intrônico asTI	79789	CLMN	calmin (calponin-like, transmembrane)	Downregulated
8798	Intrônico asTI	50863	-	-	Downregulated
8988	Intrônico asTI	64754	SMYD3	SET and MYND domain containing 3	Downregulated
9570	Intrônico asTI	2204	FCAR	Fc fragment of IgA, receptor for	Downregulated
10017	Intrônico asTI	79034	C7orf26	chromosome 7 open reading frame 26	Downregulated
10031	Intrônico asTI	53407	STX18	syntaxin 18	Downregulated
10121	Intrônico asTI	64778	FNDC3B	fibronectin type III domain containing 3B	Downregulated
10278	Intrônico asTI	5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide	Downregulated
10478	Intrônico asTI	7003	TEAD1	TEA domain family member 1 (SV40 transcriptional enhancer factor)	Downregulated

10595	Intrônico asTI	3778	KCNMA1	potassium large conductance calcium-activated channel, subfamily M, alpha member 1	Downregulated
10813	Intrônico asTI	577	BAI3	brain-specific angiogenesis inhibitor 3	Downregulated
11164	Intrônico asTI	54107	POLE3	polymerase (DNA directed), epsilon 3 (p17 subunit)	Downregulated
11342	Intrônico asTI	157922	CAMSAP1	calmodulin regulated spectrin-associated protein 1	Downregulated
12145	Intrônico asTI	2185	PTK2B	PTK2B protein tyrosine kinase 2 beta	Downregulated
12148	Intrônico asTI	1496	CTNNA2	catenin (cadherin-associated protein), alpha 2	Downregulated
12404	Intrônico asTI	29934	SNX12	sorting nexin 12	Downregulated
13138	Intrônico asTI	10867	TSPAN9	tetraspanin 9	Downregulated
13358	Intrônico asTI	8178	ELL	elongation factor RNA polymerase II	Downregulated
13423	Intrônico asTI	55930	MYO5C	myosin VC	Downregulated
14201	Intrônico asTI	3418	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial	Downregulated
14289	Intrônico asTI	7045	TGFB1	transforming growth factor, beta-induced, 68kDa	Downregulated
14686	Intrônico asTI	10320	ZNFN1A1	zinc finger protein, subfamily 1A, 1 (Ikaros)	Downregulated
15152	Intrônico asTI	168667	-	-	Downregulated
15189	Intrônico asTI	25819	CCRN4L	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	Downregulated
15780	Intrônico asTI	83938	C10orf11	chromosome 10 open reading frame 11	Downregulated
16015	Intrônico asTI	23368	PPP1R13B	protein phosphatase 1, regulatory (inhibitor) subunit 13B	Downregulated
16676	Intrônico asTI	2534	FYN	FYN oncogene related to SRC, FGR, YES	Downregulated
16750	Intrônico asTI	9728	-	-	Downregulated
16804	Intrônico asTI	122616	C14orf79	chromosome 14 open reading frame 79	Downregulated
16976	Intrônico asTI	5090	PBX3	pre-B-cell leukemia transcription factor 3	Downregulated
17025	Intrônico asTI	57552	AADACL1	arylacetamide deacetylase-like 1	Downregulated
17400	Intrônico asTI	1833	DSPG3	dermatan sulfate proteoglycan 3	Downregulated
17972	Intrônico asTI	23414	ZFPM2	zinc finger protein, multitype 2	Downregulated
18020	Intrônico asTI	1312	COMT	catechol-O-methyltransferase	Downregulated
18110	Intrônico asTI	4774	NFIA	nuclear factor I/A	Downregulated
18174	Intrônico asTI	27032	ATP2C1	ATPase, Ca++ transporting, type 2C, member 1	Downregulated
18237	Intrônico asTI	401013	-	-	Downregulated
18614	Intrônico asTI	7204	TRIO	triple functional domain (PTPRF interacting)	Downregulated
18652	Intrônico asTI	246721	-	-	Downregulated
18733	Intrônico asTI	3383	ICAM1	intercellular adhesion molecule 1 (CD54), human rhinovirus receptor	Downregulated
19070	Intrônico asTI	121665	-	-	Downregulated
19101	Intrônico asTI	26297	-	-	Downregulated
19812	Intrônico asTI	55346	-	-	Downregulated

20079	Intrônico asTI	1602	DACH1	dachshund homolog 1 (Drosophila)	Downregulated
20266	Intrônico asTI	6403	SELP	selectin P (granule membrane protein 140kDa, antigen CD62)	Downregulated
20477	Intrônico asTI	5287	PIK3C2B	phosphoinositide-3-kinase, class 2, beta polypeptide	Downregulated
20666	Intrônico asTI	9414	TJP2	tight junction protein 2 (zona occludens 2)	Downregulated
21061	Intrônico asTI	23646	PLD3	phospholipase D family, member 3	Downregulated
21581	Intrônico asTI	22926	ATF6	activating transcription factor 6	Downregulated
21722	Intrônico asTI	2902	GRIN1	glutamate receptor, ionotropic, N-methyl D-aspartate 1	Downregulated
22082	Intrônico asTI	54434	SSH1	slingshot homolog 1 (Drosophila)	Downregulated
22290	Intrônico asTI	51363	-	-	Downregulated
22467	Intrônico asTI	148641	SLC35F3	solute carrier family 35, member F3	Downregulated
22641	Intrônico asTI	5046	PCSK6	proprotein convertase subtilisin/kexin type 6	Downregulated
22705	Intrônico asTI	28232	SLCO3A1	solute carrier organic anion transporter family, member 3A1	Downregulated
22785	Intrônico asTI	81544	GDPD5	glycerophosphodiester phosphodiesterase domain containing 5	Downregulated
22787	Intrônico asTI	79934	ADCK4	aarF domain containing kinase 4	Downregulated
22832	Intrônico asTI	8642	DCHS1	dachsous 1 (Drosophila)	Downregulated
22983	Intrônico asTI	130535	KCTD18	potassium channel tetramerisation domain containing 18	Downregulated
23027	Intrônico asTI	10588	MTHFS	5,10-methenyltetrahydrofolate synthetase (5-formyltetrahydrofolate cycloligase)	Downregulated
23033	Intrônico asTI	9665	-	-	Downregulated
23114	Intrônico asTI	51363	-	-	Downregulated
23214	Intrônico asTI	54581	SCAND2	SCAN domain containing 2	Downregulated
23424	Intrônico asTI	50807	DDEF1	development and differentiation enhancing factor 1	Downregulated
23576	Intrônico asTI	2903	GRIN2A	glutamate receptor, ionotropic, N-methyl D-aspartate 2A	Downregulated
23663	Intrônico asTI	6733	SRPK2	SFRS protein kinase 2	Downregulated
24285	Intrônico asTI	5770	PTPN1	protein tyrosine phosphatase, non-receptor type 1	Downregulated
24290	Intrônico asTI	4771	NF2	neurofibromin 2 (bilateral acoustic neuroma)	Downregulated
24345	Intrônico asTI	55192	-	-	Downregulated
24373	Intrônico asTI	9612	NCOR2	nuclear receptor co-repressor 2	Downregulated
24445	Intrônico asTI	55711	MLSTD1	male sterility domain containing 1	Downregulated
24892	Intrônico asTI	8635	RNASET2	ribonuclease T2	Downregulated
25361	Intrônico asTI	80704	SLC19A3	solute carrier family 19, member 3	Downregulated
25362	Intrônico asTI	80036	TRPM3	transient receptor potential cation channel, subfamily M, member 3	Downregulated
25377	Intrônico asTI	54806	AHI1	Abelson helper integration site	Downregulated

25384	Intrônico asTI	27044	-	-	Downregulated
25433	Intrônico asTI	2081	ERN1	endoplasmic reticulum to nucleus signalling 1	Downregulated
25440	Intrônico asTI	351	APP	amyloid beta (A4) precursor protein (protease nexin-II, Alzheimer disease)	Downregulated
25483	Intrônico asTI	51203	NUSAP1	nucleolar and spindle associated protein 1	Downregulated
25585	Intrônico asTI	51663	ZFR	zinc finger RNA binding protein	Downregulated
25746	Intrônico asTI	899	CCNF	cyclin F	Downregulated
25788	Intrônico asTI	55016	1-Mar	membrane-associated ring finger (C3HC4) 1	Downregulated
25794	Intrônico asTI	29990	PILRB	paired immunoglobulin-like type 2 receptor beta	Downregulated
25876	Intrônico asTI	80205	CHD9	chromodomain helicase DNA binding protein 9	Downregulated
25883	Intrônico asTI	57522	SRGAP1	SLIT-ROBO Rho GTPase activating protein 1	Downregulated
25890	Intrônico asTI	55090	MED9	mediator of RNA polymerase II transcription, subunit 9 homolog (yeast)	Downregulated
26067	Intrônico asTI	283349	RASSF3	Ras association (RalGDS/AF-6) domain family 3	Downregulated
26282	Intrônico asTI	92259	MRPS36	mitochondrial ribosomal protein S36	Downregulated
26292	Intrônico asTI	64478	CSMD1	CUB and Sushi multiple domains 1	Downregulated
26299	Intrônico asTI	55706	TMEM48	transmembrane protein 48	Downregulated
26501	Intrônico asTI	57521	-	-	Downregulated
26692	Intrônico asTI	121260	SLC15A4	solute carrier family 15, member 4	Downregulated
26713	Intrônico asTI	55256	-	-	Downregulated
26788	Intrônico asTI	284900	-	-	Downregulated
26810	Intrônico asTI	57619	-	-	Downregulated
26827	Intrônico asTI	26115	-	-	Downregulated
26867	Intrônico asTI	4292	MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	Downregulated
26896	Intrônico asTI	143458	-	-	Downregulated
26930	Intrônico asTI	25992	SNED1	sushi, nidogen and EGF-like domains 1	Downregulated
27025	Intrônico asTI	54808	DYM	dymeclin	Downregulated
27391	Intrônico asTI	1740	DLG2	discs, large homolog 2, chapsyn-110 (Drosophila)	Downregulated
27425	Intrônico asTI	64855	C9orf88	chromosome 9 open reading frame 88	Downregulated
27516	Intrônico asTI	121260	SLC15A4	solute carrier family 15, member 4	Downregulated
27717	Intrônico asTI	201266	SLC39A11	solute carrier family 39 (metal ion transporter), member 11	Downregulated
27834	Intrônico asTI	80036	TRPM3	transient receptor potential cation channel, subfamily M, member 3	Downregulated
28239	Intrônico asTI	112399	EGLN3	egl nine homolog 3 (C. elegans)	Downregulated
28374	Intrônico asTI	23514	-	-	Downregulated
28404	Intrônico asTI	5908	RAP1B	RAP1B, member of RAS oncogene family	Downregulated
28747	Intrônico asTI	201266	SLC39A11	solute carrier family 39 (metal ion transporter), member 11	Downregulated

28752	Intrônico asTI	121227	LRIG3	leucine-rich repeats and immunoglobulin-like domains 3	Downregulated
28770	Intrônico asTI	55841	-	-	Downregulated
28793	Intrônico asTI	11057	ABHD2	abhydrolase domain containing 2	Downregulated
28838	Intrônico asTI	463	ATBF1	AT-binding transcription factor 1	Downregulated
28847	Intrônico asTI	375387	LRRC33	leucine rich repeat containing 33	Downregulated
28968	Intrônico asTI	79701	-	-	Downregulated
28988	Intrônico asTI	27132	CPNE7	copine VII	Downregulated
29055	Intrônico asTI	256536	TCERG1L	transcription elongation regulator 1-like	Downregulated
29058	Intrônico asTI	158947	ARMCX4	armadillo repeat containing, X-linked 4	Downregulated
29064	Intrônico asTI	91057	-	-	Downregulated
29091	Intrônico asTI	28957	MRPS28	mitochondrial ribosomal protein S28	Downregulated
29468	Intrônico asTI	221037	JMJD1C	jumonji domain containing 1C	Downregulated
29497	Intrônico asTI	54812	-	-	Downregulated
29575	Intrônico asTI	130074	-	-	Downregulated
29987	Intrônico asTI	130074	-	-	Downregulated
30007	Intrônico asTI	55703	POLR3B	polymerase (RNA) III (DNA directed) polypeptide B	Downregulated
30113	Intrônico asTI	55198	-	-	Downregulated
30292	Intrônico asTI	221037	JMJD1C	jumonji domain containing 1C	Downregulated
30721	Intrônico asTI	64901	RANBP17	RAN binding protein 17	Downregulated
30915	Intrônico asTI	126668	TDRD10	tudor domain containing 10	Downregulated
30925	Intrônico asTI	79848	-	-	Downregulated
30953	Intrônico asTI	23112	TNRC6B	trinucleotide repeat containing 6B	Downregulated
31016	Intrônico asTI	143187	VTI1A	vesicle transport through interaction with t-SNAREs homolog 1A (yeast)	Downregulated
31025	Intrônico asTI	81285	OR51E2	olfactory receptor, family 51, subfamily E, member 2	Downregulated
31075	Intrônico asTI	7627	ZNF75A	zinc finger protein 75a	Downregulated
31215	Intrônico asTI	399474	-	-	Downregulated
31224	Intrônico asTI	120863	DEPDC4	DEP domain containing 4	Downregulated
31226	Intrônico asTI	92002	FAM58A	family with sequence similarity 58, member A	Downregulated
31229	Intrônico asTI	84269	CHCHD5	coiled-coil-helix-coiled-coil-helix domain containing 5	Downregulated
31245	Intrônico asTI	55252	ASXL2	additional sex combs like 2 (Drosophila)	Downregulated
31257	Intrônico asTI	25771	TBC1D22A	TBC1 domain family, member 22A	Downregulated
31260	Intrônico asTI	23254	-	-	Downregulated
31339	Intrônico asTI	64902	AGXT2	alanine-glyoxylate aminotransferase 2	Downregulated
31345	Intrônico asTI	56172	ANKH	ankylosis, progressive homolog (mouse)	Downregulated
31437	Intrônico asTI	81285	OR51E2	olfactory receptor, family 51, subfamily E, member 2	Downregulated
31446	Intrônico asTI	57154	SMURF1	SMAD specific E3 ubiquitin protein ligase 1	Downregulated
31539	Intrônico asTI	84186	ZCCHC7	zinc finger, CCHC domain containing 7	Downregulated

31542	Intrônico asTI	80059	LRRTM4	leucine rich repeat transmembrane neuronal 4	Downregulated
31670	Intrônico asTI	23512	SUZ12	suppressor of zeste 12 homolog (Drosophila)	Downregulated
31735	Intrônico asTI	196527	TMEM16F	transmembrane protein 16F	Downregulated
31750	Intrônico asTI	79611	-	-	Downregulated
31846	Intrônico asTI	84750	FUT10	fucosyltransferase 10 (alpha (1,3) fucosyltransferase)	Downregulated
31870	Intrônico asTI	50618	ITSN2	intersectin 2	Downregulated
31963	Intrônico asTI	56204	-	-	Downregulated
32084	Intrônico asTI	23254	-	-	Downregulated
32159	Intrônico asTI	80741	LY6G5C	lymphocyte antigen 6 complex, locus G5C	Downregulated
32353	Intrônico asTI	196527	TMEM16F	transmembrane protein 16F	Downregulated
32367	Intrônico asTI	79866	-	-	Downregulated
32376	Intrônico asTI	55790	-	-	Downregulated
32391	Intrônico asTI	23673	STX12	syntaxin 12	Downregulated
32520	Intrônico asTI	6917	TCEA1	transcription elongation factor A (SII), 1	Downregulated
32583	Intrônico asTI	55666	-	-	Downregulated
32661	Intrônico asTI	201134	-	-	Downregulated
32673	Intrônico asTI	81035	COLEC12	collectin sub-family member 12	Downregulated
32772	Intrônico asTI	91133	L3MBTL4	l(3)mbt-like 4 (Drosophila)	Downregulated
32906	Intrônico asTI	23505	-	-	Downregulated
33091	Intrônico asTI	60676	PAPPA2	pappalysin 2	Downregulated
33285	Intrônico asTI	114788	CSMD3	CUB and Sushi multiple domains 3	Downregulated
33386	Intrônico asTI	138241	C9orf85	chromosome 9 open reading frame 85	Downregulated
33399	Intrônico asTI	64924	SLC30A5	solute carrier family 30 (zinc transporter), member 5	Downregulated
33424	Intrônico asTI	23235	SNF1LK2	SNF1-like kinase 2	Downregulated
33898	Intrônico asTI	163702	IL28RA	interleukin 28 receptor, alpha (interferon, lambda receptor)	Downregulated
33906	Intrônico asTI	84733	CBX2	chromobox homolog 2 (Pc class homolog, Drosophila)	Downregulated
33997	Intrônico asTI	378805	-	-	Downregulated
34008	Intrônico asTI	91156	-	-	Downregulated
34145	Intrônico asTI	23144	ZC3H3	zinc finger CCCH-type containing 3	Downregulated
34310	Intrônico asTI	163702	IL28RA	interleukin 28 receptor, alpha (interferon, lambda receptor)	Downregulated
34319	Intrônico asTI	84260	-	-	Downregulated
34348	Intrônico asTI	23503	ZFYVE26	zinc finger, FYVE domain containing 26	Downregulated
34449	Intrônico asTI	26133	TRPC4AP	transient receptor potential cation channel, subfamily C member 4 associated protein	Downregulated
34466	Intrônico asTI	9753	-	-	Downregulated
34536	Intrônico asTI	57153	-	-	Downregulated
34624	Intrônico asTI	116987	CENTG2	centaurin, gamma 2 tRNA splicing	Downregulated
34631	Intrônico asTI	80746	TSEN2	endonuclease 2 homolog (SEN2, S. cerevisiae)	Downregulated
34636	Intrônico asTI	64319	FBS1	fibrosin 1	Downregulated
34861	Intrônico asTI	26133	TRPC4AP	transient receptor potential cation channel, subfamily C member 4 associated protein	Downregulated

35068	Intrônico asTI	25927	C2orf32	chromosome 2 open reading frame 32	Downregulated
35167	Intrônico asTI	29072	-	-	Downregulated
35177	Intrônico asTI	22905	EPN2	epsin 2	Downregulated
35179	Intrônico asTI	11044	POLS	polymerase (DNA directed) sigma	Downregulated
35294	Intrônico asTI	8997	KALRN	kalirin, RhoGEF kinase	Downregulated
35335	Intrônico asTI	388815	C21orf34	chromosome 21 open reading frame 34	Downregulated
35463	Intrônico asTI	57479	KIAA1205	KIAA1205	Downregulated
35466	Intrônico asTI	55796	MBNL3	muscleblind-like 3 (Drosophila)	Downregulated
35581	Intrônico asTI	26229	B3GAT3	beta-1,3-glucuronyltransferase 3 (glucuronosyltransferase I)	Downregulated
35597	Intrônico asTI	9912	-	-	Downregulated
35720	Intrônico asTI	5527	PPP2R5C	protein phosphatase 2, regulatory subunit B (B56), gamma isoform	Downregulated
35750	Intrônico asTI	221935	SDK1	sidekick homolog 1 (chicken)	Downregulated
35960	Intrônico asTI	140885	PTPNS1	protein tyrosine phosphatase, non-receptor type substrate 1	Downregulated
35987	Intrônico asTI	51741	WWOX	WW domain containing oxidoreductase	Downregulated
36001	Intrônico asTI	22905	EPN2	epsin 2	Downregulated
36016	Intrônico asTI	8621	CDC2L5	cell division cycle 2-like 5 (cholinesterase-related cell division controller)	Downregulated
36293	Intrônico asTI	55209	-	-	Downregulated
36329	Intrônico asTI	7520	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen, 80kDa)	Downregulated
36390	Intrônico asTI	57148	-	-	Downregulated
36921	Intrônico asTI	26137	ZBTB20	zinc finger and BTB domain containing 20	Downregulated
36941	Intrônico asTI	9213	XPR1	xenotropic and polytropic retrovirus receptor	Downregulated
37016	Intrônico asTI	54517	-	-	Downregulated
37202	Intrônico asTI	84725	PLEKHA8	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 8	Downregulated
37334	Intrônico asTI	25932	CLIC4	chloride intracellular channel 4	Downregulated
37339	Intrônico asTI	23122	CLASP2	cytoplasmic linker associated protein 2	Downregulated
37550	Intrônico asTI	10745	PHTF1	putative homeodomain transcription factor 1	Downregulated
37609	Intrônico asTI	129531	-	-	Downregulated
37863	Intrônico asTI	9910	RABGAP1L	RAB GTPase activating protein 1-like	Downregulated
38047	Intrônico asTI	51741	WWOX	WW domain containing oxidoreductase	Downregulated
38550	Intrônico asTI	64324	NSD1	nuclear receptor binding SET domain protein 1	Downregulated
38556	Intrônico asTI	55799	CACNA2D3	calcium channel, voltage-dependent, alpha 2/delta 3 subunit	Downregulated
38569	Intrônico asTI	26146	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1	Downregulated

38602	Intrônico asTI	5868	RAB5A	RAB5A, member RAS oncogene family	Downregulated
38662	Intrônico asTI	55064	C9orf68	chromosome 9 open reading frame 68	Downregulated
38755	Intrônico asTI	65055	C2orf23	chromosome 2 open reading frame 23	Downregulated
38911	Intrônico asTI	5728	PTEN	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)	Downregulated
38962	Intrônico asTI	64324	NSD1	nuclear receptor binding SET domain protein 1	Downregulated
38984	Intrônico asTI	23476	BRD4	bromodomain containing 4	Downregulated
39000	Intrônico asTI	9456	HOMER1	homer homolog 1 (Drosophila)	Downregulated
39181	Intrônico asTI	51704	GPRC5B	G protein-coupled receptor, family C, group 5, member B	Downregulated
39314	Intrônico asTI	8031	NCOA4	nuclear receptor coactivator 4	Downregulated
39516	Intrônico asTI	9043	SPAG9	sperm associated antigen 9	Downregulated
39519	Intrônico asTI	8451	CUL4A	cullin 4A	Downregulated
39593	Intrônico asTI	51704	GPRC5B	G protein-coupled receptor, family C, group 5, member B	Downregulated
39607	Intrônico asTI	22876	INPP5F	inositol polyphosphate-5-phosphatase F	Downregulated
39693	Intrônico asTI	54839	LRRC49	leucine rich repeat containing 49	Downregulated
39731	Intrônico asTI	6695	SPOCK	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican)	Downregulated
39808	Intrônico asTI	23476	BRD4	bromodomain containing 4	Downregulated
39841	Intrônico asTI	5315	PKM2	pyruvate kinase, muscle	Downregulated
39928	Intrônico asTI	9043	SPAG9	sperm associated antigen 9	Downregulated
40040	Intrônico asTI	6872	TAF1	TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa	Downregulated
40108	Intrônico asTI	51441	YTHDF2	YTH domain family, member 2	Downregulated
40322	Intrônico asTI	23499	MACF1	microtubule-actin crosslinking factor 1	Downregulated
40495	Intrônico asTI	114569	MAL2	mal, T-cell differentiation protein 2	Downregulated
41587	Intrônico asTI	6164	RPL34	ribosomal protein L34	Downregulated
41873	Intrônico asTI	22884	WDR37	WD repeat domain 37	Downregulated
42193	Intrônico asTI	9229	DLGAP1	discs, large (Drosophila) homolog-associated protein 1	Downregulated
42239	Intrônico asTI	259230	TMEM23	transmembrane protein 23	Downregulated
42279	Intrônico asTI	23762	OSBP2	oxysterol binding protein 2	Downregulated
42682	Intrônico asTI	54502	-	-	Downregulated
42918	Intrônico asTI	8613	PPAP2B	phosphatidic acid phosphatase type 2B	Downregulated
43101	Intrônico asTI	26155	-	-	Downregulated
43782	Intrônico asTI	285590	KIAA1295	KIAA1295	Downregulated
43795	Intrônico asTI	84236	-	-	Downregulated
44100	Intrônico asTI	117156	SCGB3A2	secretoglobin, family 3A, member 2	Downregulated
44245	Intrônico asTI	10558	SPTLC1	serine palmitoyltransferase, long chain base subunit 1	Downregulated

1945	Intrônico asPI	3101	HK3	hexokinase 3 (white cell)	Upregulated
4519	Intrônico asPI	3479	IGF1	insulin-like growth factor 1 (somatomedin C)	Upregulated
5745	Intrônico asPI	5789	PTPRD	protein tyrosine phosphatase, receptor type, D	Upregulated
6361	Intrônico asPI	6303	SAT	spermidine/spermine N1-acetyltransferase	Upregulated
6561	Intrônico asPI	7803	PTP4A1	protein tyrosine phosphatase type IVA, member 1	Upregulated
8416	Intrônico asPI	7399	USH2A	Usher syndrome 2A (autosomal recessive, mild)	Upregulated
9284	Intrônico asPI	134526	-	-	Upregulated
9763	Intrônico asPI	5636	PRPSAP2	phosphoribosyl pyrophosphate synthetase-associated protein 2	Upregulated
9957	Intrônico asPI	8720	MBTPS1	membrane-bound transcription factor protease, site 1	Upregulated
10029	Intrônico asPI	54899	PXK	PX domain containing serine/threonine kinase	Upregulated
10843	Intrônico asPI	63967	CLSPN	claspin homolog (Xenopus laevis)	Upregulated
11037	Intrônico asPI	114880	OSBPL6	oxysterol binding protein-like 6	Upregulated
11284	Intrônico asPI	10873	ME3	malic enzyme 3, NADP(+)-dependent, mitochondrial	Upregulated
11391	Intrônico asPI	10142	AKAP9	A kinase (PRKA) anchor protein (yotiao) 9	Upregulated
11396	Intrônico asPI	9382	COG1	component of oligomeric golgi complex 1	Upregulated
11709	Intrônico asPI	8500	PPFIA1	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 1	Upregulated
12241	Intrônico asPI	4245	MGAT1	mannosyl (alpha-1,3-)-glycoprotein beta-1,2-N-acetylglucosaminyltransferase	Upregulated
12339	Intrônico asPI	5426	POLE	polymerase (DNA directed), epsilon	Upregulated
12436	Intrônico asPI	6794	STK11	serine/threonine kinase 11 (Peutz-Jeghers syndrome)	Upregulated
12797	Intrônico asPI	79811	-	-	Upregulated
13024	Intrônico asPI	26999	CYFIP2	cytoplasmic FMR1 interacting protein 2	Upregulated
13447	Intrônico asPI	10942	PRSS21	protease, serine, 21 (testisin)	Upregulated
14747	Intrônico asPI	89846	FGD3	FYVE, RhoGEF and PH domain containing 3	Upregulated
15410	Intrônico asPI	9819	TSC22D2	TSC22 domain family, member 2	Upregulated
15689	Intrônico asPI	55964	3-Sep	septin 3	Upregulated
16119	Intrônico asPI	23305	ACSL6	acyl-CoA synthetase long-chain family member 6	Upregulated
16215	Intrônico asPI	29128	UHRF1	ubiquitin-like, containing PHD and RING finger domains, 1	Upregulated
18880	Intrônico asPI	57189	-	-	Upregulated

21199	Intrônico asPI	5047	PAEP	progesterone-associated endometrial protein (placental protein 14, pregnancy-associated endometrial alpha-2- globulin, alpha uterine protein)	Upregulated
21966	Intrônico asPI	64598	MOSPD3	motile sperm domain containing 3	Upregulated
21997	Intrônico asPI	10589	DRAP1	DR1-associated protein 1 (negative cofactor 2 alpha)	Upregulated
23079	Intrônico asPI	374659	-	-	Upregulated
23247	Intrônico asPI	7701	ZNF142	zinc finger protein 142 (clone pHZ-49)	Upregulated
24224	Intrônico asPI	84812	PLCD4	phospholipase C, delta 4	Upregulated
24683	Intrônico asPI	9271	PIWIL1	piwi-like 1 (Drosophila)	Upregulated
28191	Intrônico asPI	7633	ZNF79	zinc finger protein 79 (pT7)	Upregulated
28941	Intrônico asPI	617	BCS1L	BCS1-like (yeast) sterol regulatory element	Upregulated
29019	Intrônico asPI	6721	SREBF2	binding transcription factor 2	Upregulated
29251	Intrônico asPI	213	ALB	albumin	Upregulated
29734	Intrônico asPI	8570	KHSRP	KH-type splicing regulatory protein (FUSE binding protein 2)	Upregulated
30269	Intrônico asPI	3337	DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1	Upregulated
31096	Intrônico asPI	2335	FN1	fibronectin 1	Upregulated
31150	Intrônico asPI	29997	GLTSCR2	glioma tumor suppressor candidate region gene 2	Upregulated
33308	Intrônico asPI	54531	KIAA1193	KIAA1193	Upregulated
33731	Intrônico asPI	23352	-	-	Upregulated
35002	Intrônico asPI	4736	RPL10A	ribosomal protein L10a ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	Upregulated
35012	Intrônico asPI	1994	ELAVL1		
36243	Intrônico asPI	3294	HSD17B2	hydroxysteroid (17-beta) dehydrogenase 2	Upregulated
36767	Intrônico asPI	1106	CHD2	chromodomain helicase DNA binding protein 2	Upregulated
36863	Intrônico asPI	2766	GMPR	guanosine monophosphate reductase	Upregulated
38096	Intrônico asPI	3675	ITGA3	integrin, alpha 3 (antigen CD49C, alpha 3 subunit of VLA-3 receptor)	Upregulated
38259	Intrônico asPI	26224	FBXL3	F-box and leucine-rich repeat protein 3	Upregulated
39936	Intrônico asPI	6901	TAZ	tafazzin (cardiomyopathy, dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome)	Upregulated
40053	Intrônico asPI	3859	KRT12	keratin 12 (Meesmann corneal dystrophy)	Upregulated
40067	Intrônico asPI	166	AES	amino-terminal enhancer of split	Upregulated
40265	Intrônico asPI	2119	ETV5	ets variant gene 5 (ets- related molecule)	Upregulated
40988	Intrônico asPI	1400	CRMP1	collapsin response mediator protein 1	Upregulated
41908	Intrônico asPI	3659	IRF1	interferon regulatory factor 1	Upregulated
42189	Intrônico asPI	9910	RABGAP1L	RAB GTPase activating protein 1-like	Upregulated

42319	Intrônico asPI	3899	LAF4	lymphoid nuclear protein related to AF4	Upregulated
42596	Intrônico asPI	10768	AHCYL1	S-adenosylhomocysteine hydrolase-like 1	Upregulated
43464	Intrônico asPI	405	ARNT	aryl hydrocarbon receptor nuclear translocator	Upregulated
43556	Intrônico asPI	3661	IRF3	interferon regulatory factor 3	Upregulated
44065	Intrônico asPI	4929	NR4A2	nuclear receptor subfamily 4, group A, member 2	Upregulated
43	Intrônico asPI	51031	C17orf25	chromosome 17 open reading frame 25	Downregulated
45	Intrônico asPI	27332	ZNF638	zinc finger protein 638	Downregulated
196	Intrônico asPI	2617	GARS	glycyl-tRNA synthetase	Downregulated
221	Intrônico asPI	124401	-	-	Downregulated
279	Intrônico asPI	7407	VARs2	valyl-tRNA synthetase 2	Downregulated
327	Intrônico asPI	90139	TSPAN18	tetraspanin 18	Downregulated
422	Intrônico asPI	219402	-	-	Downregulated
568	Intrônico asPI	23085	RAB6IP2	RAB6 interacting protein 2	Downregulated
857	Intrônico asPI	55904	MLL5	myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila)	Downregulated
922	Intrônico asPI	1287	COL4A5	collagen, type IV, alpha 5 (Alport syndrome)	Downregulated
1052	Intrônico asPI	81691	-	-	Downregulated
1100	Intrônico asPI	8530	CST7	cystatin F (leukocystatin)	Downregulated
1131	Intrônico asPI	533	ATP6V0B	ATPase, H+ transporting, lysosomal 21kDa, V0 subunit c"	Downregulated
1269	Intrônico asPI	55904	MLL5	myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila)	Downregulated
1307	Intrônico asPI	8291	DYSF	dysferlin, limb girdle muscular dystrophy 2B (autosomal recessive)	Downregulated
1419	Intrônico asPI	5792	PTPRF	protein tyrosine phosphatase, receptor type, F	Downregulated
1572	Intrônico asPI	64760	RAI16	retinoic acid induced 16	Downregulated
1589	Intrônico asPI	29855	UBN1	ubiquitin 1	Downregulated
1604	Intrônico asPI	10480	-	-	Downregulated
1614	Intrônico asPI	8672	EIF4G3	eukaryotic translation initiation factor 4 gamma, 3	Downregulated
1709	Intrônico asPI	10131	TRAP1	TNF receptor-associated protein 1	Downregulated
1723	Intrônico asPI	7030	TFE3	transcription factor binding to IGHE enhancer 3	Downregulated
1750	Intrônico asPI	309	ANXA6	annexin A6	Downregulated
1759	Intrônico asPI	283989	TSEN54	tRNA splicing endonuclease 54 homolog (SEN54, S. cerevisiae)	Downregulated
1777	Intrônico asPI	79083	MLPH	melanophilin	Downregulated
1817	Intrônico asPI	9354	UBE4A	ubiquitination factor E4A (UFD2 homolog, yeast)	Downregulated
1888	Intrônico asPI	55748	CNDP2	CNDP dipeptidase 2 (metallopeptidase M20 family)	Downregulated
1925	Intrônico asPI	8291	DYSF	dysferlin, limb girdle muscular dystrophy 2B (autosomal recessive)	Downregulated

1955	Intrônico asPI	535	ATP6V0A1	ATPase, H ⁺ transporting, lysosomal V0 subunit a isoform 1	Downregulated
1966	Intrônico asPI	253827	MSRB3	methionine sulfoxide reductase B3	Downregulated
1987	Intrônico asPI	57572	DOCK6	dedicator of cytokinesis 6	Downregulated
1990	Intrônico asPI	55902	ACAS2	acetyl-Coenzyme A synthetase 2 (ADP forming)	Downregulated
2039	Intrônico asPI	5440	POLR2K	polymerase (RNA) II (DNA directed) polypeptide K, 7.0kDa	Downregulated
2048	Intrônico asPI	3093	HIP2	huntingtin interacting protein 2	Downregulated
2178	Intrônico asPI	124220	-	-	Downregulated
2222	Intrônico asPI	10480	-	-	Downregulated
2310	Intrônico asPI	29889	GNL2	guanine nucleotide binding protein-like 2 (nucleolar)	Downregulated
2335	Intrônico asPI	8697	CDC23	CDC23 (cell division cycle 23, yeast, homolog)	Downregulated
2383	Intrônico asPI	132989	-	-	Downregulated
2400	Intrônico asPI	57226	-	-	Downregulated
2439	Intrônico asPI	8522	GAS7	growth arrest-specific 7	Downregulated
2533	Intrônico asPI	10133	OPTN	optineurin	Downregulated
2608	Intrônico asPI	55902	ACAS2	acetyl-Coenzyme A synthetase 2 (ADP forming)	Downregulated
2805	Intrônico asPI	79979	CXorf34	chromosome X open reading frame 34	Downregulated
2924	Intrônico asPI	54058	C21orf58	chromosome 21 open reading frame 58	Downregulated
2952	Intrônico asPI	8883	APPBP1	amyloid beta precursor protein binding protein 1	Downregulated
3045	Intrônico asPI	10629	TAF6L	TAF6-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa	Downregulated
3127	Intrônico asPI	55143	CDCA8	cell division cycle associated 8	Downregulated
3219	Intrônico asPI	79081	-	-	Downregulated
3318	Intrônico asPI	81693	AMN	amniotless homolog (mouse)	Downregulated
3340	Intrônico asPI	29894	CPSF1	cleavage and polyadenylation specific factor 1, 160kDa	Downregulated
3421	Intrônico asPI	81618	ITM2C	integral membrane protein 2C	Downregulated
3424	Intrônico asPI	79735	TBC1D17	TBC1 domain family, member 17	Downregulated
3482	Intrônico asPI	5214	PFKP	phosphofructokinase, platelet	Downregulated
3550	Intrônico asPI	25855	BRMS1	breast cancer metastasis suppressor 1	Downregulated
3577	Intrônico asPI	7035	TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	Downregulated
3578	Intrônico asPI	6787	NEK4	NIMA (never in mitosis gene a)-related kinase 4	Downregulated
3580	Intrônico asPI	6311	ATXN2	ataxin 2	Downregulated
3635	Intrônico asPI	57568	SIPA1L2	signal-induced proliferation-associated 1 like 2	Downregulated
3658	Intrônico asPI	23081	JMJD2C	jumonji domain containing 2C	Downregulated

3660	Intrônico asPI	11336	SEC6L1	SEC6-like 1 (S. cerevisiae)	Downregulated
3770	Intrônico asPI	9970	NR1I3	nuclear receptor subfamily 1, group I, member 3	Downregulated
3869	Intrônico asPI	10628	TXNIP	thioredoxin interacting protein	Downregulated
3955	Intrônico asPI	51593	-	-	Downregulated
3984	Intrônico asPI	8536	CAMK1	calcium/calmodulin-dependent protein kinase I	Downregulated
4143	Intrônico asPI	80314	EPC1	enhancer of polycomb homolog 1 (Drosophila)	Downregulated
4173	Intrônico asPI	23092	ARHGAP26	Rho GTPase activating protein 26	Downregulated
4271	Intrônico asPI	25836	NIPBL	Nipped-B homolog (Drosophila)	Downregulated
4289	Intrônico asPI	9352	TXNL1	thioredoxin-like 1	Downregulated
4360	Intrônico asPI	55749	CCAR1	cell division cycle and apoptosis regulator 1	Downregulated
4395	Intrônico asPI	8703	B4GALT3	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 3	Downregulated
4485	Intrônico asPI	11103	HRB2	HIV-1 rev binding protein 2	Downregulated
4605	Intrônico asPI	7414	VCL	vinculin	Downregulated
4821	Intrônico asPI	5250	SLC25A3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3	Downregulated
4887	Intrônico asPI	26528	DAZAP1	DAZ associated protein 1	Downregulated
4992	Intrônico asPI	25862	USP49	ubiquitin specific protease 49	Downregulated
5098	Intrônico asPI	23275	POFUT2	protein O-fucosyltransferase 2	Downregulated
5103	Intrônico asPI	11103	HRB2	HIV-1 rev binding protein 2	Downregulated
5228	Intrônico asPI	6314	ATXN7	ataxin 7	Downregulated
5393	Intrônico asPI	55145	THAP1	THAP domain containing, apoptosis associated protein 1	Downregulated
5411	Intrônico asPI	22801	ITGA11	integrin, alpha 11	Downregulated
5522	Intrônico asPI	9832	-	-	Downregulated
5716	Intrônico asPI	23275	POFUT2	protein O-fucosyltransferase 2	Downregulated
5832	Intrônico asPI	9703	-	-	Downregulated
5908	Intrônico asPI	55127	-	-	Downregulated
6000	Intrônico asPI	79109	MAPKAP1	mitogen-activated protein kinase associated protein 1	Downregulated
6073	Intrônico asPI	996	CDC27	cell division cycle 27	Downregulated
6231	Intrônico asPI	23294	ANKS1	ankyrin repeat and sterile alpha motif domain containing 1	Downregulated
6247	Intrônico asPI	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
6419	Intrônico asPI	55914	ERBB2IP	erbb2 interacting protein	Downregulated
6427	Intrônico asPI	51602	-	-	Downregulated
6568	Intrônico asPI	5965	RECQL	RecQ protein-like (DNA helicase Q1-like)	Downregulated
6639	Intrônico asPI	26057	ANKRD17	ankyrin repeat domain 17	Downregulated
6855	Intrônico asPI	10919	EHMT2	euchromatic histone-lysine N-methyltransferase 2	Downregulated
6862	Intrônico asPI	9703	-	-	Downregulated
6930	Intrônico asPI	57801	HES4	hairy and enhancer of split 4 (Drosophila)	Downregulated

7045	Intrônico asPI	51602	-	-	Downregulated
7254	Intrônico asPI	29904	EEF2K	eukaryotic elongation factor 2 kinase	Downregulated
7277	Intrônico asPI	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
7300	Intrônico asPI	3491	CYR61	cysteine-rich, angiogenic inducer, 61	Downregulated
7371	Intrônico asPI	10625	IVNS1ABP	influenza virus NS1A binding protein	Downregulated
7549	Intrônico asPI	57561	ARRDC3	arrestin domain containing 3	Downregulated
7690	Intrônico asPI	8888	MCM3AP	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae) associated protein	Downregulated
7864	Intrônico asPI	55329	-	-	Downregulated
7876	Intrônico asPI	25874	BRP44	brain protein 44	Downregulated
7925	Intrônico asPI	1580	CYP4B1	cytochrome P450, family 4, subfamily B, polypeptide 1	Downregulated
7952	Intrônico asPI	83992	CTTNBP2	cortactin binding protein 2	Downregulated
7955	Intrônico asPI	79969	C6orf134	chromosome 6 open reading frame 134	Downregulated
8181	Intrônico asPI	29789	-	-	Downregulated
8269	Intrônico asPI	58494	JAM2	junctional adhesion molecule 2	Downregulated
8271	Intrônico asPI	57407	-	-	Downregulated
8474	Intrônico asPI	64097	EPB41L4A	erythrocyte membrane protein band 4.1 like 4A	Downregulated
8583	Intrônico asPI	55735	DNAJC11	DnaJ (Hsp40) homolog, subfamily C, member 11	Downregulated
8627	Intrônico asPI	6296	SAH	SA hypertension-associated homolog (rat)	Downregulated
8700	Intrônico asPI	25874	BRP44	brain protein 44	Downregulated
8788	Intrônico asPI	55872	PBK	PDZ binding kinase	Downregulated
8883	Intrônico asPI	79799	-	-	Downregulated
8984	Intrônico asPI	80232	WDR26	WD repeat domain 26	Downregulated
9003	Intrônico asPI	51256	TBC1D7	TBC1 domain family, member 7	Downregulated
9296	Intrônico asPI	79139	DERL1	Der1-like domain family, member 1	Downregulated
9489	Intrônico asPI	147172	-	-	Downregulated
9632	Intrônico asPI	23078	-	-	Downregulated
9682	Intrônico asPI	71	ACTG1	actin, gamma 1	Downregulated
9720	Intrônico asPI	54970	TTC12	tetratricopeptide repeat domain 12	Downregulated
9723	Intrônico asPI	51604	PIGT	phosphatidylinositol glycan, class T	Downregulated
9814	Intrônico asPI	57794	SF4	splicing factor 4	Downregulated
10064	Intrônico asPI	7392	USF2	upstream transcription factor 2, c-fos interacting	Downregulated
10287	Intrônico asPI	3431	SP110	SP110 nuclear body protein	Downregulated
10387	Intrônico asPI	4241	MFI2	antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5	Downregulated
10491	Intrônico asPI	3978	LIG1	ligase I, DNA, ATP-dependent	Downregulated
10538	Intrônico asPI	56957	ZA20D1	zinc finger, A20 domain containing 1	Downregulated
10911	Intrônico asPI	1783	DNCL12	dynein, cytoplasmic, light intermediate polypeptide 2	Downregulated
11001	Intrônico asPI	5265	SERPINA1	serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 1	Downregulated

11020	Intrônico asPI	317	APAF1	apoptotic protease activating factor	Downregulated
11270	Intrônico asPI	50853	VILL	villin-like	Downregulated
11536	Intrônico asPI	59	ACTA2	actin, alpha 2, smooth muscle, aorta	Downregulated
12042	Intrônico asPI	2224	FDPS	farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, dimethylallyltranstransferase, geranyltranstransferase)	Downregulated
12291	Intrônico asPI	55732	C1orf112	chromosome 1 open reading frame 112	Downregulated
12306	Intrônico asPI	23545	ATP6V0A2	ATPase, H+ transporting, lysosomal V0 subunit a isoform 2	Downregulated
12308	Intrônico asPI	23271	CAMSAP1L1	calmodulin regulated spectrin-associated protein 1-like 1	Downregulated
12661	Intrônico asPI	2064	ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)	Downregulated
12696	Intrônico asPI	64753	-	-	Downregulated
12865	Intrônico asPI	2648	GCN5L2	GCN5 general control of amino-acid synthesis 5-like 2 (yeast)	Downregulated
13180	Intrônico asPI	958	CD40	CD40 antigen (TNF receptor superfamily member 5)	Downregulated
13319	Intrônico asPI	56834	-	-	Downregulated
13371	Intrônico asPI	5054	SERPINE1	serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	Downregulated
13568	Intrônico asPI	6950	TCP1	t-complex 1	Downregulated
13668	Intrônico asPI	7881	KCNAB1	potassium voltage-gated channel, shaker-related subfamily, beta member 1	Downregulated
13682	Intrônico asPI	4649	MYO9A	myosin IXA	Downregulated
13793	Intrônico asPI	2181	ACSL3	acyl-CoA synthetase long-chain family member 3	Downregulated
13839	Intrônico asPI	55178	RNMTL1	RNA methyltransferase like 1	Downregulated
13986	Intrônico asPI	5590	PRKCZ	protein kinase C, zeta	Downregulated
13989	Intrônico asPI	5054	SERPINE1	serine (or cysteine) proteinase inhibitor, clade E (nexin, plasminogen activator inhibitor type 1), member 1	Downregulated
14186	Intrônico asPI	6950	TCP1	t-complex 1	Downregulated
14200	Intrônico asPI	3705	ITPK1	inositol 1,3,4-triphosphate 5/6 kinase	Downregulated
14255	Intrônico asPI	51631	LUC7L2	LUC7-like 2 (S. cerevisiae)	Downregulated
14301	Intrônico asPI	4254	KITLG	KIT ligand	Downregulated
14620	Intrônico asPI	1495	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa	Downregulated
14814	Intrônico asPI	4781	NFIB	nuclear factor I/B	Downregulated
14967	Intrônico asPI	56731	SLC2A4RG	SLC2A4 regulator	Downregulated
15051	Intrônico asPI	147700	-	-	Downregulated

15134	Intrônico asPI	1828	DSG1	desmoglein 1	Downregulated
15216	Intrônico asPI	6949	TCOF1	Treacher Collins-Franceschetti syndrome 1	Downregulated
15324	Intrônico asPI	5801	PTPRR	protein tyrosine phosphatase, receptor type, R	Downregulated
15478	Intrônico asPI	64147	KIF9	kinesin family member 9	Downregulated
15514	Intrônico asPI	9728	-	-	Downregulated
15835	Intrônico asPI	6741	SSB	Sjogren syndrome antigen B (autoantigen La)	Downregulated
15856	Intrônico asPI	1495	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa	Downregulated
16312	Intrônico asPI	54991	-	-	Downregulated
16347	Intrônico asPI	7436	VLDLR	very low density lipoprotein receptor	Downregulated
16375	Intrônico asPI	593	BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)	Downregulated
16450	Intrônico asPI	7384	UQCRC1	ubiquinol-cytochrome c reductase core protein I	Downregulated
16466	Intrônico asPI	3698	ITI1H2	inter-alpha (globulin) inhibitor H2	Downregulated
16574	Intrônico asPI	2244	FGB	fibrinogen, B beta polypeptide	Downregulated
16739	Intrônico asPI	23098	SARM1	sterile alpha and TIR motif containing 1	Downregulated
16899	Intrônico asPI	353174	-	-	Downregulated
16924	Intrônico asPI	56984	TNFSF5IP1	tumor necrosis factor superfamily, member 5-induced protein 1	Downregulated
17128	Intrônico asPI	57609	-	-	Downregulated
17193	Intrônico asPI	2067	ERCC1	excision repair cross-complementing rodent repair deficiency, complementation group 1 (includes overlapping antisense sequence)	Downregulated
17243	Intrônico asPI	51232	CRIM1	cysteine-rich motor neuron 1	Downregulated
17269	Intrônico asPI	8844	KSR	kinase suppressor of ras	Downregulated
17368	Intrônico asPI	9730	-	-	Downregulated
17435	Intrônico asPI	63894	C14orf133	chromosome 14 open reading frame 133	Downregulated
17453	Intrônico asPI	26298	EHF	ets homologous factor	Downregulated
17510	Intrônico asPI	50	ACO2	aconitase 2, mitochondrial	Downregulated
17557	Intrônico asPI	26065	FAM61A	family with sequence similarity 61, member A	Downregulated
17963	Intrônico asPI	51642	MRPL48	mitochondrial ribosomal protein L48	Downregulated
18341	Intrônico asPI	373156	GSTK1	glutathione S-transferase kappa 1	Downregulated
18366	Intrônico asPI	56986	-	-	Downregulated
18416	Intrônico asPI	5501	PPP1CC	protein phosphatase 1, catalytic subunit, gamma isoform	Downregulated
18511	Intrônico asPI	7150	TOP1	topoisomerase (DNA) I	Downregulated
18681	Intrônico asPI	54887	C6orf107	chromosome 6 open reading frame 107	Downregulated
18831	Intrônico asPI	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Downregulated
18885	Intrônico asPI	55276	PGM2	phosphoglucomutase 2	Downregulated

18928	Intrônico asPI	5921	RASA1	RAS p21 protein activator (GTPase activating protein) 1	Downregulated
18952	Intrônico asPI	50	ACO2	aconitase 2, mitochondrial	Downregulated
18998	Intrônico asPI	27039	PKD2L2	polycystic kidney disease 2- like 2	Downregulated
19179	Intrônico asPI	84167	-	-	Downregulated
19222	Intrônico asPI	9731	-	-	Downregulated
19268	Intrônico asPI	338657	-	-	Downregulated
19299	Intrônico asPI	54887	C6orf107	chromosome 6 open reading frame 107	Downregulated
19347	Intrônico asPI	4547	MTP	microsomal triglyceride transfer protein (large polypeptide, 88kDa)	Downregulated
19511	Intrônico asPI	29121	CLEC2D	C-type lectin superfamily 2, member D	Downregulated
19516	Intrônico asPI	23524	SRRM2	serine/arginine repetitive matrix 2	Downregulated
19765	Intrônico asPI	2803	GOLGA4	golgi autoantigen, golgin subfamily a, 4	Downregulated
19785	Intrônico asPI	255967	-	-	Downregulated
19874	Intrônico asPI	1314	COPA	coatamer protein complex, subunit alpha	Downregulated
19886	Intrônico asPI	338657	-	-	Downregulated
20278	Intrônico asPI	3607	FOXK2	forkhead box K2	Downregulated
20360	Intrônico asPI	8650	NUMB	numb homolog (Drosophila)	Downregulated
20456	Intrônico asPI	10000	AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	Downregulated
20654	Intrônico asPI	23019	CNOT1	CCR4-NOT transcription complex, subunit 1	Downregulated
20743	Intrônico asPI	51780	JMJD1B	jumonji domain containing 1B	Downregulated
20790	Intrônico asPI	4174	MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	Downregulated
21073	Intrônico asPI	10153	CEBPZ	CCAAT/enhancer binding protein zeta	Downregulated
21206	Intrônico asPI	3065	HDAC1	histone deacetylase 1	Downregulated
21318	Intrônico asPI	800	CALD1	caldesmon 1	Downregulated
21505	Intrônico asPI	5685	PSMA4	proteasome (prosome, macropain) subunit, alpha type, 4	Downregulated
21582	Intrônico asPI	11273	ATXN2L	ataxin 2-like	Downregulated
21615	Intrônico asPI	3934	LCN2	lipocalin 2 (oncogene 24p3)	Downregulated
21773	Intrônico asPI	51776	-	-	Downregulated
22009	Intrônico asPI	8481	OFD1	oral-facial-digital syndrome 1	Downregulated
22392	Intrônico asPI	51497	TH1L	TH1-like (Drosophila)	Downregulated
22397	Intrônico asPI	26286	ARFGAP3	ADP-ribosylation factor GTPase activating protein 3	Downregulated
22490	Intrônico asPI	55357	TBC1D2	TBC1 domain family, member 2	Downregulated
22493	Intrônico asPI	54806	AHI1	Abelson helper integration site	Downregulated
22514	Intrônico asPI	10346	TRIM22	tripartite motif-containing 22	Downregulated
22524	Intrônico asPI	8562	DENR	density-regulated protein	Downregulated
22577	Intrônico asPI	84292	-	-	Downregulated
22649	Intrônico asPI	2800	GOLGA1	golgi autoantigen, golgin subfamily a, 1	Downregulated
22786	Intrônico asPI	80208	-	-	Downregulated
23317	Intrônico asPI	54806	AHI1	Abelson helper integration site	Downregulated

23480	Intrônico asPI	908	CCT6A	chaperonin containing TCP1, subunit 6A (zeta 1)	Downregulated
23832	Intrônico asPI	54581	SCAND2	SCAN domain containing 2	Downregulated
23833	Intrônico asPI	51773	HBXAP	hepatitis B virus x associated protein	Downregulated
24298	Intrônico asPI	2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma)	Downregulated
24371	Intrônico asPI	9879	DDX46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	Downregulated
24404	Intrônico asPI	1859	DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	Downregulated
24455	Intrônico asPI	29105	-	-	Downregulated
24493	Intrônico asPI	5373	PMM2	phosphomannomutase 2	Downregulated
24657	Intrônico asPI	51773	HBXAP	hepatitis B virus x associated protein	Downregulated
24802	Intrônico asPI	5519	PPP2R1B	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	Downregulated
24893	Intrônico asPI	8475	EPR1	effector cell protease receptor 1	Downregulated
24912	Intrônico asPI	3684	ITGAM	integrin, alpha M (complement component receptor 3, alpha; also known as CD11b (p170), macrophage antigen alpha polypeptide)	Downregulated
24964	Intrônico asPI	55013	-	-	Downregulated
25111	Intrônico asPI	5373	PMM2	phosphomannomutase 2	Downregulated
25179	Intrônico asPI	26112	-	-	Downregulated
25188	Intrônico asPI	11176	BAZ2A	bromodomain adjacent to zinc finger domain, 2A	Downregulated
25347	Intrônico asPI	256126	-	-	Downregulated
25356	Intrônico asPI	90957	DHX57	DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	Downregulated
25372	Intrônico asPI	55779	WDR52	WD repeat domain 52	Downregulated
25386	Intrônico asPI	25909	-	-	Downregulated
25460	Intrônico asPI	84812	PLCD4	phospholipase C, delta 4	Downregulated
25569	Intrônico asPI	79831	-	-	Downregulated
25682	Intrônico asPI	55565	-	-	Downregulated
25764	Intrônico asPI	137964	-	-	Downregulated
25809	Intrônico asPI	10524	HTATIP	HIV-1 Tat interacting protein, 60kDa	Downregulated
25864	Intrônico asPI	166647	GPR125	G protein-coupled receptor 125	Downregulated
26078	Intrônico asPI	84812	PLCD4	phospholipase C, delta 4	Downregulated
26097	Intrônico asPI	54873	PALMD	palmdelphin	Downregulated
26099	Intrônico asPI	51763	-	-	Downregulated
26160	Intrônico asPI	471	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	Downregulated
26171	Intrônico asPI	256281	NUDT14	nudix (nucleoside diphosphate linked moiety X)-type motif 14	Downregulated
26311	Intrônico asPI	26271	FBXO5	F-box protein 5	Downregulated
26393	Intrônico asPI	79832	-	-	Downregulated
26510	Intrônico asPI	54552	GNL3L	guanine nucleotide binding protein-like 3 (nucleolar)-like	Downregulated
26515	Intrônico asPI	29102	RNASEN	ribonuclease III, nuclear	Downregulated
26607	Intrônico asPI	56063	C1orf91	chromosome 1 open reading frame 91	Downregulated

26725	Intrônico asPI	25792	CIZ1	CDKN1A interacting zinc finger protein 1	Downregulated
26773	Intrônico asPI	1741	DLG3	discs, large homolog 3 (neuroendocrine-dlg, Drosophila)	Downregulated
26791	Intrônico asPI	196515	-	-	Downregulated
26794	Intrônico asPI	137970	UNC5D	unc-5 homolog D (C. elegans)	Downregulated
27006	Intrônico asPI	84572	GNPTG	N-acetylglucosamine-1-phosphotransferase, gamma subunit	Downregulated
27018	Intrônico asPI	57002	C7orf36	chromosome 7 open reading frame 36	Downregulated
27095	Intrônico asPI	399671	-	-	Downregulated
27174	Intrônico asPI	4763	NF1	neurofibromin 1 (neurofibromatosis, von Recklinghausen disease, Watson disease)	Downregulated
27233	Intrônico asPI	51668	C1orf41	chromosome 1 open reading frame 41	Downregulated
27348	Intrônico asPI	23053	KIAA0913	KIAA0913	Downregulated
27406	Intrônico asPI	284992	-	-	Downregulated
27498	Intrônico asPI	817	CAMK2D	calcium/calmodulin-dependent protein kinase (CaM kinase) II delta	Downregulated
27528	Intrônico asPI	64478	CSMD1	CUB and Sushi multiple domains 1	Downregulated
27599	Intrônico asPI	1164	CKS2	CDC28 protein kinase regulatory subunit 2	Downregulated
27623	Intrônico asPI	85007	-	-	Downregulated
27724	Intrônico asPI	92140	-	-	Downregulated
27736	Intrônico asPI	57706	KIAA1608	KIAA1608	Downregulated
27744	Intrônico asPI	55088	C10orf118	chromosome 10 open reading frame 118	Downregulated
27805	Intrônico asPI	1164	CKS2	CDC28 protein kinase regulatory subunit 2	Downregulated
27938	Intrônico asPI	79705	LRRK1	leucine-rich repeat kinase 1	Downregulated
27959	Intrônico asPI	26266	SLC13A4	solute carrier family 13 (sodium/sulfate symporters), member 4	Downregulated
27965	Intrônico asPI	23160	WDR43	WD repeat domain 43	Downregulated
28049	Intrônico asPI	56159	TEX11	testis expressed sequence 11	Downregulated
28231	Intrônico asPI	256471	-	-	Downregulated
28380	Intrônico asPI	11253	MAN1B1	mannosidase, alpha, class 1B, member 1	Downregulated
28450	Intrônico asPI	83549	UCK1	uridine-cytidine kinase 1	Downregulated
28460	Intrônico asPI	57018	CCNL1	cyclin L1	Downregulated
28537	Intrônico asPI	399512	SLC25A35	solute carrier family 25, member 35	Downregulated
28572	Intrônico asPI	51479	ANKFY1	ankyrin repeat and FYVE domain containing 1	Downregulated
28579	Intrônico asPI	25776	PGEA1	PKD2 interactor, golgi and endoplasmic reticulum associated 1	Downregulated
28669	Intrônico asPI	55661	DDX27	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	Downregulated
28673	Intrônico asPI	54812	-	-	Downregulated
28681	Intrônico asPI	26122	EPC2	enhancer of polycomb homolog 2 (Drosophila)	Downregulated
28775	Intrônico asPI	54862	-	-	Downregulated
28967	Intrônico asPI	79932	KIAA0319L	KIAA0319-like	Downregulated
28977	Intrônico asPI	55704	KIAA1212	KIAA1212	Downregulated
29059	Intrônico asPI	149041	-	-	Downregulated
29062	Intrônico asPI	116985	CENTD2	centaurin, delta 2	Downregulated
29178	Intrônico asPI	57704	GBA2	glucosidase, beta (bile acid) 2	Downregulated
29404	Intrônico asPI	23514	-	-	Downregulated

29475	Intrônico asPI	112495	C6orf51	chromosome 6 open reading frame 51	Downregulated
29592	Intrônico asPI	57157	PHTF2	putative homeodomain transcription factor 2	Downregulated
29635	Intrônico asPI	7110	TMF1	TATA element modulatory factor 1	Downregulated
29680	Intrônico asPI	116985	CENTD2	centaurin, delta 2 polymerase (DNA-directed), delta interacting protein 3	Downregulated
29787	Intrônico asPI	84271	POLDIP3	-	Downregulated
29906	Intrônico asPI	55374	-	-	Downregulated
29992	Intrônico asPI	84765	ZNF577	zinc finger protein 577	Downregulated
30017	Intrônico asPI	29087	-	-	Downregulated
30091	Intrônico asPI	126526	-	-	Downregulated
30095	Intrônico asPI	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
30116	Intrônico asPI	54462	KIAA1128	KIAA1128	Downregulated
30291	Intrônico asPI	256949	-	-	Downregulated
30313	Intrônico asPI	57472	CNOT6	CCR4-NOT transcription complex, subunit 6	Downregulated
30616	Intrônico asPI	79699	ZYG11	zyg-11 homolog (C. elegans)	Downregulated
30740	Intrônico asPI	27063	ANKRD1	ankyrin repeat domain 1 (cardiac muscle)	Downregulated
30812	Intrônico asPI	120863	DEPDC4	DEP domain containing 4	Downregulated
30908	Intrônico asPI	285148	-	-	Downregulated
31021	Intrônico asPI	88455	ANKRD13	ankyrin repeat domain 13	Downregulated
31024	Intrônico asPI	83733	SLC25A18	solute carrier family 25 (mitochondrial carrier), member 18	Downregulated
31053	Intrônico asPI	23352	-	-	Downregulated
31157	Intrônico asPI	23326	USP22	ubiquitin specific protease 22	Downregulated
31331	Intrônico asPI	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
31549	Intrônico asPI	57473	-	-	Downregulated
31572	Intrônico asPI	23028	AOF2	amine oxidase (flavin containing) domain 2	Downregulated
31578	Intrônico asPI	10392	CARD4	caspase recruitment domain family, member 4	Downregulated
31661	Intrônico asPI	51742	ARID4B	AT rich interactive domain 4B (RBP1- like)	Downregulated
31756	Intrônico asPI	57038	RARSL	arginyl-tRNA synthetase-like	Downregulated
31872	Intrônico asPI	27127	SMC1L2	SMC1 structural maintenance of chromosomes 1-like 2 (yeast)	Downregulated
31974	Intrônico asPI	30000	TNPO2	transportin 2 (importin 3, karyopherin beta 2b)	Downregulated
32196	Intrônico asPI	10392	CARD4	caspase recruitment domain family, member 4	Downregulated
32288	Intrônico asPI	23512	SUZ12	suppressor of zeste 12 homolog (Drosophila)	Downregulated
32456	Intrônico asPI	164684	-	-	Downregulated
32464	Intrônico asPI	84750	FUT10	fucosyltransferase 10 (alpha (1,3) fucosyltransferase)	Downregulated
32570	Intrônico asPI	83607	-	-	Downregulated
32706	Intrônico asPI	11235	PDCD10	programmed cell death 10	Downregulated
32707	Intrônico asPI	11051	CPSF5	cleavage and polyadenylation specific factor 5, 25 kDa	Downregulated

32785	Intrônico asPI	57475	PLEKHH1	pleckstrin homology domain containing, family H (with MyTH4 domain) member 1	Downregulated
32911	Intrônico asPI	22906	-	-	Downregulated
32980	Intrônico asPI	84632	KIAA1914	KIAA1914	Downregulated
33024	Intrônico asPI	9750	C6orf32	chromosome 6 open reading frame 32	Downregulated
33106	Intrônico asPI	50618	ITSN2	intersectin 2	Downregulated
33179	Intrônico asPI	149428	BNIP1	BCL2/adenovirus E1B 19kD interacting protein like	Downregulated
33198	Intrônico asPI	57060	PCBP4	poly(rC) binding protein 4	Downregulated
33464	Intrônico asPI	2934	GSN	gelsolin (amyloidosis, Finnish type)	Downregulated
33511	Intrônico asPI	55247	NEIL3	nei endonuclease VIII-like 3 (E. coli)	Downregulated
33565	Intrônico asPI	3309	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	Downregulated
33600	Intrônico asPI	83608	C18orf21	chromosome 18 open reading frame 21	Downregulated
33636	Intrônico asPI	10734	STAG3	stromal antigen 3	Downregulated
33821	Intrônico asPI	55206	SBNO1	sno, strawberry notch homolog 1 (Drosophila)	Downregulated
33855	Intrônico asPI	8419	BFSP2	beaded filament structural protein 2, phakinin	Downregulated
33897	Intrônico asPI	200894	ARL2L1	ADP-ribosylation factor-like 2-like 1	Downregulated
33932	Intrônico asPI	27125	-	-	Downregulated
34007	Intrônico asPI	113246	-	-	Downregulated
34357	Intrônico asPI	10576	CCT2	chaperonin containing TCP1, subunit 2 (beta)	Downregulated
34445	Intrônico asPI	51123	-	-	Downregulated
34457	Intrônico asPI	22872	SEC31L1	SEC31-like 1 (S. cerevisiae)	Downregulated
34459	Intrônico asPI	11010	GLIPR1	GLI pathogenesis-related 1 (glioma)	Downregulated
34828	Intrônico asPI	138639	PTPDC1	protein tyrosine phosphatase domain containing 1	Downregulated
34855	Intrônico asPI	51700	-	-	Downregulated
34937	Intrônico asPI	84260	-	-	Downregulated
34949	Intrônico asPI	56474	CTPS2	CTP synthase II	Downregulated
35040	Intrônico asPI	84645	C22orf23	chromosome 22 open reading frame 23	Downregulated
35047	Intrônico asPI	64960	MRPS15	mitochondrial ribosomal protein S15	Downregulated
35102	Intrônico asPI	5527	PPP2R5C	protein phosphatase 2, regulatory subunit B (B56), gamma isoform	Downregulated
35154	Intrônico asPI	57153	-	-	Downregulated
35161	Intrônico asPI	54843	SYTL2	synaptotagmin-like 2	Downregulated
35555	Intrônico asPI	84260	-	-	Downregulated
35584	Intrônico asPI	23500	DAAM2	dishevelled associated activator of morphogenesis 2	Downregulated
35646	Intrônico asPI	285331	-	-	Downregulated
35656	Intrônico asPI	91252	SLC39A13	solute carrier family 39 (zinc transporter), member 13	Downregulated
35695	Intrônico asPI	11010	GLIPR1	GLI pathogenesis-related 1 (glioma)	Downregulated
35783	Intrônico asPI	51179	HAO2	hydroxyacid oxidase 2 (long chain)	Downregulated
35809	Intrônico asPI	8826	IQGAP1	IQ motif containing GTPase activating protein 1	Downregulated

35953	Intrônico asPI	388815	C21orf34	chromosome 21 open reading frame 34	Downregulated
36017	Intrônico asPI	8454	CUL1	cullin 1	Downregulated
36080	Intrônico asPI	57644	MYH7B	myosin, heavy polypeptide 7B, cardiac muscle, beta	Downregulated
36110	Intrônico asPI	10395	DLC1	deleted in liver cancer 1	Downregulated
36160	Intrônico asPI	285973	NOS3AS	nitric oxide synthase 3 antisense	Downregulated
36294	Intrônico asPI	55034	MOCOS	molybdenum cofactor sulfurase	Downregulated
36397	Intrônico asPI	54843	SYTL2	synaptotagmin-like 2	Downregulated
36471	Intrônico asPI	257160	-	-	Downregulated
36590	Intrônico asPI	79682	MLF1IP	MLF1 interacting protein nudix (nucleoside diphosphate linked moiety X)-type motif 13	Downregulated
36612	Intrônico asPI	25961	NUDT13	NUDT13	Downregulated
36619	Intrônico asPI	22905	EPN2	epsin 2	Downregulated
36732	Intrônico asPI	9757	-	-	Downregulated
37307	Intrônico asPI	84196	USP48	ubiquitin specific protease 48	Downregulated
37318	Intrônico asPI	57096	RPGRIP1	retinitis pigmentosa GTPase regulator interacting protein 1	Downregulated
37341	Intrônico asPI	22872	SEC31L1	SEC31-like 1 (S. cerevisiae)	Downregulated
37436	Intrônico asPI	25961	NUDT13	nudix (nucleoside diphosphate linked moiety X)-type motif 13	Downregulated
37465	Intrônico asPI	6710	SPTB	spectrin, beta, erythrocytic (includes spherocytosis, clinical type I)	Downregulated
37586	Intrônico asPI	2314	FLII	flightless I homolog (Drosophila)	Downregulated
37643	Intrônico asPI	24148	C20orf14	chromosome 20 open reading frame 14	Downregulated
37783	Intrônico asPI	4927	NUP88	nucleoporin 88kDa	Downregulated
37856	Intrônico asPI	11231	SEC63	SEC63-like (S. cerevisiae)	Downregulated
37866	Intrônico asPI	9481	SLC25A27	solute carrier family 25, member 27	Downregulated
37868	Intrônico asPI	9044	BTAF1	BTAF1 RNA polymerase II, B-TFIID transcription factor-associated, 170kDa (Mot1 homolog, S. cerevisiae)	Downregulated
37913	Intrônico asPI	257218	SHPRH	SNF2 histone linker PHD RING helicase	Downregulated
37963	Intrônico asPI	10535	RNASEH2A	ribonuclease H2, large subunit	Downregulated
38032	Intrônico asPI	79677	SMC6L1	SMC6 structural maintenance of chromosomes 6-like 1 (yeast)	Downregulated
38059	Intrônico asPI	23141	-	-	Downregulated
38165	Intrônico asPI	22873	DZIP1	DAZ interacting protein 1	Downregulated
38384	Intrônico asPI	9001	HAP1	huntingtin-associated protein 1 (neuroan 1)	Downregulated
38389	Intrônico asPI	7528	YY1	YY1 transcription factor	Downregulated
38447	Intrônico asPI	60559	SPCS3	signal peptidase complex subunit 3 homolog (S. cerevisiae)	Downregulated
38552	Intrônico asPI	57649	PHF12	PHD finger protein 12	Downregulated
38563	Intrônico asPI	51704	GPRC5B	G protein-coupled receptor, family C, group 5, member B	Downregulated

38706	Intrônico asPI	5533	PPP3CC	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma)	Downregulated
38787	Intrônico asPI	10535	RNASEH2A	ribonuclease H2, large subunit	Downregulated
38790	Intrônico asPI	10038	PARP2	poly (ADP-ribose) polymerase family, member 2	Downregulated
38791	Intrônico asPI	9898	UBAP2L	ubiquitin associated protein 2-like	Downregulated
38904	Intrônico asPI	7320	UBE2B	ubiquitin-conjugating enzyme E2B (RAD6 homolog)	Downregulated
39070	Intrônico asPI	55827	IQWD1	IQ motif and WD repeats 1	Downregulated
39076	Intrônico asPI	54509	RHOF	ras homolog gene family, member F (in filopodia)	Downregulated
39281	Intrônico asPI	54839	LRRC49	leucine rich repeat containing 49	Downregulated
39304	Intrônico asPI	10051	SMC4L1	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	Downregulated
39413	Intrônico asPI	9219	MTA2	metastasis associated 1 family, member 2	Downregulated
39421	Intrônico asPI	7082	TJP1	tight junction protein 1 (zonula occludens 1)	Downregulated
39684	Intrônico asPI	57674	C17orf27	chromosome 17 open reading frame 27	Downregulated
39812	Intrônico asPI	23035	PHLPPL	PH domain and leucine rich repeat protein phosphatase like	Downregulated
39814	Intrônico asPI	11213	IRAK3	interleukin-1 receptor-associated kinase 3	Downregulated
39829	Intrônico asPI	8438	RAD54L	RAD54-like (S. cerevisiae)	Downregulated
39894	Intrônico asPI	55827	IQWD1	IQ motif and WD repeats 1	Downregulated
39948	Intrônico asPI	4134	MAP4	microtubule-associated protein 4	Downregulated
39994	Intrônico asPI	57650	KIAA1524	KIAA1524	Downregulated
40031	Intrônico asPI	9219	MTA2	metastasis associated 1 family, member 2	Downregulated
40310	Intrônico asPI	55063	ZCWPW1	zinc finger, CW type with PWWP domain 1	Downregulated
40316	Intrônico asPI	50485	SMARCAL1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1	Downregulated
40441	Intrônico asPI	9644	SH3MD1	SH3 multiple domains 1	Downregulated
40506	Intrônico asPI	64398	MPP5	membrane protein, palmitoylated 5 (MAGUK p55 subfamily member 5)	Downregulated
40835	Intrônico asPI	26152	ZNF337	zinc finger protein 337	Downregulated
40844	Intrônico asPI	11214	AKAP13	A kinase (PRKA) anchor protein 13	Downregulated
40855	Intrônico asPI	9221	NOLC1	nucleolar and coiled-body phosphoprotein 1	Downregulated
41472	Intrônico asPI	9467	SH3BP5	SH3-domain binding protein 5 (BTK-associated)	Downregulated
41577	Intrônico asPI	8819	SAP30	sin3-associated polypeptide, 30kDa	Downregulated
41590	Intrônico asPI	5531	PPP4C	protein phosphatase 4 (formerly X), catalytic subunit	Downregulated

41595	Intrônico asPI	4343	MOV10	Mov10, Moloney leukemia virus 10, homolog (mouse)	Downregulated
41740	Intrônico asPI	79668	PARP8	poly (ADP-ribose) polymerase family, member 8	Downregulated
41888	Intrônico asPI	8613	PPAP2B	phosphatidic acid phosphatase type 2B	Downregulated
41947	Intrônico asPI	65117	-	-	Downregulated
42001	Intrônico asPI	5725	PTBP1	polypyrimidine tract binding protein 1	Downregulated
42275	Intrônico asPI	28991	COMMD5	COMM domain containing 5	Downregulated
42286	Intrônico asPI	11214	AKAP13	A kinase (PRKA) anchor protein 13	Downregulated
42479	Intrônico asPI	51131	PHF11	PHD finger protein 11	Downregulated
42967	Intrônico asPI	114294	LACTB	lactamase, beta	Downregulated
43211	Intrônico asPI	22889	-	-	Downregulated
43394	Intrônico asPI	57129	MRPL47	mitochondrial ribosomal protein L47	Downregulated
43743	Intrônico asPI	8443	GNPAT	glyceronephosphate O-acyltransferase	Downregulated
43878	Intrônico asPI	2	A2M	alpha-2-macroglobulin	Downregulated
44153	Intrônico asPI	8814	CDKL1	cyclin-dependent kinase-like 1 (CDC2-related kinase)	Downregulated
44227	Intrônico asPI	51710	ZNF44	zinc finger protein 44 (KOX 7)	Downregulated
578	Exônico	9844	ELMO1	engulfment and cell motility 1 (ced-12 homolog, C. elegans)	Upregulated
1633	Exônico	3991	LIPE	lipase, hormone-sensitive	Upregulated
2258	Exônico	2036	EPB41L1	erythrocyte membrane protein band 4.1-like 1	Upregulated
2675	Exônico	759	CA1	carbonic anhydrase I	Upregulated
2925	Exônico	51593	-	-	Upregulated
3013	Exônico	79081	-	-	Upregulated
3667	Exônico	9962	SLC23A2	solute carrier family 23 (nucleobase transporters), member 2	Upregulated
4430	Exônico				Upregulated
4529	Exônico	753	C18orf1	chromosome 18 open reading frame 1	Upregulated
4533	Exônico	220382	-	-	Upregulated
4919	Exônico	6304	SATB1	special AT-rich sequence binding protein 1 (binds to nuclear matrix/scaffold-associating DNA's)	Upregulated
5030	Exônico	4643	MYO1E	myosin IE	Upregulated
5714	Exônico	23569	PADI4	peptidyl arginine deiminase, type IV	Upregulated
5952	Exônico	5602	MAPK10	mitogen-activated protein kinase 10	Upregulated
6407	Exônico	84107	ZIC4	Zic family member 4	Upregulated
6769	Exônico	7169	TPM2	tropomyosin 2 (beta)	Upregulated
7089	Exônico	4846	NOS3	nitric oxide synthase 3 (endothelial cell)	Upregulated
7269	Exônico	10489	LRRC41	leucine rich repeat containing 41	Upregulated
7285	Exônico	7038	TG	thyroglobulin	Upregulated
7507	Exônico	3133	HLA-E	major histocompatibility complex, class I, E	Upregulated
7727	Exônico	81033	KCNH6	potassium voltage-gated channel, subfamily H (eag-related), member 6	Upregulated
7743	Exônico	89927	C16orf45	chromosome 16 open reading frame 45	Upregulated

7914	Exônico	4644	MYO5A	myosin VA (heavy polypeptide 12, myoxin)	Upregulated
8020	Exônico	3709	ITPR2	inositol 1,4,5-triphosphate receptor, type 2	Upregulated
8035	Exônico				Upregulated
8177	Exônico	53616	ADAM22	a disintegrin and metalloproteinase domain 22	Upregulated
8551	Exônico				Upregulated
8663	Exônico	171024	SYNPO2	synaptopodin 2	Upregulated
8822	Exônico	9123	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Upregulated
8963	Exônico	11156	PTP4A3	protein tyrosine phosphatase type IVA, member 3	Upregulated
9006	Exônico	27301	APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2	Upregulated
9007	Exônico	26509	FER1L3	fer-1-like 3, myoferlin (C. elegans)	Upregulated
9379	Exônico	404266	-	-	Upregulated
9858	Exônico	7392	USF2	upstream transcription factor 2, c-fos interacting	Upregulated
10094	Exônico	59	ACTA2	actin, alpha 2, smooth muscle, aorta	Upregulated
10718	Exônico	441220	-	-	Upregulated
11124	Exônico	59	ACTA2	actin, alpha 2, smooth muscle, aorta	Upregulated
11208	Exônico	5082	PDCL	phosducin-like	Upregulated
11339	Exônico	254268	C1orf62	chromosome 1 open reading frame 62	Upregulated
11771	Exônico	58508	MLL3	myeloid/lymphoid or mixed-lineage leukemia 3	Upregulated
11860	Exônico	122953	-	-	Upregulated
11968	Exônico	84467	FBN3	fibrillin 3	Upregulated
12335	Exônico	6277	S100A6	S100 calcium binding protein A6 (calcyclin)	Upregulated
12368	Exônico	284217	LAMA1	laminin, alpha 1	Upregulated
13011	Exônico	55930	MYO5C	myosin VC	Upregulated
13108	Exônico	64753	-	-	Upregulated
13669	Exônico	7431	VIM	vimentin	Upregulated
13880	Exônico	6347	CCL2	chemokine (C-C motif) ligand 2	Upregulated
14027	Exônico	84940	CORO6	coronin 6	Upregulated
14215	Exônico	3117	HLA-DQA1	major histocompatibility complex, class II, DQ alpha 1	Upregulated
14271	Exônico	10942	PRSS21	protease, serine, 21 (testisin)	Upregulated
14287	Exônico	7431	VIM	vimentin	Upregulated
14627	Exônico				Upregulated
14834	Exônico				Upregulated
14956	Exônico	83938	C10orf11	chromosome 10 open reading frame 11	Upregulated
15476	Exônico	79158	-	-	Upregulated
15533	Exônico	5274	SERPINI1	serine (or cysteine) proteinase inhibitor, clade I (neuroserpin), member 1	Upregulated
15760	Exônico				Upregulated
16061	Exônico	1773	DNASE1	deoxyribonuclease I	Upregulated
16084	Exônico	116028	-	-	Upregulated
16120	Exônico	23216	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	Upregulated
16264	Exônico	2534	FYN	FYN oncogene related to SRC, FGR, YES	Upregulated

16393	Exônico	114819	-	-	Upregulated
16581	Exônico	593	BCKDHA	branched chain keto acid dehydrogenase E1, alpha polypeptide (maple syrup urine disease)	Upregulated
16601	Exônico	89845	ABCC10	ATP-binding cassette, sub-family C (CFTR/MRP), member 10	Upregulated
16736	Exônico	23406	COTL1	coactosin-like 1 (Dictyostelium)	Upregulated
16932	Exônico	54331	GNG2	guanine nucleotide binding protein (G protein), gamma 2	Upregulated
16996	Exônico				Upregulated
17019	Exônico	79949	C10orf81	chromosome 10 open reading frame 81	Upregulated
17099	Exônico	162417	NAGS	N-acetylglutamate synthase	Upregulated
17119	Exônico	84162	KIAA1109	KIAA1109	Upregulated
17214	Exônico	143903	-	-	Upregulated
17225	Exônico	79949	C10orf81	chromosome 10 open reading frame 81	Upregulated
17509	Exônico	259	AMBP	alpha-1-microglobulin/bikunin precursor	Upregulated
18732	Exônico	3691	ITGB4	integrin, beta 4	Upregulated
19011	Exônico	10516	FBLN5	fibulin 5	Upregulated
19138	Exônico	5176	SERPINF1	serine (or cysteine) proteinase inhibitor, clade F (alpha-2 antiplasmin, pigment epithelium derived factor), member 1	Upregulated
19392	Exônico	64167	-	-	Upregulated
19806	Exônico	57609	-	-	Upregulated
20105	Exônico	84299	C17orf37	chromosome 17 open reading frame 37	Upregulated
20312	Exônico	83872	HMCN1	hemicentin 1	Upregulated
20903	Exônico	1604	DAF	decay accelerating factor for complement (CD55, Cromer blood group system)	Upregulated
21007	Exônico	1191	CLU	clusterin (complement lysis inhibitor, SP-40,40, sulfated glycoprotein 2, testosterone repressed prostate message 2, apolipoprotein J)	Upregulated
21022	Exônico	220164	DOK6	docking protein 6	Upregulated
21136	Exônico	83871	RAB34	RAB34, member RAS oncogene family	Upregulated
21400	Exônico	5918	RARRES1	retinoic acid receptor responder (tazarotene induced) 1	Upregulated
21965	Exônico	78994	-	-	Upregulated
22029	Exônico	3371	TNC	tenascin C (hexabrachion)	Upregulated
22166	Exônico	83857	-	-	Upregulated
22172	Exônico	64582	GPR135	G protein-coupled receptor 135	Upregulated
22234	Exônico	3687	ITGAX	integrin, alpha X (antigen CD11C (p150), alpha polypeptide)	Upregulated
22975	Exônico	400566	-	-	Upregulated
23103	Exônico	57458	TMCC3	transmembrane and coiled-coil domains 3	Upregulated
23450	Exônico	8635	RNASET2	ribonuclease T2	Upregulated

23502	Exônico	90865	C9orf26	chromosome 9 open reading frame 26 (NF-HEV)	Upregulated
24027	Exônico	63036	-	-	Upregulated
24110	Exônico	284618	C1orf104	chromosome 1 open reading frame 104	Upregulated
24287	Exônico	5373	PMM2	phosphomannomutase 2	Upregulated
24293	Exônico	3921	RPSA	ribosomal protein SA	Upregulated
24367	Exônico	10523	CHERP	calcium homeostasis endoplasmic reticulum protein	Upregulated
25130	Exônico	473	RERE	arginine-glutamic acid dipeptide (RE) repeats	Upregulated
25353	Exônico	126248	-	-	Upregulated
25394	Exônico	11176	BAZ2A	bromodomain adjacent to zinc finger domain, 2A	Upregulated
25735	Exônico	3920	LAMP2	lysosomal-associated membrane protein 2	Upregulated
26045	Exônico	3831	KNS2	kinesin 2 60/70kDa	Upregulated
26521	Exônico	23358	USP24	ubiquitin specific protease 24	Upregulated
26532	Exônico	10075	HUWE1	HECT, UBA and WWE domain containing 1	Upregulated
26771	Exônico	2146	EZH2	enhancer of zeste homolog 2 (Drosophila)	Upregulated
27035	Exônico	23650	TRIM29	tripartite motif-containing 29	Upregulated
27144	Exônico	11253	MAN1B1	mannosidase, alpha, class 1B, member 1	Upregulated
27168	Exônico	5909	RAP1GA1	RAP1, GTPase activating protein 1	Upregulated
27169	Exônico	5768	QSCN6	quiescin Q6	Upregulated
27388	Exônico	2354	FOSB	FBJ murine osteosarcoma viral oncogene homolog B	Upregulated
27591	Exônico	3339	HSPG2	heparan sulfate proteoglycan 2 (perlecan)	Upregulated
27651	Exônico	26115	-	-	Upregulated
27948	Exônico	55558	PLXNA3	plexin A3	Upregulated
28221	Exônico	213	ALB	albumin	Upregulated
28619	Exônico	3914	LAMB3	laminin, beta 3	Upregulated
28714	Exônico	5836	PYGL	phosphorylase, glycogen; liver (Hers disease, glycogen storage disease type VI)	Upregulated
29138	Exônico	2915	GRM5	glutamate receptor, metabotropic 5	Upregulated
29460	Exônico	10586	MAB21L2	mab-21-like 2 (C. elegans)	Upregulated
30354	Exônico	7975	MAFK	v-maf musculoaponeurotic fibrosarcoma oncogene homolog K (avian)	Upregulated
30379	Exônico	1634	DCN	decorin	Upregulated
30765	Exônico	8412	BCAR3	breast cancer anti-estrogen resistance 3	Upregulated
30902	Exônico				Upregulated
31186	Exônico	5858	PZP	pregnancy-zone protein	Upregulated
31416	Exônico				Upregulated
31726	Exônico	3494	IGHA2	immunoglobulin heavy constant alpha 2 (A2m marker)	Upregulated
32282	Exônico	50618	ITSN2	intersectin 2	Upregulated
32317	Exônico	6203	RPS9	ribosomal protein S9	Upregulated
32326	Exônico	4151	MB	myoglobin	Upregulated
32855	Exônico	633	BGN	biglycan	Upregulated

32932	Exônico	6916	TBXAS1	thromboxane A synthase 1 (platelet, cytochrome P450, family 5, subfamily A)	Upregulated
33663	Exônico	4921	DDR2	discoidin domain receptor family, member 2	Upregulated
33940	Exônico	23043	TNIK	TRAF2 and NCK interacting kinase	Upregulated
34109	Exônico	114785	MBD6	methyl-CpG binding domain protein 6	Upregulated
35113	Exônico	2733	GLE1L	GLE1 RNA export mediator like (yeast)	Upregulated
35203	Exônico	5747	PTK2	PTK2 protein tyrosine kinase 2	Upregulated
35304	Exônico	6446	SGK	serum/glucocorticoid regulated kinase	Upregulated
35323	Exônico	1650	DDOST	dolichyl- diphosphooligosaccharide- protein glycosyltransferase	Upregulated
35412	Exônico	5151	PDE8A	phosphodiesterase 8A	Upregulated
37696	Exônico	416	ARSF	arylsulfatase F	Upregulated
37790	Exônico	2952	GSTT1	glutathione S-transferase theta 1	Upregulated
37796	Exônico	1384	CRAT	carnitine acetyltransferase	Upregulated
37901	Exônico	667	DST	dystonin	Upregulated
38095	Exônico	3910	LAMA4	laminin, alpha 4	Upregulated
38299	Exônico	4354	MPP1	membrane protein, palmitoylated 1, 55kDa	Upregulated
38606	Exônico	5142	PDE4B	phosphodiesterase 4B, cAMP-specific (phosphodiesterase E4 dunce homolog, Drosophila)	Upregulated
38730	Exônico	2835	GPR12	G protein-coupled receptor 12	Upregulated
39032	Exônico	1384	CRAT	carnitine acetyltransferase	Upregulated
39038	Exônico	81491	GPR63	G protein-coupled receptor 63	Upregulated
39655	Exônico	165	AEBP1	AE binding protein 1	Upregulated
40145	Exônico	6185	RPN2	ribophorin II	Upregulated
40163	Exônico	1672	DEFB1	defensin, beta 1	Upregulated
40383	Exônico	387601	-	-	Upregulated
40557	Exônico	6185	RPN2	ribophorin II	Upregulated
40943	Exônico	23139	MAST2	microtubule associated serine/threonine kinase 2	Upregulated
41495	Exônico	3860	KRT13	keratin 13	Upregulated
41698	Exônico	4707	NDUFB1	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 1, 7kDa	Upregulated
41708	Exônico	1950	EGF	epidermal growth factor (beta-urogastrone)	Upregulated
41901	Exônico	5319	PLA2G1B	phospholipase A2, group IB (pancreas)	Upregulated
41905	Exônico	4313	MMP2	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)	Upregulated
42018	Exônico	1396	CRIP1	cysteine-rich protein 1 (intestinal)	Upregulated
42411	Exônico	6144	RPL21	ribosomal protein L21	Upregulated
42519	Exônico	5320	PLA2G2A	phospholipase A2, group IIA (platelets, synovial fluid)	Upregulated
42924	Exônico	6876	TAGLN	transgelin	Upregulated

6	Exônico	440275	EIF2AK4	eukaryotic translation initiation factor 2 alpha kinase 4	Downregulated
31	Exônico	57291	-	-	Downregulated
33	Exônico	55902	ACAS2	acetyl-Coenzyme A synthetase 2 (ADP forming)	Downregulated
40	Exônico	54014	BRWD1	bromodomain and WD repeat domain containing 1	Downregulated
41	Exônico	51585	-	-	Downregulated
61	Exônico	10128	LRPPRC	leucine-rich PPR-motif containing	Downregulated
68	Exônico	8878	SQSTM1	sequestosome 1	Downregulated
72	Exônico	7840	ALMS1	Alstrom syndrome 1	Downregulated
73	Exônico	7407	VARs2	valyl-tRNA synthetase 2	Downregulated
76	Exônico	6777	STAT5B	signal transducer and activator of transcription 5B	Downregulated
80	Exônico	5792	PTPRF	protein tyrosine phosphatase, receptor type, F	Downregulated
111	Exônico	284001	-	-	Downregulated
120	Exônico	93058	-	-	Downregulated
124	Exônico	84067	-	-	Downregulated
126	Exônico	80273	GRPEL1	GrpE-like 1, mitochondrial (E. coli)	Downregulated
130	Exônico	64761	PARP12	poly (ADP-ribose) polymerase family, member 12	Downregulated
143	Exônico	54014	BRWD1	bromodomain and WD repeat domain containing 1	Downregulated
153	Exônico	23387	-	-	Downregulated
168	Exônico	9569	GTF2IRD1	GTF2I repeat domain containing 1	Downregulated
170	Exônico	9136	RNU3IP2	RNA, U3 small nucleolar interacting protein 2	Downregulated
173	Exônico	8527	DGKD	diacylglycerol kinase, delta 130kDa	Downregulated
177	Exônico	7174	TPP2	tripeptidyl peptidase II	Downregulated
191	Exônico	4000	LMNA	lamin A/C	Downregulated
202	Exônico	976	CD97	CD97 antigen	Downregulated
220	Exônico	133015	-	-	Downregulated
223	Exônico	93081	-	-	Downregulated
251	Exônico	27332	ZNF638	zinc finger protein 638	Downregulated
263	Exônico	10905	MAN1A2	mannosidase, alpha, class 1A, member 2	Downregulated
266	Exônico	10280	OPRS1	opioid receptor, sigma 1	Downregulated
273	Exônico	9136	RNU3IP2	RNA, U3 small nucleolar interacting protein 2	Downregulated
280	Exônico	7175	TPR	translocated promoter region (to activated MET oncogene)	Downregulated
294	Exônico	4001	LMNB1	lamin B1	Downregulated
297	Exônico	3093	HIP2	huntingtin interacting protein 2	Downregulated
308	Exônico	308	ANXA5	annexin A5	Downregulated
323	Exônico	133015	-	-	Downregulated
333	Exônico	79980	C20orf172	chromosome 20 open reading frame 172	Downregulated
349	Exônico	54014	BRWD1	bromodomain and WD repeat domain containing 1	Downregulated
352	Exônico	51029	-	-	Downregulated
357	Exônico	25842	ASF1A	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	Downregulated
374	Exônico	9569	GTF2IRD1	GTF2I repeat domain containing 1	Downregulated

382	Exônico	7405	UVRAG	UV radiation resistance associated gene	Downregulated
398	Exônico	3720	JARID2	Jumonji, AT rich interactive domain 2	Downregulated
402	Exônico	2597	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	Downregulated
409	Exônico	770	CA11	carbonic anhydrase XI	Downregulated
448	Exônico	55316	RSAD1	radical S-adenosyl methionine domain containing 1	Downregulated
451	Exônico	54762	-	-	Downregulated
466	Exônico	22986	SORCS3	sortilin-related VPS10 domain containing receptor 3	Downregulated
495	Exônico	5218	PFTK1	PFTAIRES protein kinase 1	Downregulated
501	Exônico	3735	KARS	lysyl-tRNA synthetase	Downregulated
505	Exônico	2617	GARS	glycyl-tRNA synthetase	Downregulated
506	Exônico	2195	FAT	FAT tumor suppressor homolog 1 (Drosophila)	Downregulated
511	Exônico	976	CD97	CD97 antigen	Downregulated
516	Exônico	5081	PAX7	paired box gene 7	Downregulated
532	Exônico	93058	-	-	Downregulated
542	Exônico	64760	RAI16	retinoic acid induced 16	Downregulated
547	Exônico	56907	SPIRE1	spire homolog 1 (Drosophila)	Downregulated
553	Exônico	54914	KIAA1797	KIAA1797	Downregulated
554	Exônico	54758	KLHDC4	kelch domain containing 4	Downregulated
560	Exônico	27332	ZNF638	zinc finger protein 638	Downregulated
573	Exônico	10640	SEC10L1	SEC10-like 1 (S. cerevisiae)	Downregulated
577	Exônico	9967	THRAP3	thyroid hormone receptor associated protein 3	Downregulated
582	Exônico	9135	RABEP1	rabaptin, RAB GTPase binding effector protein 1	Downregulated
616	Exônico	513	ATP5D	ATP synthase, H+ transporting, mitochondrial F1 complex, delta subunit	Downregulated
624	Exônico	440275	EIF2AK4	eukaryotic translation initiation factor 2 alpha kinase 4	Downregulated
626	Exônico	284004	-	-	Downregulated
628	Exônico	219402	-	-	Downregulated
642	Exônico	79980	C20orf172	chromosome 20 open reading frame 172	Downregulated
650	Exônico	56907	SPIRE1	spire homolog 1 (Drosophila)	Downregulated
653	Exônico	55609	SUHW3	suppressor of hairy wing homolog 3 (Drosophila)	Downregulated
666	Exônico	25850	ZNF345	zinc finger protein 345	Downregulated
689	Exônico	8289	ARID1A	AT rich interactive domain 1A (SWI-like)	Downregulated
690	Exônico	7849	PAX8	paired box gene 8	Downregulated
710	Exônico	2869	GRK5	G protein-coupled receptor kinase 5	Downregulated
734	Exônico	145837	-	-	Downregulated
751	Exônico	57572	DOCK6	dedicator of cytokinesis 6	Downregulated
769	Exônico	25842	ASF1A	ASF1 anti-silencing function 1 homolog A (S. cerevisiae)	Downregulated
770	Exônico	23588	KLHDC2	kelch domain containing 2	Downregulated
774	Exônico	23082	PPRC1	peroxisome proliferative activated receptor, gamma, coactivator-related 1	Downregulated

780	Exônico	10480	-	-	Downregulated
795	Exônico	7174	TPP2	tripeptidyl peptidase II	Downregulated
806	Exônico	4809	NHP2L1	NHP2 non-histone chromosome protein 2-like 1 (<i>S. cerevisiae</i>)	Downregulated
826	Exônico				Downregulated
836	Exônico	153830	-	-	Downregulated
851	Exônico	64761	PARP12	poly (ADP-ribose) polymerase family, member 12	Downregulated
866	Exônico	51300	C3orf1	chromosome 3 open reading frame 1	Downregulated
868	Exônico	29880	ALG5	asparagine-linked glycosylation 5 homolog (yeast, dolichyl-phosphate beta-glucosyltransferase)	Downregulated
873	Exônico	23592	LEMD3	LEM domain containing 3	Downregulated
899	Exônico	7026	NR2F2	nuclear receptor subfamily 2, group F, member 2	Downregulated
917	Exônico	2618	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimida zole synthetase	Downregulated
918	Exônico	2196	FAT2	FAT tumor suppressor homolog 2 (<i>Drosophila</i>)	Downregulated
929	Exônico				Downregulated
935	Exônico	283989	TSEN54	tRNA splicing endonuclease 54 homolog (SEN54, <i>S. cerevisiae</i>)	Downregulated
942	Exônico	124245	-	-	Downregulated
951	Exônico	79979	CXorf34	chromosome X open reading frame 34	Downregulated
953	Exônico	79083	MLPH	melanophilin	Downregulated
959	Exônico	56905	-	-	Downregulated
971	Exônico	29855	UBN1	ubiquitin 1	Downregulated
984	Exônico	10902	BRD8	bromodomain containing 8	Downregulated
988	Exônico	10128	LRPPRC	leucine-rich PPR-motif containing	Downregulated
990	Exônico	9841	ZBTB24	zinc finger and BTB domain containing 24	Downregulated
1004	Exônico	6591	SNAI2	snail homolog 2 (<i>Drosophila</i>)	Downregulated
1009	Exônico	5442	POLRMT	polymerase (RNA) mitochondrial (DNA directed)	Downregulated
1021	Exônico	2194	FASN	fatty acid synthase	Downregulated
1031	Exônico				Downregulated
1032	Exônico	57176	VARS2L	valyl-tRNA synthetase 2-like	Downregulated
1055	Exônico	79745	RSNL2	restin-like 2	Downregulated
1066	Exônico	55316	RSAD1	radical S-adenosyl methionine domain containing 1	Downregulated
1069	Exônico	54764	ZRANB1	zinc finger, RAN-binding domain containing 1	Downregulated
1103	Exônico	7407	VARS2	valyl-tRNA synthetase 2 SWI/SNF related, matrix associated, actin	Downregulated
1107	Exônico	6595	SMARCA2	dependent regulator of chromatin, subfamily a, member 2	Downregulated
1112	Exônico	5447	POR	P450 (cytochrome) oxidoreductase	Downregulated

1134	Exônimo				Downregulated
1140	Exônimo	340481	ZDHHC21	zinc finger, DHHC-type containing 21	Downregulated
1141	Exônimo	283989	TSEN54	tRNA splicing endonuclease 54 homolog (SEN54, <i>S. cerevisiae</i>)	Downregulated
1156	Exônimo	80267	C1orf22	chromosome 1 open reading frame 22	Downregulated
1158	Exônimo	79739	-	-	Downregulated
1174	Exônimo	51585	-	-	Downregulated
1196	Exônimo	9841	ZBTB24	zinc finger and BTB domain containing 24	Downregulated
1199	Exônimo	9354	UBE4A	ubiquitination factor E4A (UFD2 homolog, yeast)	Downregulated
1209	Exônimo	6776	STAT5A	signal transducer and activator of transcription 5A	Downregulated
1229	Exônimo	1793	DOCK1	dedicator of cytokinesis 1	Downregulated
1238	Exônimo	389250	-	-	Downregulated
1246	Exônimo	219402	-	-	Downregulated
1258	Exônimo	81691	-	-	Downregulated
1267	Exônimo	57325	CSRP2BP	CSRP2 binding protein	Downregulated
1270	Exônimo	55746	NUP133	nucleoporin 133kDa	Downregulated
1273	Exônimo	55139	-	-	Downregulated
1280	Exônimo	29882	ANAPC2	anaphase promoting complex subunit 2	Downregulated
1288	Exônimo	23196	C9orf10	chromosome 9 open reading frame 10	Downregulated
1300	Exônimo	9698	PUM1	pumilio homolog 1 (<i>Drosophila</i>)	Downregulated
1305	Exônimo	8672	EIF4G3	eukaryotic translation initiation factor 4 gamma, 3	Downregulated
1312	Exônimo	6778	STAT6	signal transducer and activator of transcription 6, interleukin-4 induced	Downregulated
1322	Exônimo	4636	MYL5	myosin, light polypeptide 5, regulatory	Downregulated
1337	Exônimo	533	ATP6V0B	ATPase, H ⁺ transporting, lysosomal 21kDa, V0 subunit c''	Downregulated
1338	Exônimo	308	ANXA5	annexin A5	Downregulated
1340	Exônimo	6013	RLN1	relaxin 1	Downregulated
1341	Exônimo	1577	CYP3A5	cytochrome P450, family 3, subfamily A, polypeptide 5	Downregulated
1352	Exônimo	145748	-	-	Downregulated
1361	Exônimo	81624	DIAPH3	diaphanous homolog 3 (<i>Drosophila</i>)	Downregulated
1364	Exônimo	79739	-	-	Downregulated
1388	Exônimo	23588	KLHDC2	kelch domain containing 2	Downregulated
1413	Exônimo	7174	TPP2	tripeptidyl peptidase II	Downregulated
1421	Exônimo	5440	POLR2K	polymerase (RNA) II (DNA directed) polypeptide K, 7.0kDa	Downregulated
1432	Exônimo	2590	GALNT2	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 2 (GalNAc-T2)	Downregulated
1434	Exônimo	2036	EPB41L1	erythrocyte membrane protein band 4.1-like 1	Downregulated
1443	Exônimo				Downregulated
1448	Exônimo	440275	EIF2AK4	eukaryotic translation initiation factor 2 alpha kinase 4	Downregulated

1474	Exônico	56913	C1GALT1	core 1 synthase, glycoprotein-N- acetylgalactosamine 3-beta galactosyltransferase, 1	Downregulated
1478	Exônico	55320	C14orf106	chromosome 14 open reading frame 106	Downregulated
1486	Exônico	29882	ANAPC2	anaphase promoting complex subunit 2	Downregulated
1494	Exônico	23196	C9orf10	chromosome 9 open reading frame 10	Downregulated
1503	Exônico	10130	PDIA6	protein disulfide isomerase- associated 6	Downregulated
1515	Exônico	7407	VAR52	valyl-tRNA synthetase 2	Downregulated
1546	Exônico				Downregulated
1553	Exônico	283989	TSEN54	tRNA splicing endonuclease 54 homolog (SEN54, <i>S. cerevisiae</i>)	Downregulated
1567	Exônico	81619	TSPAN14	tetraspanin 14	Downregulated
1586	Exônico	51582	AZIN1	antizyme inhibitor 1	Downregulated
1591	Exônico	26610	ELP4	elongation protein 4 homolog (<i>S. cerevisiae</i>)	Downregulated
1596	Exônico	23279	NUP160	nucleoporin 160kDa	Downregulated
1621	Exônico	6776	STAT5A	signal transducer and activator of transcription 5A	Downregulated
1630	Exônico	4802	NFYC	nuclear transcription factor Y, gamma	Downregulated
1631	Exônico	4628	MYH10	myosin, heavy polypeptide 10, non-muscle	Downregulated
1662	Exônico	133015	-	-	Downregulated
1671	Exônico	80306	MED28	mediator of RNA polymerase II transcription, subunit 28 homolog (yeast)	Downregulated
1682	Exônico	55746	NUP133	nucleoporin 133kDa	Downregulated
1689	Exônico	51592	TRIM33	tripartite motif-containing 33	Downregulated
1691	Exônico	51031	C17orf25	chromosome 17 open reading frame 25	Downregulated
1696	Exônico	25852	ARMC8	armadillo repeat containing 8	Downregulated
1699	Exônico	23286	-	-	Downregulated
1708	Exônico	10282	BET1	BET1 homolog (<i>S.</i> <i>cerevisiae</i>)	Downregulated
1710	Exônico	9968	MED12	mediator of RNA polymerase II transcription, subunit 12 homolog (yeast)	Downregulated
1714	Exônico	9360	PPIG	peptidyl-prolyl isomerase G (cyclophilin G)	Downregulated
1717	Exônico	8678	BECN1	beclin 1 (coiled-coil, myosin like BCL2 interacting protein)	Downregulated
1722	Exônico	7175	TPR	translocated promoter region (to activated MET oncogene)	Downregulated
1729	Exônico	5611	DNAJC3	DnaJ (Hsp40) homolog, subfamily C, member 3	Downregulated
1734	Exônico	4636	MYL5	myosin, light polypeptide 5, regulatory	Downregulated
1739	Exônico	3098	HK1	hexokinase 1	Downregulated
1742	Exônico	2197	FAU	Finkel-Biskis-Reilly murine sarcoma virus (FBR-MuSV) ubiquitously expressed (fox derived); ribosomal protein S30	Downregulated

1749	Exônico	533	ATP6V0B	ATPase, H+ transporting, lysosomal 21kDa, V0 subunit c"	Downregulated
1752	Exônico				Downregulated
1767	Exônico	114885	OSBPL11	oxysterol binding protein-like 11	Downregulated
1776	Exônico	79739	-	-	Downregulated
1779	Exônico	64062	C13orf10	chromosome 13 open reading frame 10	Downregulated
1792	Exônico	51574	-	-	Downregulated
1794	Exônico	51018	-	-	Downregulated
1795	Exônico	29843	SEN1	SUMO1/sentrin specific protease 1	Downregulated
1804	Exônico	23082	PPRC1	peroxisome proliferative activated receptor, gamma, coactivator-related 1	Downregulated
1811	Exônico	10276	NET1	neuroepithelial cell transforming gene 1	Downregulated
1820	Exônico	8671	SLC4A4	solute carrier family 4, sodium bicarbonate cotransporter, member 4	Downregulated
1833	Exônico	5440	POLR2K	polymerase (RNA) II (DNA directed) polypeptide K, 7.0kDa	Downregulated
1842	Exônico	3093	HIP2	huntingtin interacting protein 2	Downregulated
1863	Exônico	253959	GARNL1	GTPase activating Rap/RanGAP domain-like 1	Downregulated
1869	Exônico	124454	-	-	Downregulated
1879	Exônico	79746	ECHDC3	enoyl Coenzyme A hydratase domain containing 3	Downregulated
1902	Exônico	25852	ARMC8	armadillo repeat containing 8	Downregulated
1922	Exônico	8879	SGPL1	sphingosine-1-phosphate lyase 1	Downregulated
1936	Exônico	5451	POU2F1	POU domain, class 2, transcription factor 1	Downregulated
1947	Exônico	2618	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	Downregulated
1952	Exônico	1287	COL4A5	collagen, type IV, alpha 5 (Alport syndrome)	Downregulated
1953	Exônico	987	LRBA	LPS-responsive vesicle trafficking, beach and anchor containing	Downregulated
1959	Exônico				Downregulated
1972	Exônico	124220	-	-	Downregulated
1994	Exônico	55132	-	-	Downregulated
2002	Exônico	27327	TNRC6A	trinucleotide repeat containing 6A	Downregulated
2006	Exônico	23588	KLHDC2	kelch domain containing 2	Downregulated
2021	Exônico	9695	EDEM1	ER degradation enhancer, mannosidase alpha-like 1	Downregulated
2038	Exônico	5608	MAP2K6	mitogen-activated protein kinase kinase 6	Downregulated
2046	Exônico	3717	JAK2	Janus kinase 2 (a protein tyrosine kinase)	Downregulated
2058	Exônico	506	ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	Downregulated

2069	Exônico	253959	GARNL1	GTPase activating Rap/RanGAP domain-like 1	Downregulated
2077	Exônico	93109	TMEM44	transmembrane protein 44	Downregulated
2092	Exônico	56916	SMARCA4	SWI/SNF-related, matrix-associated actin-dependent regulator of chromatin, subfamily a, containing DEAD/H box 1	Downregulated
2095	Exônico	55610	-	-	Downregulated
2103	Exônico	51042	ZNF593	zinc finger protein 593	Downregulated
2115	Exônico	22796	COG2	component of oligomeric golgi complex 2	Downregulated
2129	Exônico	8678	BECN1	beclin 1 (coiled-coil, myosin like BCL2 interacting protein)	Downregulated
2152	Exônico	2879	GPX4	glutathione peroxidase 4 (phospholipid hydroperoxidase)	Downregulated
2153	Exônico	2618	GART	phosphoribosylglycinamide formyltransferase, phosphoribosylglycinamide synthetase, phosphoribosylaminoimidazole synthetase	Downregulated
2205	Exônico	51280	GOLPH2	golgi phosphoprotein 2	Downregulated
2224	Exônico	10120	ACTR1B	ARP1 actin-related protein 1 homolog B, cetractin beta (yeast)	Downregulated
2234	Exônico	8242	SMCX	Smcy homolog, X-linked (mouse)	Downregulated
2245	Exônico	5438	POLR2I	polymerase (RNA) II (DNA directed) polypeptide I, 14.5kDa	Downregulated
2246	Exônico	5214	PFKP	phosphofructokinase, platelet	Downregulated
2276	Exônico	219402	-	-	Downregulated
2302	Exônico	55320	C14orf106	chromosome 14 open reading frame 106	Downregulated
2303	Exônico	55140	ELP3	elongation protein 3 homolog (S. cerevisiae)	Downregulated
2325	Exônico	10486	CAP2	CAP, adenylate cyclase-associated protein, 2 (yeast)	Downregulated
2326	Exônico	10283	SDCCAG10	serologically defined colon cancer antigen 10	Downregulated
2327	Exônico	10131	TRAP1	TNF receptor-associated protein 1	Downregulated
2336	Exônico	8531	CSDA	cold shock domain protein A	Downregulated
2341	Exônico	7030	TFE3	transcription factor binding to IGHM enhancer 3	Downregulated
2342	Exônico	6780	STAU	staufen, RNA binding protein (Drosophila)	Downregulated
2356	Exônico	3482	IGF2R	insulin-like growth factor 2 receptor	Downregulated
2363	Exônico	1523	CUTL1	cut-like 1, CCAAT displacement protein (Drosophila)	Downregulated
2369	Exônico	86	ACTL6A	actin-like 6A	Downregulated
2378	Exônico	253827	MSRB3	methionine sulfoxide reductase B3	Downregulated
2393	Exônico	79979	CXorf34	chromosome X open reading frame 34	Downregulated
2398	Exônico	57826	RAP2C	RAP2C, member of RAS oncogene family	Downregulated

2411	Exônico	51279	C1RL	complement component 1, i	Downregulated
2412	Exônico	51010	EXOSC3	subcomponent-like exosome component 3	Downregulated
2418	Exônico	23586	DDX58	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58	Downregulated
2425	Exônico	11104	KATNA1	katanin p60 (ATPase- containing) subunit A 1	Downregulated
2427	Exônico	10629	TAF6L	TAF6-like RNA polymerase II, p300/CBP-associated factor (PCAF)-associated factor, 65kDa	Downregulated
2441	Exônico	7837	-	-	Downregulated
2450	Exônico	5607	MAP2K5	mitogen-activated protein kinase kinase 5	Downregulated
2460	Exônico	3093	HIP2	huntingtin interacting protein 2	Downregulated
2473	Exônico	23217	KIAA1086	KIAA1086	Downregulated
2511	Exônico	54764	ZRANB1	zinc finger, RAN-binding domain containing 1	Downregulated
2526	Exônico	22990	PCNX	pecanex homolog (Drosophila)	Downregulated
2538	Exônico	9364	RAB28	RAB28, member RAS oncogene family	Downregulated
2539	Exônico	9139	CBFA2T2	core-binding factor, runt domain, alpha subunit 2; translocated to, 2	Downregulated
2551	Exônico	5980	REV3L	REV3-like, catalytic subunit of DNA polymerase zeta (yeast)	Downregulated
2586	Exônico	170506	DHX36	DEAH (Asp-Glu-Ala-His) box polypeptide 36	Downregulated
2596	Exônico	84059	MASS1	monogenic, audiogenic seizure susceptibility 1 homolog (mouse)	Downregulated
2599	Exônico	79979	CXorf34	chromosome X open reading frame 34	Downregulated
2601	Exônico	79081	-	-	Downregulated
2626	Exônico	23276	KLHL18	kelch-like 18 (Drosophila)	Downregulated
2628	Exônico	23081	JMJD2C	jumonji domain containing 2C	Downregulated
2635	Exônico	10274	STAG1	stromal antigen 1	Downregulated
2639	Exônico	9694	KIAA0103	KIAA0103	Downregulated
2642	Exônico	9134	CCNE2	cyclin E2	Downregulated
2643	Exônico	8869	ST3GAL5	ST3 beta-galactoside alpha 2,3-sialyltransferase 5	Downregulated
2645	Exônico	8518	IKBKAP	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase complex- associated protein	Downregulated
2657	Exônico	5438	POLR2I	polymerase (RNA) II (DNA directed) polypeptide I, 14.5kDa	Downregulated
2658	Exônico	5214	PFKP	phosphofructokinase, platelet	Downregulated
2676	Exônico	506	ATP5B	ATP synthase, H+ transporting, mitochondrial F1 complex, beta polypeptide	Downregulated
2704	Exônico	79091	-	-	Downregulated
2707	Exônico	58155	PTBP2	polypyrimidine tract binding protein 2	Downregulated
2713	Exônico	55611	OTUB1	OTU domain, ubiquitin aldehyde binding 1	Downregulated
2715	Exônico	55143	CDCA8	cell division cycle associated 8	Downregulated

2733	Exônico	22796	COG2	component of oligomeric golgi complex 2	Downregulated
2746	Exônico	8881	CDC16	CDC16 cell division cycle 16 homolog (S. cerevisiae)	Downregulated
2752	Exônico	7175	TPR	translocated promoter region (to activated MET oncogene)	Downregulated
2777	Exônico	987	LRBA	LPS-responsive vesicle trafficking, beach and anchor containing	Downregulated
2782	Exônico	5630	PRPH	peripherin	Downregulated
2796	Exônico	124152	-	-	Downregulated
2804	Exônico	80262	-	-	Downregulated
2812	Exônico	57223	-	-	Downregulated
2817	Exônico	55297	-	-	Downregulated
2819	Exônico	54908	-	-	Downregulated
2827	Exônico	26574	AATF	apoptosis antagonizing transcription factor	Downregulated
2836	Exônico	11336	SEC6L1	SEC6-like 1 (S. cerevisiae)	Downregulated
2837	Exônico	11104	KATNA1	katanin p60 (ATPase-containing) subunit A 1	Downregulated
2867	Exônico	4628	MYH10	myosin, heavy polypeptide 10, non-muscle	Downregulated
2878	Exônico	1512	CTSH	cathepsin H	Downregulated
2905	Exônico	84079	ANKRD27	ankyrin repeat domain 27 (VPS9 domain)	Downregulated
2908	Exônico	79989	-	-	Downregulated
2919	Exônico	55611	OTUB1	OTU domain, ubiquitin aldehyde binding 1	Downregulated
2920	Exônico	55321	C20orf46	chromosome 20 open reading frame 46	Downregulated
2941	Exônico	10907	TXNL4A	thioredoxin-like 4A	Downregulated
2943	Exônico	10486	CAP2	CAP, adenylate cyclase-associated protein, 2 (yeast)	Downregulated
2954	Exônico	8533	COPS3	COP9 constitutive photomorphogenic homolog subunit 3 (Arabidopsis)	Downregulated
2957	Exônico	7410	VAV2	vav 2 oncogene	Downregulated
2963	Exônico	5981	RFC1	replication factor C (activator 1) 1, 145kDa	Downregulated
2969	Exônico	4835	NQO2	NAD(P)H dehydrogenase, quinone 2	Downregulated
2977	Exônico	2620	GAS2	growth arrest-specific 2	Downregulated
2998	Exônico	170506	DHX36	DEAH (Asp-Glu-Ala-His) box polypeptide 36	Downregulated
3000	Exônico	145567	TTC7B	tetratricopeptide repeat domain 7B	Downregulated
3028	Exônico	51574	-	-	Downregulated
3035	Exônico	25836	NIPBL	Nipped-B homolog (Drosophila)	Downregulated
3044	Exônico	10890	RAB10	RAB10, member RAS oncogene family	Downregulated
3053	Exônico	9352	TXNL1	thioredoxin-like 1	Downregulated
3058	Exônico	8242	SMCX	Smcy homolog, X-linked (mouse)	Downregulated
3069	Exônico	5437	POLR2H	polymerase (RNA) II (DNA directed) polypeptide H	Downregulated
3075	Exônico	3987	LIMS1	LIM and senescent cell antigen-like domains 1	Downregulated
3104	Exônico	133522	PPARGC1B	peroxisome proliferative activated receptor, gamma, coactivator 1, beta	Downregulated

3134	Exônico	29889	GNL2	guanine nucleotide binding protein-like 2 (nucleolar)	Downregulated
3154	Exônico	9701	KIAA0685	KIAA0685	Downregulated
3158	Exônico	8883	APPBP1	amyloid beta precursor protein binding protein 1	Downregulated
3159	Exônico	8697	CDC23	CDC23 (cell division cycle 23, yeast, homolog)	Downregulated
3195	Exônico	114824	PNMA5	paraneoplastic antigen like 5	Downregulated
3208	Exônico	123811	-	-	Downregulated
3251	Exônico	10628	TXNIP	thioredoxin interacting protein	Downregulated
3261	Exônico	8867	SYNJ1	synaptojanin 1	Downregulated
3265	Exônico	7832	BTG2	BTG family, member 2	Downregulated
3286	Exônico	2580	GAK	cyclin G associated kinase	Downregulated
3352	Exônico	11113	CIT	citron (rho-interacting, serine/threonine kinase 21)	Downregulated
3355	Exônico	10487	CAP1	CAP, adenylate cyclase-associated protein 1 (yeast)	Downregulated
3366	Exônico	8534	CHST1	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1	Downregulated
3381	Exônico	4836	NMT1	N-myristoyltransferase 1	Downregulated
3387	Exônico	3104	HKR3	GLI-Kruppel family member HKR3	Downregulated
3388	Exônico	2885	GRB2	growth factor receptor-bound protein 2	Downregulated
3395	Exônico	988	CDC5L	CDC5 cell division cycle 5-like (S. pombe)	Downregulated
3397	Exônico	537	ATP6AP1	ATPase, H ⁺ transporting, lysosomal accessory protein 1	Downregulated
3411	Exônico	153396	-	-	Downregulated
3412	Exônico	145567	TTC7B	tetratricopeptide repeat domain 7B	Downregulated
3425	Exônico	79075	-	-	Downregulated
3430	Exônico	57222	-	-	Downregulated
3446	Exônico	26046	ZNF294	zinc finger protein 294	Downregulated
3449	Exônico	23386	NUDCD3	NudC domain containing 3	Downregulated
3455	Exônico	11104	KATNA1	katanin p60 (ATPase-containing) subunit A 1	Downregulated
3464	Exônico	9568	GPR51	G protein-coupled receptor 51	Downregulated
3478	Exônico	5976	RENT1	regulator of nonsense transcripts 1	Downregulated
3481	Exônico	5437	POLR2H	polymerase (RNA) II (DNA directed) polypeptide H	Downregulated
3511	Exônico	253982	-	-	Downregulated
3523	Exônico	84081	-	-	Downregulated
3526	Exônico	79991	OBFC1	oligonucleotide/oligosaccharide-binding fold containing 1	Downregulated
3543	Exônico	51593	-	-	Downregulated
3546	Exônico	29894	CPSF1	cleavage and polyadenylation specific factor 1, 160kDa	Downregulated
3557	Exônico	22796	COG2	component of oligomeric golgi complex 2	Downregulated
3559	Exônico	10908	-	-	Downregulated
3589	Exônico	4216	MAP3K4	mitogen-activated protein kinase kinase kinase 4	Downregulated
3607	Exônico				Downregulated
3611	Exônico	440145	-	-	Downregulated

3616	Exônico	170506	DHX36	DEAH (Asp-Glu-Ala-His) box polypeptide 36	Downregulated
3624	Exônico	84897	TBRG1	transforming growth factor beta regulator 1	Downregulated
3626	Exônico	84056	KATNAL1	katanin p60 subunit A-like 1	Downregulated
3630	Exônico	79734	KCTD17	potassium channel tetramerisation domain containing 17	Downregulated
3634	Exônico	57820	CCNB1IP1	cyclin B1 interacting protein 1	Downregulated
3639	Exônico	55740	ENAH	enabled homolog (Drosophila)	Downregulated
3643	Exônico	54904	WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1	Downregulated
3663	Exônico	10628	TXNIP	thioredoxin interacting protein	Downregulated
3665	Exônico	10274	STAG1	stromal antigen 1	Downregulated
3679	Exônico	7171	TPM4	tropomyosin 4	Downregulated
3681	Exônico	6774	STAT3	signal transducer and activator of transcription 3 (acute-phase response factor)	Downregulated
3731	Exônico	80314	EPC1	enhancer of polycomb homolog 1 (Drosophila)	Downregulated
3733	Exônico	79752	ZFAND1	zinc finger, AN1-type domain 1	Downregulated
3746	Exônico	54927	CHCHD3	coiled-coil-helix-coiled-coil-helix domain containing 3	Downregulated
3748	Exônico	54058	C21orf58	chromosome 21 open reading frame 58	Downregulated
3756	Exônico	25855	BRMS1	breast cancer metastasis suppressor 1	Downregulated
3775	Exônico	9145	SYNGR1	synaptogyrin 1	Downregulated
3783	Exônico	7035	TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	Downregulated
3789	Exônico	5613	PRKX	protein kinase, X-linked	Downregulated
3793	Exônico	4836	NMT1	N-myristoyltransferase 1	Downregulated
3812	Exônico				Downregulated
3832	Exônico	84056	KATNAL1	katanin p60 subunit A-like 1	Downregulated
3842	Exônico	57222	-	-	Downregulated
3844	Exônico	55898	-	-	Downregulated
3867	Exônico	11103	HRB2	HIV-1 rev binding protein 2	Downregulated
3886	Exônico	7016	TESK1	testis-specific kinase 1	Downregulated
3889	Exônico	6307	SC4MOL	sterol-C4-methyl oxidase-like	Downregulated
3892	Exônico	5604	MAP2K1	mitogen-activated protein kinase kinase 1	Downregulated
3916	Exônico				Downregulated
3927	Exônico	146198	ZFP90	zinc finger protein 90 homolog (mouse)	Downregulated
3938	Exônico	79993	ELOVL7	ELOVL family member 7, elongation of long chain fatty acids (yeast)	Downregulated
3940	Exônico	79091	-	-	Downregulated
3946	Exônico	56927	GPR108	G protein-coupled receptor 108	Downregulated
3947	Exônico	55904	MLL5	myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila)	Downregulated
3950	Exônico	55322	-	-	Downregulated
3959	Exônico	27333	GOLPH4	golgi phosphoprotein 4	Downregulated
3963	Exônico	23598	ZNF278	zinc finger protein 278	Downregulated

3982	Exônico	8883	APPBP1	amyloid beta precursor protein binding protein 1	Downregulated
3993	Exônico	5982	RFC2	replication factor C (activator 1) 2, 40kDa	Downregulated
3996	Exônico	5464	PP	pyrophosphatase (inorganic)	Downregulated
4001	Exônico	4217	MAP3K5	mitogen-activated protein kinase kinase kinase 5	Downregulated
4005	Exônico	3109	HLA-DMB	major histocompatibility complex, class II, DM beta	Downregulated
4019	Exônico	29978	UBQLN2	ubiquilin 2	Downregulated
4029	Exônico	153241	-	-	Downregulated
4030	Exônico	145508	C14orf145	chromosome 14 open reading frame 145	Downregulated
4043	Exônico	79072	-	-	Downregulated
4052	Exônico	55604	LRRC16	leucine rich repeat containing 16	Downregulated
4060	Exônico	51006	SLC35C2	solute carrier family 35, member C2	Downregulated
4075	Exônico	10627	-	-	Downregulated
4078	Exônico	10111	RAD50	RAD50 homolog (S. cerevisiae)	Downregulated
4081	Exônico	9693	RAPGEF2	Rap guanine nucleotide exchange factor (GEF) 2	Downregulated
4082	Exônico	9567	GTPBP1	GTP binding protein 1	Downregulated
4083	Exônico	9352	TXNL1	thioredoxin-like 1	Downregulated
4086	Exônico	8668	EIF3S2	eukaryotic translation initiation factor 3, subunit 2 beta, 36kDa	Downregulated
4089	Exônico	7804	LRP8	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	Downregulated
4098	Exônico	5603	MAPK13	mitogen-activated protein kinase 13	Downregulated
4102	Exônico	4799	NFX1	nuclear transcription factor, X-box binding 1	Downregulated
4122	Exônico				Downregulated
4132	Exônico	154807	VKORC1L1	vitamin K epoxide reductase complex, subunit 1-like 1	Downregulated
4138	Exônico	90268	-	-	Downregulated
4149	Exônico	58476	TP53INP2	tumor protein p53 inducible nuclear protein 2	Downregulated
4154	Exônico	55749	CCAR1	cell division cycle and apoptosis regulator 1	Downregulated
4157	Exônico	55143	CDCA8	cell division cycle associated 8	Downregulated
4158	Exônico	54934	-	-	Downregulated
4161	Exônico	51593	-	-	Downregulated
4164	Exônico	29894	CPSF1	cleavage and polyadenylation specific factor 1, 160kDa	Downregulated
4166	Exônico	26959	HBP1	HMG-box transcription factor 1	Downregulated
4174	Exônico	22992	FBXL11	F-box and leucine-rich repeat protein 11	Downregulated
4179	Exônico	10487	CAP1	CAP, adenylate cyclase-associated protein 1 (yeast)	Downregulated
4180	Exônico	10289	-	-	Downregulated
4184	Exônico	9701	KIAA0685	KIAA0685	Downregulated
4185	Exônico	9577	BRE	brain and reproductive organ-expressed (TNFRSF1A modulator)	Downregulated

4189	Exônico	8702	B4GALT4	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 4	Downregulated
4190	Exônico	8536	CAMK1	calcium/calmodulin-dependent protein kinase I	Downregulated
4195	Exônico	7035	TFPI	tissue factor pathway inhibitor (lipoprotein-associated coagulation inhibitor)	Downregulated
4196	Exônico	6787	NEK4	NIMA (never in mitosis gene a)-related kinase 4	Downregulated
4221	Exônico	539	ATP5O	ATP synthase, H+ transporting, mitochondrial F1 complex, O subunit (oligomycin sensitivity conferring protein)	Downregulated
4224	Exônico	57573	ZNF471	zinc finger protein 471	Downregulated
4230	Exônico	340277	-	-	Downregulated
4242	Exônico	84894	LRRN6A	leucine rich repeat neuronal 6A	Downregulated
4256	Exônico	55898	-	-	Downregulated
4261	Exônico	54904	WHSC1L1	Wolf-Hirschhorn syndrome candidate 1-like 1	Downregulated
4264	Exônico	51573	-	-	Downregulated
4269	Exônico	26528	DAZAP1	DAZ associated protein 1	Downregulated
4277	Exônico	22980	-	-	Downregulated
4279	Exônico	11103	HRB2	HIV-1 rev binding protein 2	Downregulated
4295	Exônico	7804	LRP8	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	Downregulated
4298	Exônico	7016	TESK1	testis-specific kinase 1	Downregulated
4301	Exônico	6305	SBF1	SET binding factor 1	Downregulated
4306	Exônico	5210	PFKFB4	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4	Downregulated
4311	Exônico	3985	LIMK2	LIM domain kinase 2	Downregulated
4347	Exônico	84083	ZRANB3	zinc finger, RAN-binding domain containing 3	Downregulated
4350	Exônico	79993	ELOVL7	ELOVL family member 7, elongation of long chain fatty acids (yeast)	Downregulated
4356	Exônico	57584	ARHGAP21	Rho GTPase activating protein 21	Downregulated
4366	Exônico	54059	C21orf57	chromosome 21 open reading frame 57	Downregulated
4371	Exônico	27335	EIF3S12	eukaryotic translation initiation factor 3, subunit 12	Downregulated
4383	Exônico	10915	TCERG1	transcription elongation regulator 1	Downregulated
4394	Exônico	8884	SLC5A6	solute carrier family 5 (sodium-dependent vitamin transporter), member 6	Downregulated
4446	Exônico	92906	HNRPLL	heterogeneous nuclear ribonucleoprotein L-like	Downregulated
4450	Exônico	84033	OBSCN	obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF	Downregulated
4465	Exônico	55296	TBC1D19	TBC1 domain family, member 19	Downregulated
4471	Exônico	51276	ZNF571	zinc finger protein 571	Downregulated
4482	Exônico	23080	-	-	Downregulated

4492	Exônico	9833	MELK	maternal embryonic leucine zipper kinase	Downregulated
4497	Exônico	8864	PER2	period homolog 2 (Drosophila)	Downregulated
4502	Exônico	7402	UTRN	utrophin (homologous to dystrophin)	Downregulated
4503	Exônico	7170	TPM3	tropomyosin 3	Downregulated
4546	Exônico	133619	-	-	Downregulated
4556	Exônico	79993	ELOVL7	ELOVL family member 7, elongation of long chain fatty acids (yeast)	Downregulated
4559	Exônico	64768	C9orf12	chromosome 9 open reading frame 12	Downregulated
4566	Exônico	55749	CCAR1	cell division cycle and apoptosis regulator 1	Downregulated
4570	Exônico	54934	-	-	Downregulated
4574	Exônico	51306	C5orf5	chromosome 5 open reading frame 5	Downregulated
4583	Exônico	23291	FBXW11	F-box and WD-40 domain protein 11	Downregulated
4586	Exônico	22994	AZI1	5-azacytidine induced 1	Downregulated
4601	Exônico	8703	B4GALT3	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 3	Downregulated
4609	Exônico	6595	SMARCA2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2	Downregulated
4611	Exônico	5982	RFC2	replication factor C (activator 1) 2, 40kDa	Downregulated
4623	Exônico	3109	HLA-DMB	major histocompatibility complex, class II, DM beta	Downregulated
4625	Exônico	2621	GAS6	growth arrest-specific 6	Downregulated
4631	Exônico	989	7-Sep	septin 7	Downregulated
4636	Exônico				Downregulated
4655	Exônico	84365	MKI67IP	MKI67 (FHA domain) interacting nucleolar phosphoprotein	Downregulated
4660	Exônico	79731	-	-	Downregulated
4672	Exônico	55127	-	-	Downregulated
4673	Exônico	54902	TTC19	tetratricopeptide repeat domain 19	Downregulated
4678	Exônico	51006	SLC35C2	solute carrier family 35, member C2	Downregulated
4684	Exônico	23576	DDAH1	dimethylarginine dimethylaminohydrolase 1	Downregulated
4688	Exônico	23080	-	-	Downregulated
4700	Exônico	9562	MINPP1	multiple inositol polyphosphate histidine phosphatase, 1	Downregulated
4709	Exônico	7170	TPM3	tropomyosin 3	Downregulated
4725	Exônico	3475	IFRD1	interferon-related developmental regulator 1	Downregulated
4732	Exônico	1512	CTSH	cathepsin H	Downregulated
4736	Exônico	493	ATP2B4	ATPase, Ca++ transporting, plasma membrane 4	Downregulated
4764	Exônico	79092	CARD14	caspase recruitment domain family, member 14	Downregulated
4769	Exônico	57343	ZNF304	zinc finger protein 304	Downregulated
4770	Exônico	56931	DUS3L	dihydrouridine synthase 3-like (S. cerevisiae)	Downregulated
4777	Exônico	54765	TRIM44	tripartite motif-containing 44	Downregulated

4778	Exônico	54059	C21orf57	chromosome 21 open reading frame 57	Downregulated
4779	Exônico	51593	-	-	Downregulated
4792	Exônico	22994	AZI1	5-azacytidine induced 1	Downregulated
4802	Exônico	9702	-	-	Downregulated
4805	Exônico	9145	SYNGR1	synaptogyrin 1	Downregulated
4813	Exônico	7037	TFRC	transferrin receptor (p90, CD71)	Downregulated
4817	Exônico	5982	RFC2	replication factor C (activator 1) 2, 40kDa	Downregulated
4842	Exônico				Downregulated
4847	Exônico	440026	TMEM41B	transmembrane protein 41B	Downregulated
4858	Exônico	92822	ZFP276	zinc finger protein 276 homolog (mouse)	Downregulated
4860	Exônico	84893	FBXO18	F-box protein, helicase, 18	Downregulated
4867	Exônico	79072	-	-	Downregulated
4868	Exônico	64754	SMYD3	SET and MYND domain containing 3	Downregulated
4885	Exônico	29803	REPIN1	replication initiator 1	Downregulated
4886	Exônico	27309	ZNF330	zinc finger protein 330	Downregulated
4896	Exônico	11335	CBX3	chromobox homolog 3 (HP1 gamma homolog, Drosophila)	Downregulated
4903	Exônico	9962	SLC23A2	solute carrier family 23 (nucleobase transporters), member 2	Downregulated
4905	Exônico	9690	UBE3C	ubiquitin protein ligase E3C	Downregulated
4913	Exônico	7803	PTP4A1	protein tyrosine phosphatase type IVA, member 1	Downregulated
4946	Exônico	2974	GUCY1B2	guanylate cyclase 1, soluble, beta 2	Downregulated
4970	Exônico	79092	CARD14	caspase recruitment domain family, member 14	Downregulated
4976	Exônico	56942	-	-	Downregulated
4989	Exônico	27335	EIF3S12	eukaryotic translation initiation factor 3, subunit 12	Downregulated
4991	Exônico	26053	AUTS2	autism susceptibility candidate 2	Downregulated
4993	Exônico	23603	CORO1C	coronin, actin binding protein, 1C	Downregulated
5001	Exônico	10915	TCERG1	transcription elongation regulator 1	Downregulated
5013	Exônico	8703	B4GALT3	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 3	Downregulated
5027	Exônico	5250	SLC25A3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3	Downregulated
5041	Exônico	1537	CYC1	cytochrome c-1	Downregulated
5043	Exônico	991	CDC20	CDC20 cell division cycle 20 homolog (S. cerevisiae)	Downregulated
5067	Exônico	84365	MKI67IP	MKI67 (FHA domain) interacting nucleolar phosphoprotein	Downregulated
5071	Exônico	79971	-	-	Downregulated
5081	Exônico	55739	-	-	Downregulated
5088	Exônico	51573	-	-	Downregulated
5099	Exônico	23193	GANAB	glucosidase, alpha; neutral AB	Downregulated
5117	Exônico	8505	PARG	poly (ADP-ribose) glycohydrolase	Downregulated

5132	Exônico	4799	NFX1	nuclear transcription factor, X-box binding 1	Downregulated
5143	Exônico	1788	DNMT3A	DNA (cytosine-5-)- methyltransferase 3 alpha	Downregulated
5166	Exônico	115416	C7orf30	chromosome 7 open reading frame 30	Downregulated
5177	Exônico	64768	C9orf12	chromosome 9 open reading frame 12	Downregulated
5180	Exônico	57589	KIAA1432	KIAA1432	Downregulated
5181	Exônico	57380	MRS2L	MRS2-like, magnesium homeostasis factor (S. cerevisiae)	Downregulated
5187	Exônico	55144	LRRC5	leucine rich repeat containing 5	Downregulated
5200	Exônico	23390	ZDHHC17	zinc finger, DHHC-type containing 17	Downregulated
5202	Exônico	23199	-	-	Downregulated
5204	Exônico	22996	C1orf34	chromosome 1 open reading frame 34	Downregulated
5205	Exônico	22800	RRAS2	related RAS viral (r-ras) oncogene homolog 2	Downregulated
5217	Exônico	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
5218	Exônico	8884	SLC5A6	solute carrier family 5 (sodium-dependent vitamin transporter), member 6	Downregulated
5221	Exônico	8295	TRRAP	transformation/transcription domain-associated protein	Downregulated
5229	Exônico	5983	RFC3	replication factor C (activator 1) 3, 38kDa	Downregulated
5255	Exônico				Downregulated
5282	Exônico	57819	LSM2	LSM2 homolog, U6 small nuclear RNA associated (S. cerevisiae)	Downregulated
5289	Exônico	55295	-	-	Downregulated
5300	Exônico	26043	-	-	Downregulated
5303	Exônico	23383	KIAA0892	KIAA0892	Downregulated
5309	Exônico	11101	ATE1	arginyltransferase 1	Downregulated
5320	Exônico	9129	PRPF3	PRP3 pre-mRNA processing factor 3 homolog (yeast)	Downregulated
5324	Exônico	8237	USP11	ubiquitin specific protease 11	Downregulated
5335	Exônico	5435	POLR2F	polymerase (RNA) II (DNA directed) polypeptide F	Downregulated
5347	Exônico	2186	FALZ	fetal Alzheimer antigen	Downregulated
5355	Exônico	291	SLC25A4	solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	Downregulated
5382	Exônico	79101	-	-	Downregulated
5388	Exônico	56945	MRPS22	mitochondrial ribosomal protein S22	Downregulated
5400	Exônico	29896	-	-	Downregulated
5417	Exônico	10135	PBEF1	pre-B-cell colony enhancing factor 1	Downregulated
5420	Exônico	9703	-	-	Downregulated
5438	Exônico	5465	PPARA	peroxisome proliferative activated receptor, alpha	Downregulated
5439	Exônico	5250	SLC25A3	solute carrier family 25 (mitochondrial carrier; phosphate carrier), member 3	Downregulated
5453	Exônico	1537	CYC1	cytochrome c-1	Downregulated
5455	Exônico	995	CDC25C	cell division cycle 25C	Downregulated

5460	Exônimo	79118	-	-	Downregulated
5479	Exônimo	84364	ZNF289	zinc finger protein 289, ID1 regulated	Downregulated
5482	Exônimo	80254	-	-	Downregulated
5485	Exônimo	79071	ELOVL6	ELOVL family member 6, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	Downregulated
5519	Exônimo	10273	STUB1	STIP1 homology and U-box containing protein 1	Downregulated
5525	Exônimo	9344	TAOK2	TAO kinase 2	Downregulated
5558	Exônimo	965	CD58	CD58 antigen, (lymphocyte function-associated antigen 3)	Downregulated
5564	Exônimo				Downregulated
5573	Exônimo	171017	ZNF384	zinc finger protein 384	Downregulated
5575	Exônimo	146705	-	-	Downregulated
5580	Exônimo	90268	-	-	Downregulated
5589	Exônimo	64768	C9orf12	chromosome 9 open reading frame 12	Downregulated
5592	Exônimo	57589	KIAA1432	KIAA1432	Downregulated
5593	Exônimo	57380	MRS2L	MRS2-like, magnesium homeostasis factor (S. cerevisiae)	Downregulated
5600	Exônimo	54940	OCIAD1	OCIA domain containing 1	Downregulated
5607	Exônimo	27338	UBE2S	ubiquitin-conjugating enzyme E2S	Downregulated
5612	Exônimo	23392	KIAA0368	KIAA0368	Downregulated
5622	Exônimo	10293	-	-	Downregulated
5623	Exônimo	10135	PBEF1	pre-B-cell colony enhancing factor 1	Downregulated
5625	Exônimo	9848	MFAP3L	microfibrillar-associated protein 3-like	Downregulated
5633	Exônimo	8295	TRRAP	transformation/transcription domain-associated protein	Downregulated
5641	Exônimo	5983	RFC3	replication factor C (activator 1) 3, 38kDa	Downregulated
5664	Exônimo	310	ANXA7	annexin A7	Downregulated
5671	Exônimo	440026	TMEM41B	transmembrane protein 41B	Downregulated
5684	Exônimo	84893	FBXO18	F-box protein, helicase, 18	Downregulated
5697	Exônimo	56894	AGPAT3	1-acylglycerol-3-phosphate O-acyltransferase 3	Downregulated
5702	Exônimo	55127	-	-	Downregulated
5720	Exônimo	11332	-	-	Downregulated
5735	Exônimo	8504	PEX3	peroxisomal biogenesis factor 3	Downregulated
5737	Exônimo	7803	PTP4A1	protein tyrosine phosphatase type IVA, member 1	Downregulated
5759	Exônimo	2186	FALZ	fetal Alzheimer antigen	Downregulated
5811	Exônimo	51053	GMNN	geminin, DNA replication inhibitor	Downregulated
5814	Exônimo	26959	HBP1	HMG-box transcription factor 1	Downregulated
5822	Exônimo	22996	C1orf34	chromosome 1 open reading frame 34	Downregulated
5867	Exônimo	995	CDC25C	cell division cycle 25C	Downregulated
5888	Exônimo	92806	-	-	Downregulated
5889	Exônimo	89953	KNSL8	kinesin-like 8	Downregulated
5907	Exônimo	55294	FBXW7	F-box and WD-40 domain protein 7 (archipelago homolog, Drosophila)	Downregulated

5913	Exônico	51271	UBAP1	ubiquitin associated protein 1	Downregulated
5918	Exônico	26035	-	-	Downregulated
5932	Exônico	10109	ARPC2	actin related protein 2/3 complex, subunit 2, 34kDa	Downregulated
5934	Exônico	9830	TRIM14	tripartite motif-containing 14	Downregulated
5944	Exônico	7402	UTRN	utrophin (homologous to dystrophin)	Downregulated
5950	Exônico	5965	RECQL	RecQ protein-like (DNA helicase Q1-like)	Downregulated
5961	Exônico	3475	IFRD1	interferon-related developmental regulator 1	Downregulated
5972	Exônico	490	ATP2B1	ATPase, Ca++ transporting, plasma membrane 1	Downregulated
6003	Exônico	58486	-	-	Downregulated
6009	Exônico	55617	C20orf13	chromosome 20 open reading frame 13	Downregulated
6010	Exônico	55325	-	-	Downregulated
6011	Exônico	55145	THAP1	THAP domain containing, apoptosis associated protein 1	Downregulated
6018	Exônico	29899	GPSM2	G-protein signalling modulator 2 (AGS3-like, C. elegans)	Downregulated
6025	Exônico	23293	C17orf31	chromosome 17 open reading frame 31	Downregulated
6026	Exônico	23200	ATP11B	ATPase, Class VI, type 11B	Downregulated
6028	Exônico	22996	C1orf34	chromosome 1 open reading frame 34	Downregulated
6032	Exônico	10652	-	-	Downregulated
6033	Exônico	10488	CREB3	cAMP responsive element binding protein 3	Downregulated
6034	Exônico	10293	-	-	Downregulated
6038	Exônico	9703	-	-	Downregulated
6045	Exônico	8301	PICALM	phosphatidylinositol binding clathrin assembly protein	Downregulated
6047	Exônico	7414	VCL	vinculin	Downregulated
6120	Exônico	51004	COQ6	coenzyme Q6 homolog (yeast)	Downregulated
6144	Exônico	9126	CSPG6	chondroitin sulfate proteoglycan 6 (bamacan)	Downregulated
6165	Exônico	3983	ABLIM1	actin binding LIM protein 1	Downregulated
6198	Exônico	90324	-	-	Downregulated
6206	Exônico	79109	MAPKAP1	mitogen-activated protein kinase associated protein 1	Downregulated
6207	Exônico	64770	CCDC14	coiled-coil domain containing 14	Downregulated
6212	Exônico	56945	MRPS22	mitochondrial ribosomal protein S22	Downregulated
6213	Exônico	55909	BIN3	bridging integrator 3	Downregulated
6221	Exônico	51594	-	-	Downregulated
6225	Exônico	27339	PRPF19	PRP19/PSO4 pre-mRNA processing factor 19 homolog (S. cerevisiae)	Downregulated
6228	Exônico	25865	PRKD2	protein kinase D2	Downregulated
6230	Exônico	23392	KIAA0368	KIAA0368	Downregulated
6232	Exônico	23200	ATP11B	ATPase, Class VI, type 11B	Downregulated
6237	Exônico	10916	MAGED2	melanoma antigen family D 2	Downregulated

6255	Exônico	7037	TFRC	transferrin receptor (p90, CD71)	Downregulated
6259	Exônico	5984	RFC4	replication factor C (activator 1) 4, 37kDa	Downregulated
6265	Exônico	4839	NOL1	nucleolar protein 1, 120kDa	Downregulated
6271	Exônico	3131	HLF	hepatic leukemia factor	Downregulated
6314	Exônico	57217	TTC7A	tetratricopeptide repeat domain 7A	Downregulated
6320	Exônico	55125	-	-	Downregulated
6329	Exônico	26512	DDX26	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26	Downregulated
6331	Exônico	25833	POU2F3	POU domain, class 2, transcription factor 3	Downregulated
6335	Exônico	23193	GANAB	glucosidase, alpha; neutral AB	Downregulated
6338	Exônico	11331	PHB2	prohibitin 2	Downregulated
6362	Exônico	5965	RECQL	RecQ protein-like (DNA helicase Q1-like)	Downregulated
6368	Exônico	4798	NFRKB	nuclear factor related to kappaB binding protein	Downregulated
6372	Exônico	3712	IVD	isovaleryl Coenzyme A dehydrogenase	Downregulated
6373	Exônico	3460	IFNGR2	interferon gamma receptor 2 (interferon gamma transducer 1)	Downregulated
6385	Exônico	288	ANK3	ankyrin 3, node of Ranvier (ankyrin G)	Downregulated
6403	Exônico	93487	C14orf32	chromosome 14 open reading frame 32	Downregulated
6414	Exônico	64083	GOLPH3	golgi phosphoprotein 3 (coat-protein)	Downregulated
6420	Exônico	55752	11-Sep	septin 11	Downregulated
6422	Exônico	55326	AGPAT5	1-acylglycerol-3-phosphate O-acyltransferase 5 (lysophosphatidic acid acyltransferase, epsilon)	Downregulated
6429	Exônico	51056	LAP3	leucine aminopeptidase 3	Downregulated
6433	Exônico	26054	SENPA6	SUMO1/sentrin specific protease 6	Downregulated
6453	Exônico	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
6459	Exônico	7415	VCP	valosin-containing protein SWI/SNF related, matrix associated, actin	Downregulated
6463	Exônico	6596	SMARCA3	dependent regulator of chromatin, subfamily a, member 3	Downregulated
6515	Exônico	79068	-	-	Downregulated
6524	Exônico	55593	-	-	Downregulated
6532	Exônico	51000	SLC35B3	solute carrier family 35, member B3	Downregulated
6540	Exônico	23274	-	-	Downregulated
6554	Exônico	9557	CHD1L	chromodomain helicase DNA binding protein 1-like	Downregulated
6556	Exônico	9126	CSPG6	chondroitin sulfate proteoglycan 6 (bamacan)	Downregulated
6559	Exônico	8504	PEX3	peroxisomal biogenesis factor 3	Downregulated
6591	Exônico	288	ANK3	ankyrin 3, node of Ranvier (ankyrin G)	Downregulated
6594	Exônico				Downregulated
6617	Exônico	79770	C5orf14	chromosome 5 open reading frame 14	Downregulated

6624	Exônico	56947	C2orf33	chromosome 2 open reading frame 33	Downregulated
6629	Exônico	55153	SDAD1	SDA1 domain containing 1	Downregulated
6634	Exônico	51312	SLC25A37	solute carrier family 25, member 37	Downregulated
6640	Exônico	25865	PRKD2	protein kinase D2	Downregulated
6650	Exônico	10652	-	-	Downregulated
6654	Exônico	9972	NUP153	nucleoporin 153kDa	Downregulated
6663	Exônico	8301	PICALM	phosphatidylinositol binding clathrin assembly protein	Downregulated
6664	Exônico	7871	SLMAP	sarcolemma associated protein	Downregulated
6669	Exônico	6596	SMARCA3	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 3	Downregulated
6693	Exônico	547	KIF1A	kinesin family member 1A	Downregulated
6718	Exônico	80232	WDR26	WD repeat domain 26	Downregulated
6724	Exônico	57801	HES4	hairy and enhancer of split 4 (Drosophila)	Downregulated
6728	Exônico	55884	WSB2	WD repeat and SOCS box-containing 2	Downregulated
6729	Exônico	55737	VPS35	vacuolar protein sorting 35 (yeast)	Downregulated
6755	Exônico	10263	CDK2AP2	CDK2-associated protein 2	Downregulated
6759	Exônico	9690	UBE3C	ubiquitin protein ligase E3C	Downregulated
6776	Exônico	5601	MAPK9	mitogen-activated protein kinase 9	Downregulated
6785	Exônico	3460	IFNGR2	interferon gamma receptor 2 (interferon gamma transducer 1)	Downregulated
6796	Exônico	488	ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	Downregulated
6800	Exônico				Downregulated
6819	Exônico	84128	WDR75	WD repeat domain 75	Downregulated
6826	Exônico	64087	MCCC2	methylcrotonoyl-Coenzyme A carboxylase 2 (beta)	Downregulated
6829	Exônico	57403	RAB22A	RAB22A, member RAS oncogene family	Downregulated
6832	Exônico	55752	11-Sep	septin 11	Downregulated
6836	Exônico	54948	MRPL16	mitochondrial ribosomal protein L16	Downregulated
6845	Exônico	26057	ANKRD17	ankyrin repeat domain 17	Downregulated
6848	Exônico	23392	KIAA0368	KIAA0368	Downregulated
6870	Exônico	7873	ARMET	arginine-rich, mutated in early stage tumors	Downregulated
6877	Exônico	5984	RFC4	replication factor C (activator 1) 4, 37kDa	Downregulated
6882	Exônico	5073	PARN	poly(A)-specific ribonuclease (deadenylation nuclease)	Downregulated
6883	Exônico	4841	NONO	non-POU domain containing, octamer-binding	Downregulated
6942	Exônico	51567	TTRAP	TRAF and TNF receptor associated protein	Downregulated
6949	Exônico	25831	HECTD1	HECT domain containing 1	Downregulated
6967	Exônico	9343	-	-	Downregulated
6971	Exônico	8502	PKP4	plakophilin 4	Downregulated
6980	Exônico	5962	RDX	radixin	Downregulated

6988	Exônico	4204	MECP2	methyl CpG binding protein 2 (Rett syndrome)	Downregulated
6990	Exônico	3712	IVD	isovaleryl Coenzyme A dehydrogenase	Downregulated
7000	Exônico	965	CD58	CD58 antigen, (lymphocyte function-associated antigen 3)	Downregulated
7012	Exônico	284058	-	-	Downregulated
7063	Exônico	10489	LRRC41	leucine rich repeat containing 41	Downregulated
7071	Exônico	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
7074	Exônico	8540	AGPS	alkylglycerone phosphate synthase	Downregulated
7096	Exônico	2888	GRB14	growth factor receptor-bound protein 14	Downregulated
7107	Exônico	92	ACVR2A	activin A receptor, type IIA	Downregulated
7136	Exônico	57799	RAB40C	RAB40C, member RAS oncogene family	Downregulated
7141	Exônico	55737	VPS35	vacuolar protein sorting 35 (yeast)	Downregulated
7149	Exônico	51271	UBAP1	ubiquitin associated protein 1	Downregulated
7150	Exônico	51000	SLC35B3	solute carrier family 35, member B3	Downregulated
7156	Exônico	23568	ARL2BP	ADP-ribosylation factor-like 2 binding protein	Downregulated
7167	Exônico	10263	CDK2AP2	CDK2-associated protein 2	Downregulated
7172	Exônico	9555	H2AFY	H2A histone family, member Y	Downregulated
7188	Exônico	5597	MAPK6	mitogen-activated protein kinase 6	Downregulated
7192	Exônico	4798	NFRKB	nuclear factor related to kappaB binding protein	Downregulated
7237	Exônico	64771	C6orf106	chromosome 6 open reading frame 106	Downregulated
7244	Exônico	55755	CDK5RAP2	CDK5 regulatory subunit associated protein 2	Downregulated
7248	Exônico	54948	MRPL16	mitochondrial ribosomal protein L16	Downregulated
7251	Exônico	51602	-	-	Downregulated
7252	Exônico	51312	SLC25A37	solute carrier family 25, member 37	Downregulated
7258	Exônico	25865	PRKD2	protein kinase D2	Downregulated
7259	Exônico	23609	MKRN2	makorin, ring finger protein, 2	Downregulated
7270	Exônico	10295	BCKDK	branched chain ketoacid dehydrogenase kinase	Downregulated
7282	Exônico	7874	USP7	ubiquitin specific protease 7 (herpes virus-associated)	Downregulated
7289	Exônico	5984	RFC4	replication factor C (activator 1) 4, 37kDa	Downregulated
7294	Exônico	5074	PAWR	PRKC, apoptosis, WT1, regulator	Downregulated
7309	Exônico	996	CDC27	cell division cycle 27	Downregulated
7352	Exônico	54733	SLC35F2	solute carrier family 35, member F2	Downregulated
7354	Exônico	51564	HDAC7A	histone deacetylase 7A	Downregulated
7364	Exônico	23273	KIAA0367	KIAA0367	Downregulated
7379	Exônico	9341	VAMP3	vesicle-associated membrane protein 3 (cellubrevin)	Downregulated
7382	Exônico	8666	EIF3S4	eukaryotic translation initiation factor 3, subunit 4 delta, 44kDa	Downregulated

7402	Exônico	3710	ITPR3	inositol 1,4,5-triphosphate receptor, type 3	Downregulated
7409	Exônico	1786	DNMT1	DNA (cytosine-5)-methyltransferase 1	Downregulated
7438	Exônico	81706	PPP1R14C	protein phosphatase 1, regulatory (inhibitor) subunit 14C	Downregulated
7442	Exônico	79109	MAPKAP1	mitogen-activated protein kinase associated protein 1	Downregulated
7443	Exônico	64771	C6orf106	chromosome 6 open reading frame 106	Downregulated
7458	Exônico	51312	SLC25A37	solute carrier family 25, member 37	Downregulated
7476	Exônico	10295	BCKDK	branched chain ketoacid dehydrogenase kinase	Downregulated
7477	Exônico	10138	YAF2	YY1 associated factor 2	Downregulated
7482	Exônico	9373	PLAA	phospholipase A2-activating protein	Downregulated
7483	Exônico	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
7489	Exônico	7416	VDAC1	voltage-dependent anion channel 1	Downregulated
7501	Exônico	4848	CNOT2	CCR4-NOT transcription complex, subunit 2	Downregulated
7512	Exônico	1819	DRG2	developmentally regulated GTP binding protein 2	Downregulated
7538	Exônico	84883	AMID	apoptosis-inducing factor (AIF)-like mitochondrion-associated inducer of death	Downregulated
7555	Exônico	55293	-	-	Downregulated
7565	Exônico	26512	DDX26	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 26	Downregulated
7572	Exônico	23078	-	-	Downregulated
7586	Exônico	9123	SLC16A3	solute carrier family 16 (monocarboxylic acid transporters), member 3	Downregulated
7601	Exônico	5428	POLG	polymerase (DNA directed), gamma	Downregulated
7645	Exônico	80314	EPC1	enhancer of polycomb homolog 1 (Drosophila)	Downregulated
7653	Exônico	57404	CYP20A1	cytochrome P450, family 20, subfamily A, polypeptide 1	Downregulated
7658	Exônico	55326	AGPAT5	1-acylglycerol-3-phosphate O-acyltransferase 5 (lysophosphatidic acid acyltransferase, epsilon)	Downregulated
7671	Exônico	23609	MKRN2	makorin, ring finger protein, 2	Downregulated
7676	Exônico	22998	-	-	Downregulated
7682	Exônico	10295	BCKDK	branched chain ketoacid dehydrogenase kinase	Downregulated
7686	Exônico	9703	-	-	Downregulated
7688	Exônico	9375	TM9SF2	transmembrane 9 superfamily member 2	Downregulated
7696	Exônico	7184	TRA1	tumor rejection antigen (gp96) 1	Downregulated
7706	Exônico	5074	PAWR	PRKC, apoptosis, WT1, regulator	Downregulated
7719	Exônico	1573	CYP2J2	cytochrome P450, family 2, subfamily J, polypeptide 2	Downregulated
7739	Exônico	132001	-	-	Downregulated
7778	Exônico	23078	-	-	Downregulated

7779	Exônico	22976	PAXIP1	PAX interacting (with transcription-activation domain) protein 1	Downregulated
7783	Exônico	10625	IVNS1ABP	influenza virus NS1A binding protein	Downregulated
7793	Exônico	8857	FCGBP	Fc fragment of IgG binding protein	Downregulated
7821	Exônico	1786	DNMT1	DNA (cytosine-5-)-methyltransferase 1	Downregulated
7841	Exônico	146956	EME1	essential meiotic endonuclease 1 homolog 1 (S. pombe)	Downregulated
7851	Exônico	80314	EPC1	enhancer of polycomb homolog 1 (Drosophila)	Downregulated
7867	Exônico	54780	C10orf86	chromosome 10 open reading frame 86	Downregulated
7896	Exônico	8888	MCM3AP	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae) associated protein	Downregulated
7905	Exônico	6597	SMARCA4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4	Downregulated
7945	Exônico	132001	-	-	Downregulated
7964	Exônico	55872	PBK	PDZ binding kinase	Downregulated
7975	Exônico	29789	-	-	Downregulated
7986	Exônico	11329	STK38	serine/threonine kinase 38	Downregulated
8000	Exônico	8665	EIF3S5	eukaryotic translation initiation factor 3, subunit 5 epsilon, 47kDa	Downregulated
8013	Exônico	5428	POLG	polymerase (DNA directed), gamma	Downregulated
8016	Exônico	4792	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha	Downregulated
8056	Exônico	81790	RNF170	ring finger protein 170	Downregulated
8063	Exônico	58490	C20orf77	chromosome 20 open reading frame 77	Downregulated
8065	Exônico	57404	CYP20A1	cytochrome P450, family 20, subfamily A, polypeptide 1	Downregulated
8066	Exônico	56947	C2orf33	chromosome 2 open reading frame 33	Downregulated
8080	Exônico	26960	NBEA	neurobeachin	Downregulated
8084	Exônico	23392	KIAA0368	KIAA0368	Downregulated
8091	Exônico	10919	EHMT2	euchromatic histone-lysine N-methyltransferase 2	Downregulated
8093	Exônico	10491	CRTAP	cartilage associated protein	Downregulated
8095	Exônico	10140	TOB1	transducer of ERBB2, 1	Downregulated
8104	Exônico	8540	AGPS	alkylglycerone phosphate synthase	Downregulated
8133	Exônico	998	CDC42	cell division cycle 42 (GTP binding protein, 25kDa)	Downregulated
8157	Exônico	84340	GFM2	G elongation factor, mitochondrial 2	Downregulated
8166	Exônico	57798	GATAD1	GATA zinc finger domain containing 1	Downregulated
8196	Exônico	10460	TACC3	transforming, acidic coiled-coil containing protein 3	Downregulated
8230	Exônico	2562	GABRB3	gamma-aminobutyric acid (GABA) A receptor, beta 3	Downregulated

8233	Exônico	1785	DNM2	dynamín 2	Downregulated
8251	Exônico	171023	ASXL1	additional sex combs like 1 (Drosophila)	Downregulated
8253	Exônico	146956	EME1	essential meiotic endonuclease 1 homolog 1 (S. pombe)	Downregulated
8276	Exônico	55329	-	-	Downregulated
8283	Exônico	51061	TXNDC11	thioredoxin domain containing 11	Downregulated
8284	Exônico	29904	EEF2K	eukaryotic elongation factor 2 kinase	Downregulated
8285	Exônico	27340	-	-	Downregulated
8298	Exônico	10657	KHDRBS1	KH domain containing, RNA binding, signal transduction associated 1	Downregulated
8314	Exônico	7184	TRA1	tumor rejection antigen (gp96) 1	Downregulated
8316	Exônico	6790	STK6	serine/threonine kinase 6 SWI/SNF related, matrix associated, actin	Downregulated
8317	Exônico	6598	SMARCB1	dependent regulator of chromatin, subfamily b, member 1	Downregulated
8322	Exônico	5468	PPARG	peroxisome proliferative activated receptor, gamma	Downregulated
8333	Exônico	2621	GAS6	growth arrest-specific 6	Downregulated
8335	Exônico	2054	EPIM	epimorphin	Downregulated
8339	Exônico	998	CDC42	cell division cycle 42 (GTP binding protein, 25kDa)	Downregulated
8343	Exônico	92	ACVR2A	activin A receptor, type IIA	Downregulated
8359	Exônico	114882	OSBPL8	oxysterol binding protein- like 8	Downregulated
8368	Exônico	79728	-	-	Downregulated
8369	Exônico	79064	-	-	Downregulated
8379	Exônico	55291	C11orf23	chromosome 11 open reading frame 23	Downregulated
8382	Exônico	54732	TMED9	transmembrane emp24 protein transport domain containing 9	Downregulated
8384	Exônico	51560	RAB6B	RAB6B, member RAS oncogene family	Downregulated
8385	Exônico	51256	TBC1D7	TBC1 domain family, member 7	Downregulated
8425	Exônico	5427	POLE2	polymerase (DNA directed), epsilon 2 (p59 subunit)	Downregulated
8426	Exônico	5205	ATP8B1	ATPase, Class I, type 8B, member 1	Downregulated
8430	Exônico	4204	MECP2	methyl CpG binding protein 2 (Rett syndrome)	Downregulated
8455	Exônico	254128	-	-	Downregulated
8470	Exônico	80006	-	-	Downregulated
8477	Exônico	57407	-	-	Downregulated
8480	Exônico	55755	CDK5RAP2	CDK5 regulatory subunit associated protein 2	Downregulated
8481	Exônico	55622	-	-	Downregulated
8502	Exônico	11124	FAF1	Fas (TNFRSF6) associated factor 1	Downregulated
8507	Exônico	10140	TOB1	transducer of ERBB2, 1 ectonucleoside	Downregulated
8511	Exônico	9583	ENTPD4	triphosphate diphosphohydrolase 4	Downregulated
8536	Exônico	3516	RBPSUH	recombining binding protein suppressor of hairless (Drosophila)	Downregulated

8547	Exônico	550	AUP1	ancient ubiquitous protein 1	Downregulated
8561	Exônico	153129	-	-	Downregulated
8582	Exônico	55872	PBK	PDZ binding kinase	Downregulated
8586	Exônico	55120	FANCL	Fanconi anemia, complementation group L	Downregulated
8594	Exônico	27301	APEX2	APEX nuclease (apurinic/apyrimidinic endonuclease) 2	Downregulated
8600	Exônico	23273	KIAA0367	KIAA0367	Downregulated
8606	Exônico	10885	WDR3	WD repeat domain 3	Downregulated
8615	Exônico	9337	CNOT8	CCR4-NOT transcription complex, subunit 8	Downregulated
8625	Exônico	6772	STAT1	signal transducer and activator of transcription 1, 91kDa	Downregulated
8628	Exônico	5954	RCN1	reticulocalbin 1, EF-hand calcium binding domain	Downregulated
8698	Exônico	26974	ZNF285	zinc finger protein 285	Downregulated
8701	Exônico	23613	PRKCBP1	protein kinase C binding protein 1	Downregulated
8703	Exônico	23294	ANKS1	ankyrin repeat and sterile alpha motif domain containing 1	Downregulated
8704	Exônico	23207	PLEKHM2	pleckstrin homology domain containing, family IV (with RUN domain) member 2	Downregulated
8710	Exônico	10657	KHDRBS1	KH domain containing, RNA binding, signal transduction associated 1	Downregulated
8717	Exônico	9583	ENTPD4	ectonucleoside triphosphate diphosphohydrolase 4	Downregulated
8719	Exônico	9147	SDCCAG1	serologically defined colon cancer antigen 1	Downregulated
8724	Exônico	7874	USP7	ubiquitin specific protease 7 (herpes virus-associated)	Downregulated
8737	Exônico	4848	CNOT2	CCR4-NOT transcription complex, subunit 2	Downregulated
8738	Exônico	4644	MYO5A	myosin VA (heavy polypeptide 12, myoxin)	Downregulated
8739	Exônico	4236	MFAP1	microfibrillar-associated protein 1	Downregulated
8780	Exônico	79726	WDR59	WD repeat domain 59	Downregulated
8785	Exônico	57556	SEMA6A	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A	Downregulated
8797	Exônico	51256	TBC1D7	TBC1 domain family, member 7	Downregulated
8808	Exônico	23078	-	-	Downregulated
8813	Exônico	10621	POLR3F	polymerase (RNA) III (DNA directed) polypeptide F, 39 kDa	Downregulated
8817	Exônico	9957	HS3ST1	heparan sulfate (glucosamine) 3-O-sulfotransferase 1	Downregulated
8823	Exônico	8853	DDEF2	development and differentiation enhancing factor 2	Downregulated
8829	Exônico	7163	TPD52	tumor protein D52	Downregulated
8833	Exônico	6293	VPS52	vacuolar protein sorting 52 (yeast)	Downregulated
8871	Exônico	147007	C17orf32	chromosome 17 open reading frame 32	Downregulated
8889	Exônico	57407	-	-	Downregulated

8895	Exônico	55164	SHQ1	SHQ1 homolog (S. cerevisiae)	Downregulated
8899	Exônico	51602	-	-	Downregulated
8900	Exônico	51315	-	-	Downregulated
8901	Exônico	51061	TXNDC11	thioredoxin domain containing 11	Downregulated
8906	Exônico	25874	BRP44	brain protein 44 SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	Downregulated
8935	Exônico	6598	SMARCB1	dependent regulator of chromatin, subfamily b, member 1	Downregulated
8959	Exônico	550	AUP1	ancient ubiquitous protein 1	Downregulated
8971	Exônico	203069	-	-	Downregulated
8976	Exônico	123263	-	-	Downregulated
8983	Exônico	81602	CDADC1	cytidine and dCMP deaminase domain containing 1 sema domain,	Downregulated
8991	Exônico	57556	SEMA6A	transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A	Downregulated
9015	Exônico	22976	PAXIP1	PAX interacting (with transcription-activation domain) protein 1	Downregulated
9021	Exônico	10260	MYCPBP	c-myc promoter binding protein	Downregulated
9024	Exônico	9828	ARHGEF17	Rho guanine nucleotide exchange factor (GEF) 17	Downregulated
9027	Exônico	9337	CNOT8	CCR4-NOT transcription complex, subunit 8	Downregulated
9047	Exônico	4605	MYBL2	v-myb myeloblastosis viral oncogene homolog (avian)-like 2	Downregulated
9088	Exônico	80011	-	-	Downregulated
9097	Exônico	55914	ERBB2IP	erb2 interacting protein	Downregulated
9100	Exônico	55329	-	-	Downregulated
9103	Exônico	54788	DNAJB12	DnaJ (Hsp40) homolog, subfamily B, member 12	Downregulated
9118	Exônico	23001	WDFY3	WD repeat and FYVE domain containing 3	Downregulated
9128	Exônico	9706	ULK2	unc-51-like kinase 2 (C. elegans)	Downregulated
9129	Exônico	9584	RNPC2	RNA-binding region (RNP1, RRM) containing 2	Downregulated
9132	Exônico	8888	MCM3AP	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae) associated protein	Downregulated
9138	Exônico	7184	TRA1	tumor rejection antigen (gp96) 1	Downregulated
9154	Exônico	3516	RBPSUH	recombining binding protein suppressor of hairless (Drosophila)	Downregulated
9159	Exônico	2058	EPRS	glutamyl-prolyl-tRNA synthetase	Downregulated
9179	Exônico	153090	DAB2IP	DAB2 interacting protein	Downregulated
9196	Exônico	57794	SF4	splicing factor 4 sema domain,	Downregulated
9197	Exônico	57556	SEMA6A	transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A	Downregulated
9202	Exônico	55589	BMP2K	BMP2 inducible kinase	Downregulated
9207	Exônico	53407	STX18	syntaxin 18	Downregulated

9252	Exônico	4790	NFKB1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (p105)	Downregulated
9257	Exônico	3431	SP110	SP110 nuclear body protein	Downregulated
9283	Exônico	147007	C17orf32	chromosome 17 open reading frame 32	Downregulated
9305	Exônico	55625	ZDHC7	zinc finger, DHHC-type containing 7	Downregulated
9308	Exônico	54970	TTC12	tetratricopeptide repeat domain 12	Downregulated
9309	Exônico	54788	DNAJB12	DnaJ (Hsp40) homolog, subfamily B, member 12	Downregulated
9319	Exônico	23613	PRKCBP1	protein kinase C binding protein 1	Downregulated
9321	Exônico	23295	MGRN1	mahogunin, ring finger 1	Downregulated
9342	Exônico	7879	RAB7	RAB7, member RAS oncogene family	Downregulated
9352	Exônico	5469	PPARBP	PPAR binding protein	Downregulated
9355	Exônico	4849	CNOT3	CCR4-NOT transcription complex, subunit 3	Downregulated
9357	Exônico	4236	MFAP1	microfibrillar-associated protein 1	Downregulated
9383	Exônico	203069	-	-	Downregulated
9421	Exônico	25831	HECTD1	HECT domain containing 1	Downregulated
9422	Exônico	23560	GTPBP4	GTP binding protein 4	Downregulated
9432	Exônico	10459	MAD2L2	MAD2 mitotic arrest deficient-like 2 (yeast)	Downregulated
9438	Exônico	9552	SPAG7	sperm associated antigen 7	Downregulated
9449	Exônico	6768	ST14	suppression of tumorigenicity 14 (colon carcinoma, matrilysin, epithin)	Downregulated
9460	Exônico	4201	MEA1	male-enhanced antigen 1	Downregulated
9467	Exônico	2186	FALZ	fetal Alzheimer antigen	Downregulated
9470	Exônico	1499	CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	Downregulated
9476	Exônico	71	ACTG1	actin, gamma 1	Downregulated
9517	Exônico	51603	KIAA0859	KIAA0859	Downregulated
9522	Exônico	26984	SEC22L2	SEC22 vesicle trafficking protein-like 2 (S. cerevisiae)	Downregulated
9532	Exônico	11127	KIF3A	kinesin family member 3A	Downregulated
9544	Exônico	8888	MCM3AP	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae) associated protein	Downregulated
9555	Exônico	5993	RFX5	regulatory factor X, 5 (influences HLA class II expression)	Downregulated
9569	Exônico	2631	GBAS	glioblastoma amplified sequence	Downregulated
9623	Exônico	29766	TMOD3	tropomodulin 3 (ubiquitous)	Downregulated
9636	Exônico	10885	WDR3	WD repeat domain 3	Downregulated
9640	Exônico	10098	TSPAN5	tetraspanin 5	Downregulated
9642	Exônico	9827	KIAA0258	KIAA0258	Downregulated
9661	Exônico	5427	POLE2	polymerase (DNA directed), epsilon 2 (p59 subunit)	Downregulated
9667	Exônico	3980	LIG3	ligase III, DNA, ATP-dependent	Downregulated
9701	Exônico	84916	CIRH1A	cirrhosis, autosomal recessive 1A (cirhin)	Downregulated
9714	Exônico	56956	LHX9	LIM homeobox 9	Downregulated

9745	Exônico	9857	-	-	Downregulated
9752	Exônico	8545	CGGBP1	CGG triplet repeat binding protein 1	Downregulated
9767	Exônico	4849	CNOT3	CCR4-NOT transcription complex, subunit 3	Downregulated
9772	Exônico	3516	RBPSUH	recombining binding protein suppressor of hairless (Drosophila)	Downregulated
9796	Exônico	169270	ZNF596	zinc finger protein 596	Downregulated
9799	Exônico	131965	-	-	Downregulated
9810	Exônico	79726	WDR59	WD repeat domain 59	Downregulated
9813	Exônico	63971	KIF13A	kinesin family member 13A	Downregulated
9823	Exônico	54899	PXK	PX domain containing serine/threonine kinase	Downregulated
9829	Exônico	29766	TMOD3	tropomodulin 3 (ubiquitous)	Downregulated
9833	Exônico	25829	C22orf5	chromosome 22 open reading frame 5	Downregulated
9834	Exônico	23560	GTPBP4	GTP binding protein 4	Downregulated
9838	Exônico	23077	MYCBP2	MYC binding protein 2	Downregulated
9839	Exônico	22974	TPX2	TPX2, microtubule-associated protein homolog (Xenopus laevis)	Downregulated
9844	Exônico	10459	MAD2L2	MAD2 mitotic arrest deficient-like 2 (yeast)	Downregulated
9850	Exônico	9552	SPAG7	sperm associated antigen 7	Downregulated
9854	Exônico	8662	EIF3S9	eukaryotic translation initiation factor 3, subunit 9 eta, 116kDa	Downregulated
9857	Exônico	7776	ZNF236	zinc finger protein 236	Downregulated
9866	Exônico	5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide	Downregulated
9873	Exônico	3978	LIG1	ligase I, DNA, ATP-dependent	Downregulated
9882	Exônico	1497	CTNS	cystinosis, nephropathic	Downregulated
9914	Exônico	79139	DERL1	Der1-like domain family, member 1	Downregulated
9919	Exônico	57410	SCYL1	SCY1-like 1 (S. cerevisiae)	Downregulated
9926	Exônico	54970	TTC12	tetratricopeptide repeat domain 12	Downregulated
9932	Exônico	29911	HOOK2	hook homolog 2 (Drosophila)	Downregulated
9938	Exônico	23394	ADNP	activity-dependent neuroprotector	Downregulated
9940	Exônico	23210	PTDSR	phosphatidylserine receptor	Downregulated
9958	Exônico	8545	CGGBP1	CGG triplet repeat binding protein 1	Downregulated
9969	Exônico	5636	PRPSAP2	phosphoribosyl pyrophosphate synthetase-associated protein 2	Downregulated
9970	Exônico	5471	PPAT	phosphoribosyl pyrophosphate amidotransferase	Downregulated
9992	Exônico				Downregulated
9993	Exônico	201931	-	-	Downregulated
10011	Exônico	84330	ZNF414	zinc finger protein 414	Downregulated
10014	Exônico	80230	RUFY1	RUN and FYVE domain containing 1	Downregulated
10033	Exônico	51253	MRPL37	mitochondrial ribosomal protein L37	Downregulated
10093	Exônico	288	ANK3	ankyrin 3, node of Ranvier (ankyrin G)	Downregulated

10102	Exônico	284129	SLC26A11	solute carrier family 26, member 11	Downregulated
10137	Exônico	51070	NOSIP	nitric oxide synthase interacting protein	Downregulated
10150	Exônico	11128	POLR3A	polymerase (RNA) III (DNA directed) polypeptide A, 155kDa	Downregulated
10173	Exônico	5994	RFXAP	regulatory factor X-associated protein	Downregulated
10175	Exônico	5636	PRPSAP2	phosphoribosyl pyrophosphate synthetase-associated protein 2	Downregulated
10185	Exônico	3145	HMBS	hydroxymethylbilane synthase	Downregulated
10206	Exônico	253558	LYCAT	lysocardiolipin acyltransferase	Downregulated
10207	Exônico	203069	-	-	Downregulated
10218	Exônico	83986	C16orf9	chromosome 16 open reading frame 9	Downregulated
10223	Exônico	79033	PRNPIP	prion protein interacting protein	Downregulated
10232	Exônico	55585	UBE2Q	ubiquitin-conjugating enzyme E2Q (putative)	Downregulated
10249	Exônico	23187	PHLDB1	pleckstrin homology-like domain, family B, member 1	Downregulated
10251	Exônico	22955	SCMH1	sex comb on midleg homolog 1 (Drosophila)	Downregulated
10270	Exônico	7391	USF1	upstream transcription factor 1	Downregulated
10283	Exônico	4602	MYB	v-myb myeloblastosis viral oncogene homolog (avian)	Downregulated
10294	Exônico	1497	CTNS	cystinosis, nephropathic	Downregulated
10298	Exônico	488	ATP2A2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2	Downregulated
10323	Exônico	80315	CPEB4	cytoplasmic polyadenylation element binding protein 4	Downregulated
10331	Exônico	57410	SCYL1	SCY1-like 1 (S. cerevisiae)	Downregulated
10333	Exônico	55920	RCC2	regulator of chromosome condensation 2	Downregulated
10344	Exônico	29911	HOOK2	hook homolog 2 (Drosophila)	Downregulated
10352	Exônico	23210	PTDSR	phosphatidylserine receptor	Downregulated
10354	Exônico	23002	DAAM1	dishevelled associated activator of morphogenesis 1	Downregulated
10363	Exônico	9861	PSMD6	proteasome (prosome, macropain) 26S subunit, non-ATPase, 6	Downregulated
10377	Exônico	6598	SMARCB1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1	Downregulated
10386	Exônico	4646	MYO6	myosin VI	Downregulated
10393	Exônico	2632	GBE1	glucan (1,4-alpha-), branching enzyme 1 (glycogen branching enzyme, Andersen disease, glycogen storage disease type IV)	Downregulated
10395	Exônico	2059	EPS8	epidermal growth factor receptor pathway substrate 8	Downregulated

10397	Exônico	1588	CYP19A1	cytochrome P450, family 19, subfamily A, polypeptide 1	Downregulated
10441	Exônico	54899	PXK	PX domain containing serine/threonine kinase	Downregulated
10443	Exônico	53407	STX18	syntaxin 18	Downregulated
10459	Exônico	11097	NUPL2	nucleoporin like 2	Downregulated
10477	Exônico	7159	TP53BP2	tumor protein p53 binding protein, 2	Downregulated
10563	Exônico	10927	SPIN	spindlin	Downregulated
10571	Exônico	9584	RNPC2	RNA-binding region (RNP1, RRM) containing 2	Downregulated
10597	Exônico	3145	HMBS	hydroxymethylbilane synthase	Downregulated
10605	Exônico	1003	CDH5	cadherin 5, type 2, VE-cadherin (vascular epithelium)	Downregulated
10629	Exônico	84330	ZNF414	zinc finger protein 414	Downregulated
10656	Exônico	26019	UPF2	UPF2 regulator of nonsense transcripts homolog (yeast)	Downregulated
10675	Exônico	9328	GTF3C5	general transcription factor IIIC, polypeptide 5, 63kDa	Downregulated
10676	Exônico	9117	SEC22L3	SEC22 vesicle trafficking protein-like 3 (S. cerevisiae)	Downregulated
10696	Exônico	4194	MDM4	Mdm4, transformed 3T3 cell double minute 4, p53 binding protein (mouse)	Downregulated
10698	Exônico	3708	ITPR1	inositol 1,4,5-triphosphate receptor, type 1	Downregulated
10702	Exônico	2551	GABPA	GA binding protein transcription factor, alpha subunit 60kDa	Downregulated
10703	Exônico	2185	PTK2B	PTK2B protein tyrosine kinase 2 beta	Downregulated
10710	Exônico	483	ATP1B3	ATPase, Na+/K+ transporting, beta 3 polypeptide	Downregulated
10735	Exônico	80321	-	-	Downregulated
10748	Exônico	55334	SLC39A9	solute carrier family 39 (zinc transporter), member 9	Downregulated
10750	Exônico	54970	TTC12	tetratricopeptide repeat domain 12	Downregulated
10757	Exônico	27342	RABGEF1	RAB guanine nucleotide exchange factor (GEF) 1	Downregulated
10770	Exônico	10658	CUGBP1	CUG triplet repeat, RNA binding protein 1	Downregulated
10783	Exônico	8301	PICALM	phosphatidylinositol binding clathrin assembly protein	Downregulated
10789	Exônico	6599	SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	Downregulated
10822	Exônico	339487	ZBTB8OS	zinc finger and BTB domain containing 8 opposite strand	Downregulated
10825	Exônico	203069	-	-	Downregulated
10856	Exônico	51550	-	-	Downregulated
10885	Exônico	8500	PPFIA1	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 1	Downregulated

10891	Exônico	6760	SS18	synovial sarcoma translocation, chromosome 18	Downregulated
10897	Exônico	5427	POLE2	polymerase (DNA directed), epsilon 2 (p59 subunit)	Downregulated
10902	Exônico	4194	MDM4	Mdm4, transformed 3T3 cell double minute 4, p53 binding protein (mouse)	Downregulated
10961	Exônico	51071	DERA	2-deoxyribose-5-phosphate aldolase homolog (C. elegans)	Downregulated
10983	Exônico	9585	MPHOSPH1	M-phase phosphoprotein 1	Downregulated
10996	Exônico	6337	SCNN1A	sodium channel, nonvoltage gated 1 alpha	Downregulated
11008	Exônico	3516	RBPSUH	recombining binding protein suppressor of hairless (Drosophila)	Downregulated
11055	Exônico	55732	C1orf112	chromosome 1 open reading frame 112	Downregulated
11058	Exônico	55114	ARHGAP17	Rho GTPase activating protein 17	Downregulated
11061	Exônico	53373	TPCN1	two pore segment channel 1	Downregulated
11077	Exônico	11094	C9orf7	chromosome 9 open reading frame 7	Downregulated
11093	Exônico	7770	ZNF227	zinc finger protein 227	Downregulated
11094	Exônico	7390	UROS	uroporphyrinogen III synthase (congenital erythropoietic porphyria)	Downregulated
11096	Exônico	6996	TDG	thymine-DNA glycosylase	Downregulated
11097	Exônico	6749	SSRP1	structure specific recognition protein 1	Downregulated
11117	Exônico	1781	DNCI2	dynein, cytoplasmic, intermediate polypeptide 2	Downregulated
11119	Exônico	1263	PLK3	polo-like kinase 3 (Drosophila)	Downregulated
11181	Exônico	10933	MORF4L1	mortality factor 4 like 1 eukaryotic translation	Downregulated
11192	Exônico	8892	EIF2B2	initiation factor 2B, subunit 2 beta, 39kDa	Downregulated
11214	Exônico	3516	RBPSUH	recombining binding protein suppressor of hairless (Drosophila)	Downregulated
11224	Exônico	778	CACNA1F	calcium channel, voltage-dependent, alpha 1F subunit	Downregulated
11243	Exônico	114879	OSBPL5	oxysterol binding protein-like 5	Downregulated
11266	Exônico	54707	-	-	Downregulated
11267	Exônico	53373	TPCN1	two pore segment channel 1	Downregulated
11280	Exônico	23076	KIAA0179	KIAA0179	Downregulated
11287	Exônico	10257	ABCC4	ATP-binding cassette, sub-family C (CFTR/MRP), member 4	Downregulated
11288	Exônico	10097	ACTR2	ARP2 actin-related protein 2 homolog (yeast)	Downregulated
11300	Exônico	7390	UROS	uroporphyrinogen III synthase (congenital erythropoietic porphyria)	Downregulated
11301	Exônico	7155	TOP2B	topoisomerase (DNA) II beta 180kDa	Downregulated

11307	Exônico	5783	PTPN13	protein tyrosine phosphatase, non-receptor type 13 (APO-1/CD95 (Fas)-associated phosphatase)	Downregulated
11322	Exônico	2033	EP300	E1A binding protein p300	Downregulated
11387	Exônico	10933	MORF4L1	mortality factor 4 like 1	Downregulated
11390	Exônico	10314	LANCL1	LanC lantibiotic synthetase component C-like 1 (bacterial)	Downregulated
11404	Exônico	7186	TRAF2	TNF receptor-associated factor 2	Downregulated
11407	Exônico	6599	SMARCC1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1	Downregulated
11412	Exônico	5471	PPAT	phosphoribosyl pyrophosphate amidotransferase	Downregulated
11443	Exônico	203068	TUBB	tubulin, beta polypeptide	Downregulated
11461	Exônico	63916	ELMO2	engulfment and cell motility 2 (ced-12 homolog, C. elegans)	Downregulated
11468	Exônico	55578	FAM48A	family with sequence similarity 48, member A	Downregulated
11482	Exônico	23549	DNPEP	aspartyl aminopeptidase	Downregulated
11483	Exônico	23370	ARHGEF18	rho/rac guanine nucleotide exchange factor (GEF) 18	Downregulated
11487	Exônico	22948	CCT5	chaperonin containing TCP1, subunit 5 (epsilon)	Downregulated
11490	Exônico	10867	TSPAN9	tetraspanin 9	Downregulated
11509	Exônico	6749	SSRP1	structure specific recognition protein 1	Downregulated
11514	Exônico	5591	PRKDC	protein kinase, DNA-activated, catalytic polypeptide	Downregulated
11521	Exônico	3964	LGALS8	lectin, galactoside-binding, soluble, 8 (galectin 8)	Downregulated
11552	Exônico	115677	NOSTRIN	nitric oxide synthase trafficker	Downregulated
11569	Exônico	55929	DMAP1	DNA methyltransferase 1 associated protein 1	Downregulated
11600	Exônico	9712	USP6NL	USP6 N-terminal like	Downregulated
11602	Exônico	9382	COG1	component of oligomeric golgi complex 1	Downregulated
11605	Exônico	8723	SNX4	sorting nexin 4	Downregulated
11620	Exônico	5082	PDCL	phosducin-like	Downregulated
11627	Exônico	3146	HMGB1	high-mobility group box 1	Downregulated
11656	Exônico	92689	-	-	Downregulated
11664	Exônico	79720	-	-	Downregulated
11678	Exônico	54707	-	-	Downregulated
11679	Exônico	53373	TPCN1	two pore segment channel 1	Downregulated
11698	Exônico	10455	PECI	peroxisomal D3,D2-enoyl-CoA isomerase	Downregulated
11734	Exônico	2030	SLC29A1	solute carrier family 29 (nucleoside transporters), member 1	Downregulated
11788	Exônico	26993	AKAP8L	A kinase (PRKA) anchor protein 8-like	Downregulated
11796	Exônico	23008	-	-	Downregulated
11797	Exônico	22823	MTF2	metal response element binding transcription factor 2	Downregulated
11799	Exônico	10935	PRDX3	peroxiredoxin 3	Downregulated

11807	Exônico	9585	MPHOSPH1	M-phase phosphoprotein 1	Downregulated
11818	Exônico	6793	STK10	serine/threonine kinase 10	Downregulated
11824	Exônico	5476	PPGB	protective protein for beta-galactosidase (galactosialidosis)	Downregulated
11836	Exônico	2222	FDFT1	farnesyl-diphosphate farnesyltransferase 1	Downregulated
11880	Exônico	55578	FAM48A	family with sequence similarity 48, member A	Downregulated
11887	Exônico	51250	C6orf203	chromosome 6 open reading frame 203	Downregulated
11890	Exônico	27250	PDCD4	programmed cell death 4 (neoplastic transformation inhibitor)	Downregulated
11894	Exônico	23545	ATP6V0A2	ATPase, H ⁺ transporting, lysosomal V0 subunit a isoform 2	Downregulated
11896	Exônico	23271	CAMSAP1L1	calmodulin regulated spectrin-associated protein 1-like 1	Downregulated
11904	Exônico	10452	TOMM40	translocase of outer mitochondrial membrane 40 homolog (yeast)	Downregulated
11906	Exônico	10096	ACTR3	ARP3 actin-related protein 3 homolog (yeast)	Downregulated
11937	Exônico	2821	GPI	glucose phosphate isomerase	Downregulated
11950	Exônico	51291	GMIP	GEM interacting protein	Downregulated
11995	Exônico	26057	ANKRD17	ankyrin repeat domain 17	Downregulated
11998	Exônico	23397	BRRN1	barren homolog (Drosophila)	Downregulated
12010	Exônico	9987	HNRPD	heterogeneous nuclear ribonucleoprotein D-like	Downregulated
12014	Exônico	9382	COG1	component of oligomeric golgi complex 1	Downregulated
12017	Exônico	8725	C19orf2	chromosome 19 open reading frame 2	Downregulated
12039	Exônico	3148	HMGB2	high-mobility group box 2	Downregulated
12085	Exônico	55732	C1orf112	chromosome 1 open reading frame 112	Downregulated
12090	Exônico	54680	-	-	Downregulated
12114	Exônico	9821	RB1CC1	RB1-inducible coiled-coil 1	Downregulated
12125	Exônico	7155	TOP2B	topoisomerase (DNA) II beta 180kDa	Downregulated
12127	Exônico	6749	SSRP1	structure specific recognition protein 1	Downregulated
12156	Exônico	79856	SNX22	sorting nexin 22	Downregulated
12188	Exônico	55760	DHX32	DEAH (Asp-Glu-Ala-His) box polypeptide 32	Downregulated
12189	Exônico	55626	-	-	Downregulated
12191	Exônico	55172	C14orf104	chromosome 14 open reading frame 104	Downregulated
12198	Exônico	29922	NME7	non-metastatic cells 7, protein expressed in (nucleoside-diphosphate kinase)	Downregulated
12203	Exônico	23613	PRKCBP1	protein kinase C binding protein 1	Downregulated
12215	Exônico	10142	AKAP9	A kinase (PRKA) anchor protein (yotiao) 9	Downregulated
12219	Exônico	9585	MPHOSPH1	M-phase phosphoprotein 1	Downregulated
12225	Exônico	8310	ACOX3	acyl-Coenzyme A oxidase 3, pristanoyl	Downregulated

12230	Exônico	6794	STK11	serine/threonine kinase 11 (Peutz-Jeghers syndrome)	Downregulated
12248	Exônico	2224	FDPS	farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, dimethylallyltranstransferas e, geranyltranstransferase)	Downregulated
12249	Exônico	2060	EPS15	epidermal growth factor receptor pathway substrate 15	Downregulated
12255	Exônico	580	BARD1	BRCA1 associated RING domain 1	Downregulated
12264	Exônico	339448	-	-	Downregulated
12274	Exônico	92675	HARS2	histidyl-tRNA synthetase 2	Downregulated
12288	Exônico	57211	GPR126	G protein-coupled receptor 126	Downregulated
12298	Exônico	51538	-	-	Downregulated
12311	Exônico	22948	CCT5	chaperonin containing TCP1, subunit 5 (epsilon)	Downregulated
12334	Exônico	6558	SLC12A2	solute carrier family 12 (sodium/potassium/chloride transporters), member 2	Downregulated
12346	Exônico	3705	ITPK1	inositol 1,3,4-triphosphate 5/6 kinase	Downregulated
12378	Exônico	90417	C15orf23	chromosome 15 open reading frame 23	Downregulated
12407	Exônico	26058	TNRC15	trinucleotide repeat containing 15	Downregulated
12409	Exônico	23613	PRKCBP1	protein kinase C binding protein 1	Downregulated
12450	Exônico	3551	IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	Downregulated
12455	Exônico	2063	NR2F6	nuclear receptor subfamily 2, group F, member 6	Downregulated
12493	Exônico	57553	MICAL3	microtubule associated monooxygenase, calponin and LIM domain containing 3	Downregulated
12497	Exônico	55731	-	-	Downregulated
12513	Exônico	23369	PUM2	pumilio homolog 2 (Drosophila)	Downregulated
12536	Exônico	7390	UROS	uroporphyrinogen III synthase (congenital erythropoietic porphyria)	Downregulated
12539	Exônico	6747	SSR3	signal sequence receptor, gamma (translocon- associated protein gamma)	Downregulated
12545	Exônico	5425	POLD2	polymerase (DNA directed), delta 2, regulatory subunit 50kDa	Downregulated
12626	Exônico	10314	LANCL1	LanC lantibiotic synthetase component C-like 1 (bacterial)	Downregulated
12628	Exônico	9987	HNRPDL	heterogeneous nuclear ribonucleoprotein D-like	Downregulated
12637	Exônico	8310	ACOX3	acyl-Coenzyme A oxidase 3, pristanoyl	Downregulated
12638	Exônico	7879	RAB7	RAB7, member RAS oncogene family	Downregulated
12650	Exônico	5083	PAX9	paired box gene 9	Downregulated

12657	Exônico	3148	HMGB2	high-mobility group box 2	Downregulated
12667	Exônico	581	BAX	BCL2-associated X protein	Downregulated
12704	Exônico	55578	FAM48A	family with sequence similarity 48, member A	Downregulated
12712	Exônico	50813	COPS7A	COP9 constitutive photomorphogenic homolog subunit 7A (Arabidopsis)	Downregulated
12722	Exônico	23075	-	-	Downregulated
12723	Exônico	22944	KIN	KIN, antigenic determinant of recA protein homolog (mouse)	Downregulated
12725	Exônico	11083	DATF1	death associated transcription factor 1	Downregulated
12744	Exônico	6950	TCP1	t-complex 1	Downregulated
12751	Exônico	5423	POLB	polymerase (DNA directed), beta	Downregulated
12807	Exônico	55627	-	-	Downregulated
12810	Exônico	54973	CPSF3L	cleavage and polyadenylation specific factor 3-like	Downregulated
12812	Exônico	54187	NANS	N-acetylneuraminic acid synthase (sialic acid synthase)	Downregulated
12814	Exônico	51317	PHF21A	PHD finger protein 21A	Downregulated
12818	Exônico	26994	RNF11	ring finger protein 11	Downregulated
12836	Exônico	9716	AQR	aquarius homolog (mouse)	Downregulated
12858	Exônico	4649	MYO9A	myosin IXA	Downregulated
12862	Exônico	3551	IKBKB	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase beta	Downregulated
12866	Exônico	2224	FDPS	farnesyl diphosphate synthase (farnesyl pyrophosphate synthetase, dimethylallyltranstransferase, geranyltranstransferase)	Downregulated
12881	Exônico	401505	C9orf105	chromosome 9 open reading frame 105	Downregulated
12889	Exônico	131118	-	-	Downregulated
12890	Exônico	122830	C14orf35	chromosome 14 open reading frame 35	Downregulated
12896	Exônico	83940	TATDN1	TatD DNase domain containing 1	Downregulated
12912	Exônico	55112	WDR60	WD repeat domain 60	Downregulated
12915	Exônico	53371	NUP54	nucleoporin 54kDa	Downregulated
12916	Exônico	51538	-	-	Downregulated
12921	Exônico	26354	GNL3	guanine nucleotide binding protein-like 3 (nucleolar)	Downregulated
12927	Exônico	23185	KIAA0217	KIAA0217	Downregulated
12933	Exônico	10615	SPAG5	sperm associated antigen 5	Downregulated
12935	Exônico	10250	SRRM1	serine/arginine repetitive matrix 1	Downregulated
12939	Exônico	9685	-	-	Downregulated
12948	Exônico	7389	UROD	uroporphyrinogen decarboxylase	Downregulated
12950	Exônico	6950	TCP1	t-complex 1	Downregulated
12969	Exônico	2181	ACSL3	acyl-CoA synthetase long-chain family member 3	Downregulated
13003	Exônico	79811	-	-	Downregulated
13006	Exônico	64116	SLC39A8	solute carrier family 39 (zinc transporter), member 8	Downregulated

13008	Exônico	57600	-	-	Downregulated
13028	Exônico	23398	PPWD1	peptidylprolyl isomerase domain and WD repeat containing 1	Downregulated
13038	Exônico	10318	TNIP1	TNFAIP3 interacting protein 1	Downregulated
13045	Exônico	9156	EXO1	exonuclease 1	Downregulated
13052	Exônico	7203	CCT3	chaperonin containing TCP1, subunit 3 (gamma)	Downregulated
13071	Exônico	2648	GCN5L2	GCN5 general control of amino-acid synthesis 5-like 2 (yeast)	Downregulated
13088	Exônico	339287	-	-	Downregulated
13101	Exônico	84320	ACBD6	acyl-Coenzyme A binding domain containing 6	Downregulated
13113	Exônico	56834	-	-	Downregulated
13124	Exônico	50813	COPS7A	COP9 constitutive photomorphogenic homolog subunit 7A (Arabidopsis)	Downregulated
13126	Exônico	27248	C2orf30	chromosome 2 open reading frame 30	Downregulated
13132	Exônico	23269	MGA	MAX gene associated KIN, antigenic determinant of recA protein homolog (mouse)	Downregulated
13135	Exônico	22944	KIN		Downregulated
13142	Exônico	10096	ACTR3	ARP3 actin-related protein 3 homolog (yeast)	Downregulated
13160	Exônico	5936	RBM4	RNA binding motif protein 4	Downregulated
13162	Exônico	5590	PRKCZ	protein kinase C, zeta	Downregulated
13166	Exônico	4784	NFIX	nuclear factor I/X (CCAAT-binding transcription factor)	Downregulated
13221	Exônico	55174	-	-	Downregulated
13225	Exônico	51622	C7orf28A	chromosome 7 open reading frame 28A	Downregulated
13232	Exônico	25879	WDSOF1	WD repeats and SOF1 domain containing eukaryotic translation	Downregulated
13252	Exônico	8893	EIF2B5	initiation factor 2B, subunit 5 epsilon, 82kDa	Downregulated
13254	Exônico	8550	MAPKAPK5	mitogen-activated protein kinase-activated protein kinase 5	Downregulated
13267	Exônico	5269	SERPINB6	serine (or cysteine) proteinase inhibitor, clade B (ovalbumin), member 6	Downregulated
13288	Exônico				Downregulated
13293	Exônico	401505	C9orf105	chromosome 9 open reading frame 105	Downregulated
13333	Exônico	26354	GNL3	guanine nucleotide binding protein-like 3 (nucleolar)	Downregulated
13345	Exônico	10615	SPAG5	sperm associated antigen 5	Downregulated
13362	Exônico	6950	TCP1	t-complex 1	Downregulated
13363	Exônico	6745	SSR1	signal sequence receptor, alpha (translocon-associated protein alpha)	Downregulated
13391	Exônico	4105	MAGEA6	melanoma antigen family A, 6	Downregulated
13400	Exônico	219927	MRPL21	mitochondrial ribosomal protein L21	Downregulated

13420	Exônico	57602	USP36	ubiquitin specific protease 36	Downregulated
13428	Exônico	54984	-	-	Downregulated
13433	Exônico	51074	-	-	Downregulated
13442	Exônico	23214	XPO6	exportin 6	Downregulated
13464	Exônico	7203	CCT3	chaperonin containing TCP1, subunit 3 (gamma)	Downregulated
13472	Exônico	5479	PIIB	peptidylprolyl isomerase B (cyclophilin B)	Downregulated
13481	Exônico	3150	HMG1	high-mobility group nucleosome binding domain 1	Downregulated
13483	Exônico	2648	GCN5L2	GCN5 general control of amino-acid synthesis 5-like 2 (yeast)	Downregulated
13484	Exônico	2230	FDX1	ferredoxin 1	Downregulated
13491	Exônico	581	BAX	BCL2-associated X protein	Downregulated
13534	Exônico	51535	PPHLN1	periphilin 1	Downregulated
13535	Exônico	51247	PAIP2	poly(A) binding protein interacting protein 2	Downregulated
13548	Exônico	11284	PNKP	polynucleotide kinase 3'-phosphatase	Downregulated
13558	Exônico	9537	TP53L1	tumor protein p53 inducible protein 11	Downregulated
13565	Exônico	7750	ZNF198	zinc finger protein 198	Downregulated
13579	Exônico	4595	MUTYH	mutY homolog (E. coli)	Downregulated
13587	Exônico	2181	ACSL3	acyl-CoA synthetase long-chain family member 3	Downregulated
13588	Exônico	2023	ENO1	enolase 1, (alpha)	Downregulated
13598	Exônico	283131	-	-	Downregulated
13602	Exônico	441951	-	-	Downregulated
13611	Exônico	124540	MSI2	musashi homolog 2 (Drosophila)	Downregulated
13614	Exônico	90488	-	-	Downregulated
13633	Exônico	55174	-	-	Downregulated
13637	Exônico	51626	-	-	Downregulated
13664	Exônico	8893	EIF2B5	eukaryotic translation initiation factor 2B, subunit 5 epsilon, 82kDa	Downregulated
13665	Exônico	8725	C19orf2	chromosome 19 open reading frame 2	Downregulated
13667	Exônico	8312	AXIN1	axin 1	Downregulated
13685	Exônico	3778	KCNMA1	potassium large conductance calcium-activated channel, subfamily M, alpha member 1	Downregulated
13712	Exônico	144108	-	-	Downregulated
13729	Exônico	57553	MICAL3	microtubule associated monooxygenase, calponin and LIM domain containing 3	Downregulated
13775	Exônico	6745	SSR1	signal sequence receptor, alpha (translocon-associated protein alpha)	Downregulated
13780	Exônico	5590	PRKCZ	protein kinase C, zeta	Downregulated
13842	Exônico	54206	-	-	Downregulated
13848	Exônico	26999	CYFIP2	cytoplasmic FMR1 interacting protein 2	Downregulated
13865	Exônico	9863	MAGI2	membrane associated guanylate kinase, WW and PDZ domain containing 2	Downregulated
13869	Exônico	9158	FIBP	fibroblast growth factor (acidic) intracellular binding protein	Downregulated

13884	Exônico	5479	PPIB	peptidylprolyl isomerase B (cyclophilin B)	Downregulated
13896	Exônico	2235	FECH	ferrochelataase (protoporphyria)	Downregulated
13898	Exônico	1827	DSCR1	Down syndrome critical region gene 1	Downregulated
13903	Exônico	585	BBS4	Bardet-Biedl syndrome 4	Downregulated
13916	Exônico	168850	-	-	Downregulated
13947	Exônico	51241	C14orf112	chromosome 14 open reading frame 112	Downregulated
13951	Exônico	26354	GNL3	guanine nucleotide binding protein-like 3 (nucleolar)	Downregulated
13968	Exônico	9819	TSC22D2	TSC22 domain family, member 2	Downregulated
13971	Exônico	9320	TRIP12	thyroid hormone receptor interactor 12	Downregulated
13979	Exônico	7155	TOP2B	topoisomerase (DNA) II beta 180kDa	Downregulated
13980	Exônico	6950	TCP1	t-complex 1	Downregulated
13994	Exônico	3705	ITPK1	inositol 1,3,4-triphosphate 5/6 kinase	Downregulated
13999	Exônico	2180	ACSL1	acyl-CoA synthetase long-chain family member 1	Downregulated
14018	Exônico	219927	MRPL21	mitochondrial ribosomal protein L21	Downregulated
14023	Exônico	124583	CANT1	calcium activated nucleotidase 1	Downregulated
14043	Exônico	55628	ZNF407	zinc finger protein 407	Downregulated
14045	Exônico	55178	RNMTL1	RNA methyltransferase like 1	Downregulated
14049	Exônico	51626	-	-	Downregulated
14058	Exônico	23398	PPWD1	peptidylprolyl isomerase domain and WD repeat containing 1	Downregulated
14060	Exônico	23215	BAT2D1	BAT2 domain containing 1	Downregulated
14077	Exônico	8726	EED	embryonic ectoderm development	Downregulated
14078	Exônico	8550	MAPKAPK5	mitogen-activated protein kinase-activated protein kinase 5	Downregulated
14131	Exônico	84319	-	-	Downregulated
14161	Exônico	23369	PUM2	pumilio homolog 2 (Drosophila)	Downregulated
14190	Exônico	5925	RB1	retinoblastoma 1 (including osteosarcoma)	Downregulated
14193	Exônico	5420	PODXL	podocalyxin-like	Downregulated
14216	Exônico	84283	-	-	Downregulated
14282	Exônico	8894	EIF2S2	eukaryotic translation initiation factor 2, subunit 2 beta, 38kDa	Downregulated
14296	Exônico	5496	PPM1G	protein phosphatase 1G (formerly 2C), magnesium-dependent, gamma isoform	Downregulated
14308	Exônico	2235	FECH	ferrochelataase (protoporphyria)	Downregulated
14318	Exônico				Downregulated
14344	Exônico	64746	ACBD3	acyl-Coenzyme A binding domain containing 3	Downregulated
14362	Exônico	27246	ZNF364	zinc finger protein 364	Downregulated
14365	Exônico	25820	ARIH1	ariadne homolog, ubiquitin-conjugating enzyme E2 binding protein, 1 (Drosophila)	Downregulated

14373	Exônico	11079	RER1	RER1 retention in endoplasmic reticulum 1 homolog (<i>S. cerevisiae</i>)	Downregulated
14393	Exônico	6742	SSBP1	single-stranded DNA binding protein 1	Downregulated
14406	Exônico	3704	ITPA	inosine triphosphatase (nucleoside triphosphate pyrophosphatase)	Downregulated
14457	Exônico	55178	RNMTL1	RNA methyltransferase like 1	Downregulated
14463	Exônico	51077	C14orf111	chromosome 14 open reading frame 111	Downregulated
14483	Exônico	9868	TOMM70A	translocase of outer mitochondrial membrane 70 homolog A (yeast)	Downregulated
14486	Exônico	9391	WDR39	WD repeat domain 39	Downregulated
14489	Exônico	8726	EED	embryonic ectoderm development	Downregulated
14496	Exônico	6804	STX1A	syntaxin 1A (brain)	Downregulated
14501	Exônico	5652	PRSS8	protease, serine, 8 (prostasin)	Downregulated
14532	Exônico	253260	-	-	Downregulated
14536	Exônico	144100	PLEKHA7	pleckstrin homology domain containing, family A member 2	Downregulated
14537	Exônico	131118	-	-	Downregulated
14542	Exônico	84864	MINA	MYC induced nuclear antigen	Downregulated
14544	Exônico	83939	-	-	Downregulated
14567	Exônico	29760	BLNK	B-cell linker	Downregulated
14598	Exônico	6949	TCOF1	Treacher Collins-Franceschetti syndrome 1	Downregulated
14603	Exônico	5782	PTPN12	protein tyrosine phosphatase, non-receptor type 12	Downregulated
14606	Exônico	5195	PEX14	peroxisomal biogenesis factor 14	Downregulated
14613	Exônico	3418	IDH2	isocitrate dehydrogenase 2 (NADP+), mitochondrial	Downregulated
14639	Exônico	147650	-	-	Downregulated
14659	Exônico	55930	MYO5C	myosin VC	Downregulated
14668	Exônico	51322	WAC	WW domain containing adaptor with coiled-coil	Downregulated
14678	Exônico	23215	BAT2D1	BAT2 domain containing 1	Downregulated
14681	Exônico	22828	RBM16	RNA binding motif protein 16	Downregulated
14688	Exônico	9989	PPP4R1	protein phosphatase 4, regulatory subunit 1	Downregulated
14690	Exônico	9727	RAB11FIP3	RAB11 family interacting protein 3 (class II)	Downregulated
14725	Exônico	1019	CDK4	cyclin-dependent kinase 4	Downregulated
14746	Exônico	92521	SPECC1	spectrin domain with coiled-coils 1	Downregulated
14761	Exônico	56731	SLC2A4RG	SLC2A4 regulator	Downregulated
14778	Exônico	23534	TNPO3	transportin 3	Downregulated
14801	Exônico	7750	ZNF198	zinc finger protein 198	Downregulated
14802	Exônico	7385	UQCRC2	ubiquinol-cytochrome c reductase core protein II	Downregulated
14805	Exônico	6742	SSBP1	single-stranded DNA binding protein 1	Downregulated
14820	Exônico	3069	HDLBP	high density lipoprotein binding protein (vigilin)	Downregulated
14827	Exônico	1213	CLTC	clathrin, heavy polypeptide (Hc)	Downregulated
14842	Exônico	219927	MRPL21	mitochondrial ribosomal protein L21	Downregulated

14869	Exônico	55182	-	-	Downregulated
14873	Exônico	51631	LUC7L2	LUC7-like 2 (<i>S. cerevisiae</i>)	Downregulated
14890	Exônico	10664	CTCF	CCCTC-binding factor (zinc finger protein)	Downregulated
14893	Exônico	10144	FAM13A1	family with sequence similarity 13, member A1	Downregulated
14895	Exônico	9868	TOMM70A	translocase of outer mitochondrial membrane 70 homolog A (yeast)	Downregulated
14897	Exônico	9589	WTAP	Wilms tumor 1 associated protein	Downregulated
14898	Exônico	9391	WDR39	WD repeat domain 39	Downregulated
14902	Exônico	8553	BHLHB2	basic helix-loop-helix domain containing, class B, 2	Downregulated
14922	Exônico	3570	IL6R	interleukin 6 receptor	Downregulated
14923	Exônico	3156	HMGCR	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	Downregulated
14941	Exônico	401466	-	-	Downregulated
14963	Exônico	63905	MANBAL	mannosidase, beta A, lysosomal-like	Downregulated
14966	Exônico	57205	ATP10D	ATPase, Class V, type 10D	Downregulated
14980	Exônico	27244	SESN1	sestrin 1	Downregulated
15010	Exônico	6949	TCOF1	Treacher Collins-Franceschetti syndrome 1	Downregulated
15020	Exônico	4775	NFATC3	nuclear factor of activated T cells, cytoplasmic, calcineurin-dependent 3	Downregulated
15032	Exônico	1495	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa	Downregulated
15083	Exônico	27347	STK39	serine threonine kinase 39 (STE20/SPS1 homolog, yeast)	Downregulated
15091	Exônico	23097	CDC2L6	cell division cycle 2-like 6 (CDK8-like)	Downregulated
15102	Exônico	9728	-	-	Downregulated
15131	Exônico	2650	GCNT1	glucosaminyl (N-acetyl) transferase 1, core 2 (beta-1,6-N-acetylglucosaminyltransferase)	Downregulated
15137	Exônico	1019	CDK4	cyclin-dependent kinase 4	Downregulated
15176	Exônico	55576	STAB2	stabilin 2	Downregulated
15203	Exônico	9948	WDR1	WD repeat domain 1	Downregulated
15215	Exônico	7150	TOP1	topoisomerase (DNA) I	Downregulated
15217	Exônico	6741	SSB	Sjogren syndrome antigen B (autoantigen La)	Downregulated
15233	Exônico	2806	GOT2	glutamic-oxaloacetic transaminase 2, mitochondrial (aspartate aminotransferase 2)	Downregulated
15253	Exônico	254887	ZDHHC23	zinc finger, DHHC-type containing 23	Downregulated
15263	Exônico	84942	WDR73	WD repeat domain 73	Downregulated
15283	Exônico	54799	MBTD1	mbt domain containing 1	Downregulated
15305	Exônico	10146	-	-	Downregulated
15307	Exônico	9869	SETDB1	SET domain, bifurcated 1	Downregulated
15309	Exônico	9589	WTAP	Wilms tumor 1 associated protein	Downregulated
15310	Exônico	9391	WDR39	WD repeat domain 39	Downregulated

15315	Exônico	8314	BAP1	BRCA1 associated protein-1 (ubiquitin carboxy-terminal hydrolase)	Downregulated
15320	Exônico	6809	STX3A	syntaxin 3A	Downregulated
15321	Exônico	6619	SNAPC3	small nuclear RNA activating complex, polypeptide 3, 50kDa	Downregulated
15376	Exônico	57724	KIAA1632	KIAA1632	Downregulated
15392	Exônico	27244	SESN1	sestrin 1	Downregulated
15406	Exônico	10447	FAM3C	family with sequence similarity 3, member C	Downregulated
15411	Exônico	9682	JMJD2A	jumonji domain containing 2A	Downregulated
15438	Exônico	3069	HDLBP	high density lipoprotein binding protein (vigilin)	Downregulated
15441	Exônico	2180	ACSL1	acyl-CoA synthetase long-chain family member 1	Downregulated
15446	Exônico	955	ENTPD6	ectonucleoside triphosphate diphosphohydrolase 6 (putative function)	Downregulated
15466	Exônico	115939	-	-	Downregulated
15470	Exônico	84524	ZC3H8	zinc finger CCCH-type containing 8	Downregulated
15485	Exônico	55632	KIAA1333	KIAA1333	Downregulated
15496	Exônico	27000	ZRF1	zuotin related factor 1	Downregulated
15541	Exônico	3156	HMGCR	3-hydroxy-3-methylglutaryl-Coenzyme A reductase	Downregulated
15581	Exônico	63901	-	-	Downregulated
15606	Exônico	23070	KIAA0082	KIAA0082	Downregulated
15616	Exônico	9816	KIAA0133	KIAA0133	Downregulated
15620	Exônico	9110	MTMR4	myotubularin related protein 4	Downregulated
15626	Exônico	7384	UQCRC1	ubiquinol-cytochrome c reductase core protein I	Downregulated
15629	Exônico	6741	SSB	Sjogren syndrome antigen B (autoantigen La)	Downregulated
15636	Exônico	5191	PEX7	peroxisomal biogenesis factor 7	Downregulated
15649	Exônico	1773	DNASE1	deoxyribonuclease I	Downregulated
15650	Exônico	1495	CTNNA1	catenin (cadherin-associated protein), alpha 1, 102kDa	Downregulated
15664	Exônico	284273	ZADH2	zinc binding alcohol dehydrogenase, domain containing 2	Downregulated
15692	Exônico	55339	WDR33	WD repeat domain 33	Downregulated
15693	Exônico	55182	-	-	Downregulated
15705	Exônico	23633	KPNA6	karyopherin alpha 6 (importin alpha 7)	Downregulated
15716	Exônico	10325	RRAGB	Ras-related GTP binding B	Downregulated
15717	Exônico	10147	SFRS14	splicing factor, arginine/serine-rich 14	Downregulated
15721	Exônico	9589	WTAP	Wilms tumor 1 associated protein	Downregulated
15725	Exônico	8729	GBF1	golgi-specific brefeldin A resistance factor 1	Downregulated
15755	Exônico	1022	CDK7	cyclin-dependent kinase 7 (MO15 homolog, Xenopus laevis, cdk-activating kinase)	Downregulated
15781	Exônico	81558	-	-	Downregulated
15784	Exônico	79718	TBL1XR1	transducin (beta)-like 1X-linked receptor 1	Downregulated

15800	Exônico	51534	C6orf55	chromosome 6 open reading frame 55	Downregulated
15807	Exônico	25816	TNFAIP8	tumor necrosis factor, alpha induced protein 8	Downregulated
15817	Exônico	10609	-	-	Downregulated
15823	Exônico	9681	DEPDC5	DEP domain containing 5	Downregulated
15825	Exônico	9319	TRIP13	thyroid hormone receptor interactor 13	Downregulated
15831	Exônico	7748	ZNF195	zinc finger protein 195	Downregulated
15843	Exônico	5050	PAFAH1B3	platelet-activating factor acetylhydrolase, isoform Ib, gamma subunit 29kDa	Downregulated
15852	Exônico	2537	G1P3	interferon, alpha-inducible protein (clone IFI-6-16)	Downregulated
15857	Exônico	1211	CLTA	clathrin, light polypeptide (Lca)	Downregulated
15889	Exônico	64786	TBC1D15	TBC1 domain family, member 15	Downregulated
15903	Exônico	51637	C14orf166	chromosome 14 open reading frame 166	Downregulated
15907	Exônico	27348	TOR1B	torsin family 1, member B (torsin B)	Downregulated
15918	Exônico	11140	CDC37	CDC37 cell division cycle 37 homolog (S. cerevisiae)	Downregulated
15928	Exônico	9392	TGFBRAP1	transforming growth factor, beta receptor associated protein 1	Downregulated
15933	Exônico	8315	BRAP	BRCA1 associated protein	Downregulated
15941	Exônico	6009	RHEB	Ras homolog enriched in brain	Downregulated
15957	Exônico	2065	ERBB3	v-erb-b2 erythroblastic leukemia viral oncogene homolog 3 (avian)	Downregulated
15990	Exônico	79718	TBL1XR1	transducin (beta)-like 1X-linked receptor 1	Downregulated
15992	Exônico	64744	-	-	Downregulated
16020	Exônico	11277	TREX1	three prime repair exonuclease 1	Downregulated
16028	Exônico	9816	KIAA0133	KIAA0133	Downregulated
16031	Exônico	9315	C5orf13	chromosome 5 open reading frame 13	Downregulated
16041	Exônico	6733	SRPK2	SFRS protein kinase 2	Downregulated
16047	Exônico	5412	UBL3	ubiquitin-like 3	Downregulated
16051	Exônico	4580	MTX1	metaxin 1	Downregulated
16059	Exônico	2177	FANCD2	Fanconi anemia, complementation group D2	Downregulated
16062	Exônico	1491	CTH	cystathionase (cystathionine gamma-lyase)	Downregulated
16093	Exônico	79811	-	-	Downregulated
16115	Exônico	26061	-	-	Downregulated
16117	Exônico	23635	SSBP2	single-stranded DNA binding protein 2	Downregulated
16155	Exônico	4255	MGMT	O-6-methylguanine-DNA methyltransferase	Downregulated
16197	Exônico	79023	NUP37	nucleoporin 37kDa	Downregulated
16199	Exônico	63898	SH2D4A	SH2 domain containing 4A	Downregulated
16212	Exônico	51534	C6orf55	chromosome 6 open reading frame 55	Downregulated
16229	Exônico	10609	-	-	Downregulated
16244	Exônico	7384	UQCRC1	ubiquinol-cytochrome c reductase core protein I	Downregulated
16246	Exônico	6945	MLX	MAX-like protein X	Downregulated

16261	Exônimo	3399	ID3	inhibitor of DNA binding 3, dominant negative helix- loop-helix protein	Downregulated
16286	Exônimo	158158	RASEF	RAS and EF hand domain containing	Downregulated
16307	Exônimo	55964	3-Sep	septin 3	Downregulated
16316	Exônimo	51337	-	-	Downregulated
16318	Exônimo	29945	ANAPC4	anaphase promoting complex subunit 4	Downregulated
16331	Exônimo	10950	BTG3	BTG family, member 3 phenylalanine-tRNA	Downregulated
16332	Exônimo	10667	FARS2	synthetase 2 (mitochondrial)	Downregulated
16334	Exônimo	10325	RRAGB	Ras-related GTP binding B	Downregulated
16338	Exônimo	9728	-	-	Downregulated
16398	Exônimo	83930	STARD3NL	STARD3 N-terminal like	Downregulated
16408	Exônimo	57194	ATP10A	ATPase, Class V, type 10A	Downregulated
16409	Exônimo	56681	SARA1	SAR1a gene homolog 1 (S. cerevisiae)	Downregulated
16413	Exônimo	55280	CWF19L1	CWF19-like 1, cell cycle control (S. pombe)	Downregulated
16422	Exônimo	27244	SESN1	sestrin 1	Downregulated
16432	Exônimo	11277	TREX1	three prime repair exonuclease 1	Downregulated
16449	Exônimo	7738	ZNF184	zinc finger protein 184 (Kruppel-like)	Downregulated
16464	Exônimo	4179	MCP	membrane cofactor protein (CD46, trophoblast- lymphocyte cross-reactive antigen)	Downregulated
16494	Exônimo	136319	MTPN	myotrophin	Downregulated
16496	Exônimo	116028	-	-	Downregulated
16515	Exônimo	55633	TBC1D22B	TBC1 domain family, member 22B	Downregulated
16517	Exônimo	55183	RIF1	RAP1 interacting factor homolog (yeast)	Downregulated
16521	Exônimo	51637	C14orf166	chromosome 14 open reading frame 166	Downregulated
16527	Exônimo	26061	-	-	Downregulated
16536	Exônimo	11145	HRASLS3	HRAS-like suppressor 3	Downregulated
16549	Exônimo	8729	GBF1	golgi-specific brefeldin A resistance factor 1	Downregulated
16562	Exônimo	5496	PPM1G	protein phosphatase 1G (formerly 2C), magnesium- dependent, gamma isoform	Downregulated
16571	Exônimo	3157	HMGCS1	3-hydroxy-3-methylglutaryl- Coenzyme A synthase 1 (soluble)	Downregulated
16573	Exônimo	2651	GCNT2	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme	Downregulated
16579	Exônimo	1024	CDK8	cyclin-dependent kinase 8	Downregulated
16616	Exônimo	55858	-	-	Downregulated
16621	Exônimo	54891	-	-	Downregulated
16624	Exônimo	51514	DTL	denticless homolog (Drosophila)	Downregulated
16627	Exônimo	29123	ANKRD11	ankyrin repeat domain 11	Downregulated
16636	Exônimo	23065	KIAA0090	KIAA0090	Downregulated
16641	Exônimo	10608	MXD4	MAX dimerization protein 4	Downregulated
16643	Exônimo	10241	-	-	Downregulated
16647	Exônimo	9678	PHF14	PHD finger protein 14	Downregulated
16648	Exônimo	9524	GPSN2	glycoprotein, synaptic 2	Downregulated

16655	Exônico	7726	TRIM26	tripartite motif-containing 26	Downregulated
16665	Exônico	5411	PNN	pinin, desmosome associated protein	Downregulated
16667	Exônico	5048	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit 45kDa	Downregulated
16674	Exônico	3069	HDLBP	high density lipoprotein binding protein (vigilin)	Downregulated
16675	Exônico	2804	GOLGB1	golgi autoantigen, golgin subfamily b, macrogolgin (with transmembrane signal), 1	Downregulated
16703	Exônico	94081	SFXN1	sideroflexin 1	Downregulated
16713	Exônico	64786	TBC1D15	TBC1 domain family, member 15	Downregulated
16714	Exônico	64147	KIF9	kinesin family member 9	Downregulated
16718	Exônico	56980	PRDM10	PR domain containing 10	Downregulated
16733	Exônico	26063	DECR2	2,4-dienoyl CoA reductase 2, peroxisomal	Downregulated
16740	Exônico	23013	SPEN	spen homolog, transcriptional regulator (Drosophila)	Downregulated
16744	Exônico	10671	DCTN6	dynactin 6	Downregulated
16757	Exônico	8318	CDC45L	CDC45 cell division cycle 45-like (S. cerevisiae)	Downregulated
16816	Exônico	64743	WDR13	WD repeat domain 13	Downregulated
16821	Exônico	56675	NRIP3	nuclear receptor interacting protein 3	Downregulated
16823	Exônico	55727	BTBD7	BTB (POZ) domain containing 7	Downregulated
16844	Exônico	11277	TREX1	three prime repair exonuclease 1	Downregulated
16855	Exônico	9315	C5orf13	chromosome 5 open reading frame 13	Downregulated
16875	Exônico	4552	MTRR	5-methyltetrahydrofolate-homocysteine methyltransferase reductase	Downregulated
16925	Exônico	55967	-	-	Downregulated
16930	Exônico	54991	-	-	Downregulated
16944	Exônico	23216	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	Downregulated
16983	Exônico	3157	HMGCS1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)	Downregulated
17010	Exônico	122060	-	-	Downregulated
17024	Exônico	57721	-	-	Downregulated
17033	Exônico	54888	NSUN2	NOL1/NOP2/Sun domain family, member 2	Downregulated
17036	Exônico	51512	GTSE1	G-2 and S-phase expressed 1	Downregulated
17054	Exônico	10444	C9orf60	chromosome 9 open reading frame 60	Downregulated
17055	Exônico	10241	-	-	Downregulated
17057	Exônico	9943	OXS1	oxidative-stress responsive 1	Downregulated
17065	Exônico	8495	PPFIBP2	PTPRF interacting protein, binding protein 2 (liprin beta 2)	Downregulated
17070	Exônico	6944	TCFL1	transcription factor-like 1	Downregulated
17077	Exônico	5411	PNN	pinin, desmosome associated protein	Downregulated

17087	Exônico	2804	GOLGB1	golgi autoantigen, golgin subfamily b, macrogolgin (with transmembrane signal), 1	Downregulated
17093	Exônico	1207	CLNS1A	chloride channel, nucleotide sensitive, 1A	Downregulated
17096	Exônico	476	ATP1A1	ATPase, Na+/K+ transporting, alpha 1 polypeptide	Downregulated
17106	Exônico	284349	ZNF283	zinc finger protein 283	Downregulated
17130	Exônico	56984	TNFSF5IP1	tumor necrosis factor superfamily, member 5-induced protein 1	Downregulated
17132	Exônico	55763	SEC3L1	SEC3-like 1 (S. cerevisiae)	Downregulated
17135	Exônico	55183	RIF1	RAP1 interacting factor homolog (yeast)	Downregulated
17148	Exônico	23408	SIRT5	sirtuin (silent mating type information regulation 2 homolog) 5 (S. cerevisiae)	Downregulated
17150	Exônico	23216	TBC1D1	TBC1 (tre-2/USP6, BUB2, cdc16) domain family, member 1	Downregulated
17155	Exônico	10951	CBX1	chromobox homolog 1 (HP1 beta homolog Drosophila)	Downregulated
17156	Exônico	10671	DCTN6	dynactin 6	Downregulated
17189	Exônico	3157	HMGCS1	3-hydroxy-3-methylglutaryl-Coenzyme A synthase 1 (soluble)	Downregulated
17194	Exônico	1832	DSP	desmoplakin	Downregulated
17229	Exônico	63894	C14orf133	chromosome 14 open reading frame 133	Downregulated
17231	Exônico	57549	IGSF9	immunoglobulin superfamily, member 9	Downregulated
17235	Exônico	55727	BTBD7	BTB (POZ) domain containing 7	Downregulated
17238	Exônico	55105	GPATC2	G patch domain containing 2	Downregulated
17240	Exônico	54617	-	-	Downregulated
17249	Exônico	25814	ATXN10	ataxin 10	Downregulated
17266	Exônico	9522	SCAMP1	secretory carrier membrane protein 1	Downregulated
17284	Exônico	5187	PER1	period homolog 1 (Drosophila)	Downregulated
17288	Exônico	4176	MCM7	MCM7 minichromosome maintenance deficient 7 (S. cerevisiae)	Downregulated
17289	Exônico	3945	LDHB	lactate dehydrogenase B	Downregulated
17292	Exônico	3069	HDLBP	high density lipoprotein binding protein (vigilin)	Downregulated
17295	Exônico	2177	FANCD2	Fanconi anemia, complementation group D2	Downregulated
17332	Exônico	64149	-	-	Downregulated
17334	Exônico	57609	-	-	Downregulated
17339	Exônico	55634	ZNF673	zinc finger protein 673	Downregulated
17345	Exônico	51639	-	-	Downregulated
17346	Exônico	51341	ZBTB7A	zinc finger and BTB domain containing 7A	Downregulated
17358	Exônico	23014	FBXO21	F-box protein 21	Downregulated
17360	Exônico	11146	GLMN	glomulin, FKBP associated protein	Downregulated
17361	Exônico	10952	SEC61B	Sec61 beta subunit	Downregulated
17385	Exônico	5663	PSEN1	presenilin 1 (Alzheimer disease 3)	Downregulated
17407	Exônico	109	ADCY3	adenylate cyclase 3	Downregulated
17417	Exônico	201973	-	-	Downregulated

17419	Exônico	152503	-	-	Downregulated
17445	Exônico	54888	NSUN2	NOL1/NOP2/Sun domain family, member 2	Downregulated
17448	Exônico	51512	GTSE1	G-2 and S-phase expressed 1	Downregulated
17449	Exônico	51231	VRK3	vaccinia related kinase 3	Downregulated
17465	Exônico	10602	CDC42EP3	CDC42 effector protein (Rho GTPase binding) 3	Downregulated
17472	Exônico	9522	SCAMP1	secretory carrier membrane protein 1	Downregulated
17476	Exônico	8653	DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	Downregulated
17486	Exônico	5921	RASA1	RAS p21 protein activator (GTPase activating protein) 1	Downregulated
17488	Exônico	5586	PKN2	protein kinase N2	Downregulated
17542	Exônico	56984	TNFSF5IP1	tumor necrosis factor superfamily, member 5-induced protein 1	Downregulated
17543	Exônico	55970	GNG12	guanine nucleotide binding protein (G protein), gamma 12	Downregulated
17548	Exônico	54995	-	-	Downregulated
17578	Exônico	8899	PRPF4B	PRP4 pre-mRNA processing factor 4 homolog B (yeast)	Downregulated
17586	Exônico	6827	SUPT4H1	suppressor of Ty 4 homolog 1 (S. cerevisiae)	Downregulated
17587	Exônico	6627	SNRPA1	small nuclear ribonucleoprotein polypeptide A'	Downregulated
17625	Exônico	152503	-	-	Downregulated
17640	Exônico	64743	WDR13	WD repeat domain 13	Downregulated
17641	Exônico	63893	-	-	Downregulated
17665	Exônico	23178	PASK	PAS domain containing serine/threonine kinase	Downregulated
17666	Exônico	23065	KIAA0090	KIAA0090	Downregulated
17672	Exônico	10444	C9orf60	chromosome 9 open reading frame 60	Downregulated
17673	Exônico	10241	-	-	Downregulated
17678	Exônico	1977	EIF4E	eukaryotic translation initiation factor 4E	Downregulated
17693	Exônico	5780	PTPN9	protein tyrosine phosphatase, non-receptor type 9	Downregulated
17709	Exônico	1763	DNA2L	DNA2 DNA replication helicase 2-like (yeast)	Downregulated
17718	Exônico	92597	MOBK1A	MOB1, Mps One Binder kinase activator-like 1A (yeast)	Downregulated
17744	Exônico	64149	-	-	Downregulated
17745	Exônico	58513	EPS15L1	epidermal growth factor receptor pathway substrate 15-like 1	Downregulated
17751	Exônico	55635	DEPDC1	DEP domain containing 1	Downregulated
17753	Exônico	55183	RIF1	RAP1 interacting factor homolog (yeast)	Downregulated
17756	Exônico	54407	SLC38A2	solute carrier family 38, member 2	Downregulated
17757	Exônico	51639	-	-	Downregulated
17763	Exônico	26065	FAM61A	family with sequence similarity 61, member A	Downregulated
17779	Exônico	9869	SETDB1	SET domain, bifurcated 1	Downregulated
17783	Exônico	9169	SFRS2IP	splicing factor, arginine/serine-rich 2, interacting protein	Downregulated

17784	Exônimo	8899	PRPF4B	PRP4 pre-mRNA processing factor 4 homolog B (yeast)	Downregulated
17787	Exônimo	8334	HIST1H2AC	histone 1, H2ac	Downregulated
17795	Exônimo	6016	RIT1	Ras-like without CAAX 1	Downregulated
17798	Exônimo	5498	PPOX	protoporphyrinogen oxidase	Downregulated
17826	Exônimo	338692	-	-	Downregulated
17834	Exônimo	121665	-	-	Downregulated
17864	Exônimo	27230	-	-	Downregulated
17871	Exônimo	23178	PASK	PAS domain containing serine/threonine kinase	Downregulated
17875	Exônimo	11073	TOPBP1	topoisomerase (DNA) II binding protein 1	Downregulated
17888	Exônimo	8653	DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	Downregulated
17900	Exônimo	5584	PRKCI	protein kinase C, iota	Downregulated
17907	Exônimo	3945	LDHB	lactate dehydrogenase B	Downregulated
17910	Exônimo	3069	HDLBP	high density lipoprotein binding protein (vigilin)	Downregulated
17911	Exônimo	2804	GOLGB1	golgi autoantigen, golgin subfamily b, macrogolgin (with transmembrane signal), 1	Downregulated
17913	Exônimo	2176	FANCC	Fanconi anemia, complementation group C	Downregulated
17914	Exônimo	2013	EMP2	epithelial membrane protein 2	Downregulated
17948	Exônimo	79169	C1orf35	chromosome 1 open reading frame 35	Downregulated
17960	Exônimo	54996	MOSC2	MOCO sulphurase C-terminal domain containing 2	Downregulated
17961	Exônimo	54802	TRIT1	tRNA isopentenyltransferase 1	Downregulated
17966	Exônimo	29954	POMT2	protein-O-mannosyltransferase 2	Downregulated
17976	Exônimo	23015	-	-	Downregulated
18034	Exônimo	246721	-	-	Downregulated
18036	Exônimo	167227	DCP2	DCP2 decapping enzyme homolog (S. cerevisiae)	Downregulated
18052	Exônimo	64710	-	-	Downregulated
18055	Exônimo	57545	-	-	Downregulated
18083	Exônimo	10594	PRPF8	PRP8 pre-mRNA processing factor 8 homolog (yeast)	Downregulated
18090	Exônimo	9521	EEF1E1	eukaryotic translation elongation factor 1 epsilon 1	Downregulated
18097	Exônimo	7709	ZBTB17	zinc finger and BTB domain containing 17	Downregulated
18099	Exônimo	7150	TOP1	topoisomerase (DNA) I	Downregulated
18114	Exônimo	3693	ITGB5	integrin, beta 5	Downregulated
18119	Exônimo	2176	FANCC	Fanconi anemia, complementation group C	Downregulated
18121	Exônimo	1763	DNA2L	DNA2 DNA replication helicase 2-like (yeast)	Downregulated
18129	Exônimo	55623	THUMPD1	THUMP domain containing 1	Downregulated
18156	Exônimo	64151	-	-	Downregulated
18158	Exônimo	57609	-	-	Downregulated
18162	Exônimo	55764	WDR10	WD repeat domain 10	Downregulated
18179	Exônimo	23306	-	-	Downregulated
18190	Exônimo	9993	DGCR2	DiGeorge syndrome critical region gene 2	Downregulated
18192	Exônimo	9730	-	-	Downregulated

18210	Exônico	5498	PPOX	protoporphyrinogen oxidase	Downregulated
18227	Exônico	1027	CDKN1B	cyclin-dependent kinase inhibitor 1B (p27, Kip1)	Downregulated
18229	Exônico	594	BCKDHB	branched chain keto acid dehydrogenase E1, beta polypeptide (maple syrup urine disease)	Downregulated
18231	Exônico	112	ADCY6	adenylate cyclase 6	Downregulated
18232	Exônico	1746	DLX2	distal-less homeo box 2	Downregulated
18287	Exônico	11073	TOPBP1	topoisomerase (DNA) II binding protein 1	Downregulated
18301	Exônico	8493	PPM1D	protein phosphatase 1D magnesium-dependent, delta isoform	Downregulated
18319	Exônico	3939	LDHA	lactate dehydrogenase A	Downregulated
18335	Exônico	4939	OAS2	2'-5'-oligoadenylate synthetase 2, 69/71kDa	Downregulated
18346	Exônico	158234	RG9MTD3	RNA (guanine-9-) methyltransferase domain containing 3	Downregulated
18363	Exônico	58516	FAM60A	family with sequence similarity 60, member A	Downregulated
18374	Exônico	54407	SLC38A2	solute carrier family 38, member 2	Downregulated
18378	Exônico	29954	POMT2	protein-O-mannosyltransferase 2	Downregulated
18381	Exônico	26065	FAM61A	family with sequence similarity 61, member A	Downregulated
18392	Exônico	10681	GNB5	guanine nucleotide binding protein (G protein), beta 5	Downregulated
18398	Exônico	9730	-	-	Downregulated
18413	Exônico	6045	RNF2	ring finger protein 2	Downregulated
18428	Exônico	2247	FGF2	fibroblast growth factor 2 (basic)	Downregulated
18457	Exônico	84300	C6orf125	chromosome 6 open reading frame 125	Downregulated
18459	Exônico	81556	-	-	Downregulated
18468	Exônico	57189	-	-	Downregulated
18477	Exônico	51809	GALNT7	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase 7 (GalNAc-T7)	Downregulated
18496	Exônico	10438	-	-	Downregulated
18510	Exônico	7376	NR1H2	nuclear receptor subfamily 1, group H, member 2	Downregulated
18568	Exônico	64151	-	-	Downregulated
18572	Exônico	56986	-	-	Downregulated
18590	Exônico	23414	ZFPM2	zinc finger protein, multitype 2	Downregulated
18591	Exônico	23306	-	-	Downregulated
18594	Exônico	23016	EXOSC7	exosome component 7	Downregulated
18600	Exônico	10327	AKR1A1	aldo-keto reductase family 1, member A1 (aldehyde reductase)	Downregulated
18603	Exônico	9870	KIAA0317	KIAA0317	Downregulated
18611	Exônico	8341	HIST1H2BN	histone 1, H2bn	Downregulated
18617	Exônico	6635	SNRPE	small nuclear ribonucleoprotein polypeptide E	Downregulated
18622	Exônico	5501	PPP1CC	protein phosphatase 1, catalytic subunit, gamma isoform	Downregulated
18631	Exônico	3161	HMMR	hyaluronan-mediated motility receptor (RHAMM)	Downregulated
18651	Exônico	283450	-	-	Downregulated

18679	Exônico	55276	PGM2	phosphoglucomutase 2	Downregulated
18703	Exônico	10238	WDR68	WD repeat domain 68	Downregulated
18707	Exônico	9678	PHF14	PHD finger protein 14	Downregulated
18708	Exônico	9520	NPEPPS	aminopeptidase puromycin sensitive	Downregulated
18714	Exônico	8140	SLC7A5	solute carrier family 7 (cationic amino acid transporter, y+ system), member 5	Downregulated
18739	Exônico	1763	DNA2L	DNA2 DNA replication helicase 2-like (yeast)	Downregulated
18782	Exônico	55342	STRBP	spermatid perinuclear RNA binding protein	Downregulated
18783	Exônico	55183	RIF1	RAP1 interacting factor homolog (yeast)	Downregulated
18795	Exônico	23637	RABGAP1	RAB GTPase activating protein 1	Downregulated
18805	Exônico	10514	MYBBP1A	MYB binding protein (P160) 1a	Downregulated
18811	Exônico	9603	NFE2L3	nuclear factor (erythroid-derived 2)-like 3	Downregulated
18824	Exônico	6390	SDHB	succinate dehydrogenase complex, subunit B, iron sulfur (lp)	Downregulated
18827	Exônico	5664	PSEN2	presenilin 2 (Alzheimer disease 4)	Downregulated
18839	Exônico	2671	GFER	growth factor, augments liver regeneration (ERV1 homolog, S. cerevisiae)	Downregulated
18858	Exônico	246330	-	-	Downregulated
18866	Exônico	92399	MRRF	mitochondrial ribosome recycling factor	Downregulated
18878	Exônico	57714	KIAA1618	KIAA1618	Downregulated
18882	Exônico	55854	-	-	Downregulated
18889	Exônico	51809	GALNT7	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylglucosaminyltransferase 7 (GalNAc-T7)	Downregulated
18894	Exônico	27229	-	-	Downregulated
18904	Exônico	11276	AP1GBP1	AP1 gamma subunit binding protein 1	Downregulated
18920	Exônico	8131	C16orf35	chromosome 16 open reading frame 35	Downregulated
18922	Exônico	7374	UNG	uracil-DNA glycosylase	Downregulated
18923	Exônico	7150	TOP1	topoisomerase (DNA) I	Downregulated
18958	Exônico	553115	PEF1	penta-EF hand domain containing 1	Downregulated
18960	Exônico	284352	-	-	Downregulated
18975	Exônico	80347	COASY	Coenzyme A synthase	Downregulated
18981	Exônico	58517	RBM25	RNA binding motif protein 25	Downregulated
18985	Exônico	55972	-	-	Downregulated
18993	Exônico	51642	MRPL48	mitochondrial ribosomal protein L48	Downregulated
19018	Exônico	9406	ZNF265	zinc finger protein 265	Downregulated
19028	Exônico	6827	SUPT4H1	suppressor of Ty 4 homolog 1 (S. cerevisiae)	Downregulated
19029	Exônico	6636	SNRPF	small nuclear ribonucleoprotein polypeptide F	Downregulated
19033	Exônico	5664	PSEN2	presenilin 2 (Alzheimer disease 4)	Downregulated
19053	Exônico	595	CCND1	cyclin D1 (PRAD1: parathyroid adenomatosis 1)	Downregulated
19065	Exônico	201627	-	-	Downregulated

19094	Exônico	54602	NDFIP2	Nedd4 family interacting protein 2	Downregulated
19102	Exônico	26009	ZZZ3	zinc finger, ZZ-type containing 3	Downregulated
19103	Exônico	25809	TTLL1	tubulin tyrosine ligase-like family, member 1	Downregulated
19112	Exônico	10845	CLPX	ClpX caseinolytic protease X homolog (E. coli)	Downregulated
19113	Exônico	10594	PRPF8	PRP8 pre-mRNA processing factor 8 homolog (yeast)	Downregulated
19119	Exônico	9677	-	-	Downregulated
19123	Exônico	8841	HDAC3	histone deacetylase 3	Downregulated
19124	Exônico	8650	NUMB	numb homolog (Drosophila)	Downregulated
19133	Exônico	6259	RYK	RYK receptor-like tyrosine kinase	Downregulated
19142	Exônico	4175	MCM6	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, S. pombe) (S. cerevisiae)	Downregulated
19146	Exônico	3066	HDAC2	histone deacetylase 2	Downregulated
19153	Exônico	1203	CLN5	ceroid-lipofuscinosis, neuronal 5	Downregulated
19166	Exônico	284352	-	-	Downregulated
19174	Exônico	116225	ZMYND19	zinc finger, MYND domain containing 19	Downregulated
19187	Exônico	58517	RBM25	RNA binding motif protein 25	Downregulated
19195	Exônico	55183	RIF1	RAP1 interacting factor homolog (yeast)	Downregulated
19197	Exônico	54802	TRIT1	tRNA isopentenyltransferase 1	Downregulated
19200	Exônico	51347	TAOK3	TAO kinase 3	Downregulated
19201	Exônico	51088	KLHL5	kelch-like 5 (Drosophila)	Downregulated
19213	Exônico	22837	COBL1	COBL-like 1	Downregulated
19215	Exônico	10963	STIP1	stress-induced-phosphoprotein 1 (Hsp70/Hsp90-organizing protein)	Downregulated
19223	Exônico	9604	RNF14	ring finger protein 14	Downregulated
19226	Exônico	8907	AP1M1	adaptor-related protein complex 1, mu 1 subunit	Downregulated
19238	Exônico	5805	PTS	6-pyruvoyltetrahydropterin synthase	Downregulated
19242	Exônico	5094	PCBP2	poly(rC) binding protein 2	Downregulated
19253	Exônico	2071	ERCC3	excision repair cross-complementing rodent repair deficiency, complementation group 3 (xeroderma pigmentosum group B complementing)	Downregulated
19257	Exônico	1033	CDKN3	cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)	Downregulated
19259	Exônico	595	CCND1	cyclin D1 (PRAD1: parathyroid adenomatosis 1)	Downregulated
19260	Exônico	323	APBB2	amyloid beta (A4) precursor protein-binding, family B, member 2 (Fe65-like)	Downregulated
19262	Exônico	3422	IDI1	isopentenyl-diphosphate delta isomerase	Downregulated
19295	Exônico	55722	-	-	Downregulated

19301	Exônico	51780	JMJD1B	jumonji domain containing 1B	Downregulated
19317	Exônico	11065	UBE2C	ubiquitin-conjugating enzyme E2C	Downregulated
19318	Exônico	10845	CLPX	ClpX caseinolytic protease X homolog (E. coli)	Downregulated
19327	Exônico	9296	ATP6V1F	ATPase, H+ transporting, lysosomal 14kDa, V1 subunit F	Downregulated
19343	Exônico	5394	EXOSC10	exosome component 10	Downregulated
19356	Exônico	2011	MARK2	MAP/microtubule affinity-regulating kinase 2	Downregulated
19357	Exônico	1763	DNA2L	DNA2 DNA replication helicase 2-like (yeast)	Downregulated
19396	Exônico	56986	-	-	Downregulated
19398	Exônico	55768	NGLY1	N-glycanase 1	Downregulated
19399	Exônico	55636	CHD7	chromodomain helicase DNA binding protein 7	Downregulated
19400	Exônico	55346	-	-	Downregulated
19405	Exônico	51643	-	-	Downregulated
19441	Exônico	6636	SNRPF	small nuclear ribonucleoprotein polypeptide F	Downregulated
19448	Exônico	5094	PCBP2	poly(rC) binding protein 2	Downregulated
19462	Exônico	1314	COPA	coatamer protein complex, subunit alpha	Downregulated
19467	Exônico	115	ADCY9	adenylate cyclase 9	Downregulated
19500	Exônico	55854	-	-	Downregulated
19502	Exônico	55572	-	-	Downregulated
19519	Exônico	23177	KIAA0582	KIAA0582	Downregulated
19524	Exônico	10845	CLPX	ClpX caseinolytic protease X homolog (E. coli)	Downregulated
19526	Exônico	10434	LYPLA1	lysophospholipase I	Downregulated
19529	Exônico	9941	ENDOGL1	endonuclease G-like 1	Downregulated
19532	Exônico	9520	NPEPPS	aminopeptidase puromycin sensitive	Downregulated
19533	Exônico	9295	SFRS11	splicing factor, arginine/serine-rich 11	Downregulated
19551	Exônico	5048	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit 45kDa	Downregulated
19558	Exônico	3066	HDAC2	histone deacetylase 2	Downregulated
19559	Exônico	2803	GOLGA4	golgi autoantigen, golgin subfamily a, 4	Downregulated
19567	Exônico	701	BUB1B	BUB1 budding uninhibited by benzimidazoles 1 homolog beta (yeast)	Downregulated
19571	Exônico	39	ACAT2	acetyl-Coenzyme A acetyltransferase 2 (acetoacetyl Coenzyme A thiolase)	Downregulated
19581	Exônico	195828	ZNF367	zinc finger protein 367	Downregulated
19624	Exônico	23019	CNOT1	CCR4-NOT transcription complex, subunit 1	Downregulated
19626	Exônico	11146	GLMN	glomulin, FKBP associated protein	Downregulated
19627	Exônico	10966	RAB40B	RAB40B, member RAS oncogene family	Downregulated
19630	Exônico	10336	PCGF3	polycomb group ring finger 3	Downregulated
19638	Exônico	8907	AP1M1	adaptor-related protein complex 1, mu 1 subunit	Downregulated
19646	Exônico	6829	SUPT5H	suppressor of Ty 5 homolog (S. cerevisiae)	Downregulated

19655	Exônico	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Downregulated
19672	Exônico	324	APC	adenomatosis polyposis coli	Downregulated
19721	Exônico	25805	BAMBI	BMP and activin membrane bound inhibitor homolog (Xenopus laevis)	Downregulated
19740	Exônico	9100	USP10	ubiquitin specific protease 10	Downregulated
19744	Exônico	8131	C16orf35	chromosome 16 open reading frame 35	Downregulated
19754	Exônico	5583	PRKCH	protein kinase C, eta	Downregulated
19755	Exônico	5394	EXOSC10	exosome component 10	Downregulated
19769	Exônico	1763	DNA2L	DNA2 DNA replication helicase 2-like (yeast)	Downregulated
19773	Exônico	699	BUB1	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	Downregulated
19818	Exônico	51360	MBTPS2	membrane-bound transcription factor protease, site 2	Downregulated
19839	Exônico	9871	SEC24D	SEC24 related gene family, member D (S. cerevisiae)	Downregulated
19917	Exônico	54881	TEX10	testis expressed sequence 10	Downregulated
19932	Exônico	23064	ALS4	amyotrophic lateral sclerosis 4	Downregulated
19934	Exônico	11273	ATXN2L	ataxin 2-like	Downregulated
19936	Exônico	10844	TUBGCP2	tubulin, gamma complex associated protein 2	Downregulated
19939	Exônico	10236	HNRPR	heterogeneous nuclear ribonucleoprotein R	Downregulated
19944	Exônico	9518	GDF15	growth differentiation factor 15	Downregulated
19954	Exônico	6938	TCF12	transcription factor 12 (HTF4, helix-loop-helix transcription factors 4)	Downregulated
19966	Exônico	4174	MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	Downregulated
19989	Exônico	374378	GALNTL4	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase-like 4	Downregulated
19998	Exônico	116225	ZMYND19	zinc finger, MYND domain containing 19	Downregulated
20005	Exônico	80349	WDR61	WD repeat domain 61	Downregulated
20027	Exônico	28231	SLCO4A1	solute carrier organic anion transporter family, member 4A1	Downregulated
20029	Exônico	26065	FAM61A	family with sequence similarity 61, member A	Downregulated
20039	Exônico	10969	EBNA1BP2	EBNA1 binding protein 2	Downregulated
20044	Exônico	9994	CASP8AP2	CASP8 associated protein 2	Downregulated
20048	Exônico	9410	WDR57	WD repeat domain 57 (U5 snRNP specific)	Downregulated
20049	Exônico	9180	OSMR	oncostatin M receptor	Downregulated
20055	Exônico	7444	VRK2	vaccinia related kinase 2	Downregulated
20062	Exônico	5817	PVR	poliovirus receptor	Downregulated
20064	Exônico	5515	PPP2CA	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	Downregulated

20070	Exônimo	4047	LSS	lanosterol synthase (2,3-oxidosqualene-lanosterol cyclase)	Downregulated
20086	Exônimo				Downregulated
20111	Exônimo	79002	-	-	Downregulated
20126	Exônimo	51504	-	-	Downregulated
20139	Exônimo	22930	-	-	Downregulated
20142	Exônimo	10844	TUBGCP2	tubulin, gamma complex associated protein 2	Downregulated
20149	Exônimo	9670	IPO13	importin 13	Downregulated
20150	Exônimo	9517	SPTLC2	serine palmitoyltransferase, long chain base subunit 2	Downregulated
20169	Exônimo	5048	PAFAH1B1	platelet-activating factor acetylhydrolase, isoform Ib, alpha subunit 45kDa	Downregulated
20176	Exônimo	3066	HDAC2	histone deacetylase 2	Downregulated
20188	Exônimo	47	ACLY	ATP citrate lyase	Downregulated
20189	Exônimo	23532	PRAME	preferentially expressed antigen in melanoma	Downregulated
20228	Exônimo	54433	NOLA1	nucleolar protein family A, member 1 (H/ACA small nucleolar RNPs)	Downregulated
20239	Exônimo	23307	KIAA0674	KIAA0674	Downregulated
20253	Exônimo	9605	C16orf7	chromosome 16 open reading frame 7	Downregulated
20264	Exônimo	6830	SUPT6H	suppressor of Ty 6 homolog (S. cerevisiae)	Downregulated
20267	Exônimo	6051	RNPEP	arginyl aminopeptidase (aminopeptidase B)	Downregulated
20273	Exônimo	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Downregulated
20301	Exônimo	201595	-	-	Downregulated
20333	Exônimo	51228	-	-	Downregulated
20341	Exônimo	23362	PSD3	pleckstrin and Sec7 domain containing 3	Downregulated
20346	Exônimo	11273	ATXN2L	ataxin 2-like	Downregulated
20357	Exônimo	9295	SFRS11	splicing factor, arginine/serine-rich 11	Downregulated
20358	Exônimo	9100	USP10	ubiquitin specific protease 10	Downregulated
20359	Exônimo	8841	HDAC3	histone deacetylase 3	Downregulated
20460	Exônimo	9410	WDR57	WD repeat domain 57 (U5 snRNP specific)	Downregulated
20483	Exônimo	3796	KIF2	kinesin heavy chain member 2	Downregulated
20506	Exônimo	246175	CNOT6L	CCR4-NOT transcription complex, subunit 6-like	Downregulated
20527	Exônimo	57541	ZNF398	zinc finger protein 398	Downregulated
20579	Exônimo	5393	EXOSC9	exosome component 9	Downregulated
20584	Exônimo	4174	MCM5	MCM5 minichromosome maintenance deficient 5, cell division cycle 46 (S. cerevisiae)	Downregulated
20632	Exônimo	56990	CDC42SE2	CDC42 small effector 2	Downregulated
20635	Exônimo	55638	-	-	Downregulated
20640	Exônimo	54433	NOLA1	nucleolar protein family A, member 1 (H/ACA small nucleolar RNPs)	Downregulated
20641	Exônimo	51643	-	-	Downregulated
20642	Exônimo	51361	HOOK1	hook homolog 1 (Drosophila)	Downregulated
20643	Exônimo	51088	KLHL5	kelch-like 5 (Drosophila)	Downregulated
20645	Exônimo	28231	SLCO4A1	solute carrier organic anion transporter family, member 4A1	Downregulated

20658	Exônico	10691	GMEB1	glucocorticoid modulatory element binding protein 1	Downregulated
20671	Exônico	8379	MAD1L1	MAD1 mitotic arrest deficient-like 1 (yeast)	Downregulated
20672	Exônico	7916	BAT2	HLA-B associated transcript 2	Downregulated
20702	Exônico	332	BIRC5	baculoviral IAP repeat-containing 5 (survivin)	Downregulated
20710	Exônico	337867	PHGDHL1	phosphoglycerate dehydrogenase like 1	Downregulated
20755	Exônico	23173	METAP1	methionyl aminopeptidase 1	Downregulated
20757	Exônico	22926	ATF6	activating transcription factor 6	Downregulated
20765	Exônico	9939	RBM8A	RNA binding motif protein 8A	Downregulated
20783	Exônico	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	Downregulated
20807	Exônico	665	BNIP3L	BCL2/adenovirus E1B 19kDa interacting protein 3-like	Downregulated
20858	Exônico	23223	KIAA0690	KIAA0690	Downregulated
20860	Exônico	23019	CNOT1	CCR4-NOT transcription complex, subunit 1	Downregulated
20863	Exônico	10971	YWHAQ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide	Downregulated
20866	Exônico	10336	PCGF3	polycomb group ring finger 3	Downregulated
20889	Exônico	5287	PIK3C2B	phosphoinositide-3-kinase, class 2, beta polypeptide	Downregulated
20891	Exônico	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Downregulated
20951	Exônico	51227	DSCR5	Down syndrome critical region gene 5	Downregulated
20973	Exônico	9667	SAFB2	scaffold attachment factor B2	Downregulated
20976	Exônico	9100	USP10	ubiquitin specific protease 10	Downregulated
20989	Exônico	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	Downregulated
20996	Exônico	4173	MCM4	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)	Downregulated
21000	Exônico	3065	HDAC1	histone deacetylase 1	Downregulated
21047	Exônico	55644	OSGEP	O-sialoglycoprotein endopeptidase	Downregulated
21057	Exônico	28231	SLCO4A1	solute carrier organic anion transporter family, member 4A1	Downregulated
21075	Exônico	9874	TLK1	tousled-like kinase 1	Downregulated
21078	Exônico	9414	TJP2	tight junction protein 2 (zonula occludens 2)	Downregulated
21084	Exônico	7916	BAT2	HLA-B associated transcript 2	Downregulated
21094	Exônico	5515	PPP2CA	protein phosphatase 2 (formerly 2A), catalytic subunit, alpha isoform	Downregulated
21100	Exônico	4048	LTA4H	leukotriene A4 hydrolase	Downregulated
21101	Exônico	3799	KIF5B	kinesin family member 5B	Downregulated

21154	Exônico	54585	LZTFL1	leucine zipper transcription factor-like 1	Downregulated
21177	Exônico	9939	RBM8A	RNA binding motif protein 8A	Downregulated
21183	Exônico	8841	HDAC3	histone deacetylase 3	Downregulated
21219	Exônico				Downregulated
21220	Exônico	133022	TRAM1L1	translocation associated membrane protein 1-like 1	Downregulated
21262	Exônico	29959	NRBP	nuclear receptor binding protein	Downregulated
21276	Exônico	10714	POLD3	polymerase (DNA-directed), delta 3, accessory subunit	Downregulated
21279	Exônico	10153	CEBPZ	CCAAT/enhancer binding protein zeta	Downregulated
21294	Exônico	6832	SUPV3L1	suppressor of var1, 3-like 1 (S. cerevisiae)	Downregulated
21299	Exônico	5684	PSMA3	proteasome (prosome, macropain) subunit, alpha type, 3	Downregulated
21327	Exônico	400569	-	-	Downregulated
21336	Exônico	121536	AEBP2	AE binding protein 2	Downregulated
21373	Exônico	23172	KIAA0157	KIAA0157	Downregulated
21390	Exônico	8649	MAP2K1IP1	mitogen-activated protein kinase kinase 1 interacting protein 1	Downregulated
21394	Exônico	7371	UCK2	uridine-cytidine kinase 2	Downregulated
21403	Exônico	5378	PMS1	PMS1 postmeiotic segregation increased 1 (S. cerevisiae)	Downregulated
21426	Exônico	27134	TJP3	tight junction protein 3 (zonula occludens 3)	Downregulated
21435	Exônico	196074	METT5D1	methyltransferase 5 domain containing 1	Downregulated
21480	Exônico	11157	LSM6	LSM6 homolog, U6 small nuclear RNA associated (S. cerevisiae)	Downregulated
21496	Exônico	7917	BAT3	HLA-B associated transcript 3	Downregulated
21506	Exônico	5516	PPP2CB	protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform	Downregulated
21559	Exônico	56650	C3orf4	chromosome 3 open reading frame 4	Downregulated
21564	Exônico	55102	C14orf103	chromosome 14 open reading frame 103	Downregulated
21566	Exônico	54585	LZTFL1	leucine zipper transcription factor-like 1	Downregulated
21568	Exônico	51497	TH1L	TH1-like (Drosophila)	Downregulated
21576	Exônico	23518	R3HDM	R3H domain (binds single-stranded nucleic acids) containing	Downregulated
21585	Exônico	10589	DRAP1	DR1-associated protein 1 (negative cofactor 2 alpha)	Downregulated
21590	Exônico	9798	-	-	Downregulated
21596	Exônico	8648	NCOA1	nuclear receptor coactivator 1	Downregulated
21613	Exônico	4522	MTHFD1	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	Downregulated
21662	Exônico	56992	KIF15	kinesin family member 15	Downregulated

21683	Exônico	23107	MRPS27	mitochondrial ribosomal protein S27	Downregulated
21701	Exônico	8394	PIP5K1A	phosphatidylinositol-4-phosphate 5-kinase, type I, alpha	Downregulated
21707	Exônico	6643	SNX2	sorting nexin 2	Downregulated
21766	Exônico	55852	-	-	Downregulated
21771	Exônico	54879	ST7L	suppression of tumorigenicity 7 like	Downregulated
21777	Exônico	29115	-	-	Downregulated
21801	Exônico	8841	HDAC3	histone deacetylase 3	Downregulated
21808	Exônico	6936	C2orf3	chromosome 2 open reading frame 3	Downregulated
21825	Exônico	2800	GOLGA1	golgi autoantigen, golgin subfamily a, 1	Downregulated
21866	Exônico	57610	RANBP10	RAN binding protein 10	Downregulated
21869	Exônico	55973	BCAP29	B-cell receptor-associated protein 29	Downregulated
21880	Exônico	29960	FTSJ2	FtsJ homolog 2 (E. coli)	Downregulated
21883	Exônico	26091	HERC4	hect domain and RLD 4	Downregulated
21891	Exônico	22847	ZNF507	zinc finger protein 507	Downregulated
21893	Exônico	10971	YWHAQ	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, theta polypeptide	Downregulated
21895	Exônico	10521	DDX17	DEAD (Asp-Glu-Ala-Asp) box polypeptide 17	Downregulated
21917	Exônico	5685	PSMA4	proteasome (prosome, macropain) subunit, alpha type, 4	Downregulated
21918	Exônico	5516	PPP2CB	protein phosphatase 2 (formerly 2A), catalytic subunit, beta isoform	Downregulated
21935	Exônico	1054	CEBPG	CCAAT/enhancer binding protein (C/EBP), gamma	Downregulated
21938	Exônico	334	APLP2	amyloid beta (A4) precursor-like protein 2	Downregulated
21954	Exônico	121512	FGD4	FYVE, RhoGEF and PH domain containing 4	Downregulated
21977	Exônico	54879	ST7L	suppression of tumorigenicity 7 like	Downregulated
21981	Exônico	51227	DSCR5	Down syndrome critical region gene 5	Downregulated
21999	Exônico	10233	-	-	Downregulated
22003	Exônico	9665	-	-	Downregulated
22006	Exônico	9097	USP14	ubiquitin specific protease 14 (tRNA-guanine transglycosylase)	Downregulated
22017	Exônico	6251	RSU1	Ras suppressor protein 1	Downregulated
22019	Exônico	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	Downregulated
22025	Exônico	4522	MTHFD1	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1, methenyltetrahydrofolate cyclohydrolase, formyltetrahydrofolate synthetase	Downregulated
22076	Exônico	55773	-	-	Downregulated
22100	Exônico	10717	AP4B1	adaptor-related protein complex 4, beta 1 subunit	Downregulated
22103	Exônico	10153	CEBPZ	CCAAT/enhancer binding protein zeta	Downregulated

22119	Exônico	6645	SNTB2	syntrophin, beta 2 (dystrophin-associated protein A1, 59kDa, basic component 2)	Downregulated
22141	Exônico	1058	CENPA	centromere protein A, 17kDa	Downregulated
22146	Exônico				Downregulated
22151	Exônico	400566	-	-	Downregulated
22175	Exônico	57530	CGN	cingulin	Downregulated
22176	Exônico	57181	SLC39A10	solute carrier family 39 (zinc transporter), member 10	Downregulated
22186	Exônico	51497	TH1L	TH1-like (Drosophila) 5,10-	Downregulated
22203	Exônico	10588	MTHFS	methenyltetrahydrofolate synthetase (5- formyltetrahydrofolate cyclo- ligase)	Downregulated
22208	Exônico	9794	MAML1	mastermind-like 1 (Drosophila)	Downregulated
22250	Exônico				Downregulated
22281	Exônico	55975	KLHL7	kelch-like 7 (Drosophila)	Downregulated
22286	Exônico	55006	-	-	Downregulated
22291	Exônico	51094	ADIPOR1	adiponectin receptor 1	Downregulated
22296	Exônico	25902	MTHFD1L	methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like	Downregulated
22301	Exônico	23108	GARNL4	GTPase activating Rap/RanGAP domain-like 4	Downregulated
22303	Exônico	22847	ZNF507	zinc finger protein 507	Downregulated
22314	Exônico	9416	DDX23	DEAD (Asp-Glu-Ala-Asp) box polypeptide 23	Downregulated
22320	Exônico	7917	BAT3	HLA-B associated transcript 3	Downregulated
22329	Exônico	5685	PSMA4	proteasome (prosome, macropain) subunit, alpha type, 4	Downregulated
22331	Exônico	5289	PIK3C3	phosphoinositide-3-kinase, class 3	Downregulated
22333	Exônico	4891	SLC11A2	solute carrier family 11 (proton-coupled divalent metal ion transporters), member 2	Downregulated
22366	Exônico	121512	FGD4	FYVE, RhoGEF and PH domain containing 4	Downregulated
22368	Exônico	92335	-	-	Downregulated
22390	Exônico	54581	SCAND2	SCAN domain containing 2	Downregulated
22438	Exônico	4173	MCM4	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)	Downregulated
22446	Exônico	2009	EML1	echinoderm microtubule associated protein like 1	Downregulated
22450	Exônico	928	CD9	CD9 antigen (p24)	Downregulated
22451	Exônico	688	KLF5	Kruppel-like factor 5 (intestinal)	Downregulated
22473	Exônico	84961	FBXL20	F-box and leucine-rich repeat protein 20	Downregulated
22508	Exônico	23020	ASCC3L1	activating signal cointegrator 1 complex subunit 3-like 1	Downregulated
22532	Exônico	6416	MAP2K4	mitogen-activated protein kinase kinase 4	Downregulated
22534	Exônico	5819	PVRL2	poliovirus receptor-related 2 (herpesvirus entry mediator B)	Downregulated
22603	Exônico	26276	VPS33B	vacuolar protein sorting 33B (yeast)	Downregulated

22616	Exônico	10427	SEC24B	SEC24 related gene family, member B (<i>S. cerevisiae</i>)	Downregulated
22622	Exônico	9517	SPTLC2	serine palmitoyltransferase, long chain base subunit 2	Downregulated
22629	Exônico	7701	ZNF142	zinc finger protein 142 (clone pHZ-49)	Downregulated
22643	Exônico	4507	MTAP	methylthioadenosine phosphorylase	Downregulated
22650	Exônico	2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma)	Downregulated
22699	Exônico	54806	AHI1	Abelson helper integration site	Downregulated
22700	Exônico	54434	SSH1	slingshot homolog 1 (<i>Drosophila</i>)	Downregulated
22711	Exônico	23309	SIN3B	SIN3 homolog B, transcription regulator (yeast)	Downregulated
22721	Exônico	10153	CEBPZ	CCAAT/enhancer binding protein zeta	Downregulated
22726	Exônico	9416	DDX23	DEAD (Asp-Glu-Ala-Asp) box polypeptide 23	Downregulated
22730	Exônico	8562	DENR	density-regulated protein	Downregulated
22732	Exônico	7917	BAT3	HLA-B associated transcript 3	Downregulated
22741	Exônico	5686	PSMA5	proteasome (prosome, macropain) subunit, alpha type, 5	Downregulated
22764	Exônico	55343	SLC35C1	solute carrier family 35, member C1	Downregulated
22784	Exônico	83857	-	-	Downregulated
22789	Exônico	78992	-	-	Downregulated
22813	Exônico	23362	PSD3	pleckstrin and Sec7 domain containing 3	Downregulated
22826	Exônico	9793	CKAP5	cytoskeleton associated protein 5	Downregulated
22833	Exônico	8476	CDC42BPA	CDC42 binding protein kinase alpha (DMPK-like)	Downregulated
22843	Exônico	5771	PTPN2	protein tyrosine phosphatase, non-receptor type 2	Downregulated
22848	Exônico	4771	NF2	neurofibromin 2 (bilateral acoustic neuroma)	Downregulated
22850	Exônico	4172	MCM3	MCM3 minichromosome maintenance deficient 3 (<i>S. cerevisiae</i>)	Downregulated
22855	Exônico	2794	GNL1	guanine nucleotide binding protein-like 1	Downregulated
22862	Exônico	908	CCT6A	chaperonin containing TCP1, subunit 6A (zeta 1)	Downregulated
22873	Exônico	374655	-	-	Downregulated
22894	Exônico	64210	MMS19L	MMS19-like (MET18 homolog, <i>S. cerevisiae</i>)	Downregulated
22899	Exônico	55975	KLHL7	kelch-like 7 (<i>Drosophila</i>)	Downregulated
22900	Exônico	55773	-	-	Downregulated
22907	Exônico	51654	CDK5RAP1	CDK5 regulatory subunit associated protein 1	Downregulated
22944	Exônico	6421	SFPQ	splicing factor proline/glutamine rich (polypyrimidine tract binding protein associated)	Downregulated
23007	Exônico	54875	C9orf39	chromosome 9 open reading frame 39	Downregulated
23013	Exônico	29110	TBK1	TANK-binding kinase 1	Downregulated

23040	Exônimo	8100	TTC10	tetratricopeptide repeat domain 10	Downregulated
23047	Exônimo	6249	RSN	restin (Reed-Steinberg cell-expressed intermediate filament-associated protein)	Downregulated
23056	Exônimo	4171	MCM2	MCM2 minichromosome maintenance deficient 2, mitotin (<i>S. cerevisiae</i>)	Downregulated
23082	Exônimo	220906	-	-	Downregulated
23105	Exônimo	55975	KLHL7	kelch-like 7 (<i>Drosophila</i>)	Downregulated
23106	Exônimo	55773	-	-	Downregulated
23123	Exônimo	23310	-	-	Downregulated
23131	Exônimo	10522	DEAF1	deformed epidermal autoregulatory factor 1 (<i>Drosophila</i>)	Downregulated
23133	Exônimo	10153	CEBPZ	CCAAT/enhancer binding protein zeta	Downregulated
23149	Exônimo	6646	SOAT1	sterol O-acyltransferase (acyl-Coenzyme A: cholesterol acyltransferase) 1	Downregulated
23165	Exônimo	2683	B4GALT1	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypeptide 1	Downregulated
23171	Exônimo	1060	CENPC1	centromere protein C 1	Downregulated
23189	Exônimo	130507	ZNF650	zinc finger protein 650	Downregulated
23194	Exônimo	84817	TXNL5	thioredoxin-like 5	Downregulated
23200	Exônimo	79711	IPO4	importin 4	Downregulated
23201	Exônimo	78991	-	-	Downregulated
23212	Exônimo	55100	WDR70	WD repeat domain 70	Downregulated
23213	Exônimo	54875	C9orf39	chromosome 9 open reading frame 39	Downregulated
23227	Exônimo	23167	-	-	Downregulated
23234	Exônimo	10427	SEC24B	SEC24 related gene family, member B (<i>S. cerevisiae</i>)	Downregulated
23244	Exônimo	8635	RNASET2	ribonuclease T2	Downregulated
23245	Exônimo	8476	CDC42BPA	CDC42 binding protein kinase alpha (DMPK-like)	Downregulated
23274	Exônimo	908	CCT6A	chaperonin containing TCP1, subunit 6A (zeta 1)	Downregulated
23285	Exônimo	374659	-	-	Downregulated
23314	Exônimo	55357	TBC1D2	TBC1 domain family, member 2	Downregulated
23324	Exônimo	27043	-	-	Downregulated
23342	Exônimo	9735	KNTC1	kinetochore associated 1 a disintegrin and	Downregulated
23347	Exônimo	8756	ADAM7	metalloproteinase domain 7	Downregulated
23403	Exônimo	81539	SLC38A1	solute carrier family 38, member 1	Downregulated
23425	Exônimo	29110	TBK1	TANK-binding kinase 1	Downregulated
23426	Exônimo	27161	EIF2C2	eukaryotic translation initiation factor 2C, 2	Downregulated
23428	Exônimo	26001	RNF167	ring finger protein 167	Downregulated
23435	Exônimo	22921	MSRB2	methionine sulfoxide reductase B2	Downregulated
23453	Exônimo	7700	ZNF141	zinc finger protein 141 (clone pHZ-44)	Downregulated
23454	Exônimo	7360	UGP2	UDP-glucose pyrophosphorylase 2	Downregulated
23474	Exônimo	2521	FUS	fusion (involved in t(12;16) in malignant liposarcoma)	Downregulated
23479	Exônimo	1182	CLCN3	chloride channel 3	Downregulated

23481	Exônico	688	KLF5	Kruppel-like factor 5 (intestinal)	Downregulated
23524	Exônico	54437	SEMA5B	sema domain, seven thrombospondin repeats (type 1 and type 1-like), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 5B	Downregulated
23528	Exônico	29966	STRN3	striatin, calmodulin binding protein 3	Downregulated
23535	Exônico	23310	-	-	Downregulated
23554	Exônico	8562	DENR	density-regulated protein transient receptor potential cation channel, subfamily M, member 2	Downregulated
23558	Exônico	7226	TRPM2	proteasome (prosome, macropain) subunit, alpha type, 5	Downregulated
23565	Exônico	5686	PSMA5	phosphoinositide-3-kinase, regulatory subunit 1 (p85 alpha)	Downregulated
23567	Exônico	5295	PIK3R1	-	Downregulated
23604	Exônico	92305	-	-	Downregulated
23612	Exônico	79711	IPO4	importin 4	Downregulated
23627	Exônico	51773	HBXAP	hepatitis B virus x associated protein	Downregulated
23637	Exônico	23360	FNBP4	formin binding protein 4	Downregulated
23648	Exônico	10079	ATP9A	ATPase, Class II, type 9A	Downregulated
23669	Exônico	5373	PMM2	phosphomannomutase 2	Downregulated
23674	Exônico	4171	MCM2	MCM2 minichromosome maintenance deficient 2, mitotin (S. cerevisiae)	Downregulated
23685	Exônico	1182	CLCN3	chloride channel 3	Downregulated
23688	Exônico	473	RERE	arginine-glutamic acid dipeptide (RE) repeats	Downregulated
23711	Exônico	84172	POLR1B	polymerase (RNA) I polypeptide B, 128kDa	Downregulated
23731	Exônico	51660	BRP44L	brain protein 44-like	Downregulated
23733	Exônico	51095	TRNT1	tRNA nucleotidyl transferase, CCA-adding, 1	Downregulated
23737	Exônico	26100	-	-	Downregulated
23738	Exônico	25904	CNOT10	CCR4-NOT transcription complex, subunit 10	Downregulated
23747	Exônico	10980	COPS6	COP9 constitutive photomorphogenic homolog subunit 6 (Arabidopsis)	Downregulated
23748	Exônico	10721	POLQ	polymerase (DNA directed), theta	Downregulated
23758	Exônico	8925	HERC1	hect (homologous to the E6-AP (UBE3A) carboxyl terminus) domain and RCC1 (CHC1)-like domain (RLD) 1	Downregulated
23814	Exônico	83786	-	-	Downregulated
23828	Exônico	55565	-	-	Downregulated
23837	Exônico	29105	-	-	Downregulated
23847	Exônico	22920	KIFAP3	kinesin-associated protein 3	Downregulated
23852	Exônico	10427	SEC24B	SEC24 related gene family, member B (S. cerevisiae)	Downregulated
23869	Exônico	6729	SRP54	signal recognition particle 54kDa	Downregulated
23873	Exônico	5770	PTPN1	protein tyrosine phosphatase, non-receptor type 1	Downregulated

23889	Exônimo	1743	DLST	dihydrolipoamide S-succinyltransferase (E2 component of 2-oxo-glutarate complex)	Downregulated
23921	Exônimo	79828	-	-	Downregulated
23931	Exônimo	55652	-	-	Downregulated
23942	Exônimo	27044	-	-	Downregulated
23959	Exônimo	9879	DDX46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	Downregulated
23962	Exônimo	9425	CDYL	chromodomain protein, Y-like	Downregulated
23996	Exônimo	800	CALD1	caldesmon 1	Downregulated
24011	Exônimo	152185	-	-	Downregulated
24017	Exônimo	89796	NAV1	neuron navigator 1	Downregulated
24038	Exônimo	54566	EPB41L4B	erythrocyte membrane protein band 4.1 like 4B	Downregulated
24059	Exônimo	10228	STX6	syntaxin 6	Downregulated
24060	Exônimo	10078	TSSC4	tumor suppressing subtransferable candidate 4	Downregulated
24072	Exônimo	7358	UGDH	UDP-glucose dehydrogenase	Downregulated
24119	Exônimo	94134	ARHGAP12	Rho GTPase activating protein 12	Downregulated
24123	Exônimo	84173	RBED1	RNA binding motif and ELMO domain 1	Downregulated
24124	Exônimo	83541	C20orf55	chromosome 20 open reading frame 55	Downregulated
24129	Exônimo	64844	7-Mar	membrane-associated ring finger (C3HC4) 7	Downregulated
24142	Exônimo	54443	ANLN	anillin, actin binding protein (scraps homolog, Drosophila)	Downregulated
24145	Exônimo	51100	SH3GLB1	SH3-domain GRB2-like endophilin B1	Downregulated
24159	Exônimo	10980	COPS6	COP9 constitutive photomorphogenic homolog subunit 6 (Arabidopsis)	Downregulated
24171	Exônimo	8756	ADAM7	a disintegrin and metalloproteinase domain 7	Downregulated
24203	Exônimo	598	BCL2L1	BCL2-like 1	Downregulated
24225	Exônimo	84284	C1orf57	chromosome 1 open reading frame 57	Downregulated
24227	Exônimo	81539	SLC38A1	solute carrier family 38, member 1	Downregulated
24230	Exônimo	79709	GLT25D1	glycosyltransferase 25 domain containing 1	Downregulated
24231	Exônimo	78990	OTUB2	OTU domain, ubiquitin aldehyde binding 2	Downregulated
24232	Exônimo	64478	CSMD1	CUB and Sushi multiple domains 1	Downregulated
24237	Exônimo	56647	BCCIP	BRCA2 and CDKN1A interacting protein	Downregulated
24245	Exônimo	51773	HBXAP	hepatitis B virus x associated protein	Downregulated
24249	Exônimo	29105	-	-	Downregulated
24253	Exônimo	25796	PGLS	6-phosphogluconolactonase	Downregulated
24257	Exônimo	23165	NUP205	nucleoporin 205kDa	Downregulated
24276	Exônimo	8099	CDK2AP1	CDK2-associated protein 1	Downregulated
24302	Exônimo	1459	CSNK2A2	casein kinase 2, alpha prime polypeptide	Downregulated
24304	Exônimo	902	CCNH	cyclin H	Downregulated
24341	Exônimo	56006	-	-	Downregulated
24346	Exônimo	55011	-	-	Downregulated

24352	Exônico	29966	STRN3	striatin, calmodulin binding protein 3	Downregulated
24366	Exônico	10724	MGEA5	meningioma expressed antigen 5 (hyaluronidase)	Downregulated
24368	Exônico	10350	ABCA9	ATP-binding cassette, sub-family A (ABC1), member 9	Downregulated
24385	Exônico	6646	SOAT1	sterol O-acyltransferase (acyl-Coenzyme A: cholesterol acyltransferase) 1	Downregulated
24392	Exônico	5097	PCDH1	protocadherin 1 (cadherin-like 1)	Downregulated
24444	Exônico	55844	PPP2R2D	protein phosphatase 2, regulatory subunit B, delta isoform	Downregulated
24447	Exônico	55268	ECHDC2	enoyl Coenzyme A hydratase domain containing 2	Downregulated
24462	Exônico	23259	DDHD2	DDHD domain containing 2	Downregulated
24464	Exônico	23059	CLUAP1	clusterin associated protein 1	Downregulated
24483	Exônico	7681	MKRN3	makorin, ring finger protein, 3	Downregulated
24487	Exônico	6728	SRP19	signal recognition particle 19kDa	Downregulated
24494	Exônico	5165	PKD3	pyruvate dehydrogenase kinase, isoenzyme 3	Downregulated
24503	Exônico	2782	GNB1	guanine nucleotide binding protein (G protein), beta polypeptide 1	Downregulated
24511	Exônico	675	BRCA2	breast cancer 2, early onset	Downregulated
24530	Exônico	116461	C1orf19	chromosome 1 open reading frame 19	Downregulated
24531	Exônico	94239	H2AFV	H2A histone family, member V	Downregulated
24540	Exônico	79575	ABHD8	abhydrolase domain containing 8	Downregulated
24542	Exônico	64210	MMS19L	MMS19-like (MET18 homolog, S. cerevisiae)	Downregulated
24549	Exônico	55655	NALP2	NACHT, leucine rich repeat and PYD containing 2	Downregulated
24577	Exônico	9879	DDX46	DEAD (Asp-Glu-Ala-Asp) box polypeptide 46	Downregulated
24583	Exônico	8761	PABPC4	poly(A) binding protein, cytoplasmic 4 (inducible form)	Downregulated
24595	Exônico	5686	PSMA5	proteasome (prosome, macropain) subunit, alpha type, 5	Downregulated
24599	Exônico	4893	NRAS	neuroblastoma RAS viral (v-ras) oncogene homolog	Downregulated
24613	Exônico	1060	CENPC1	centromere protein C 1	Downregulated
24615	Exônico	598	BCL2L1	BCL2-like 1	Downregulated
24640	Exônico	80205	CHD9	chromodomain helicase DNA binding protein 9	Downregulated
24661	Exônico	29102	RNASEN	ribonuclease III, nuclear	Downregulated
24667	Exônico	23360	FNBP4	formin binding protein 4	Downregulated
24669	Exônico	23165	NUP205	nucleoporin 205kDa	Downregulated
24682	Exônico	9513	FXR2	fragile X mental retardation, autosomal homolog 2	Downregulated
24689	Exônico	7678	ZNF124	zinc finger protein 124 (HZF 16)	Downregulated

24690	Exônico	7357	UGCG	UDP-glucose ceramide glucosyltransferase	Downregulated
24691	Exônico	7132	TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	Downregulated
24707	Exônico	3344	HTLF	human T-cell leukemia virus enhancer factor	Downregulated
24716	Exônico	902	CCNH	cyclin H	Downregulated
24738	Exônico	90956	ADCK2	aarF domain containing kinase 2	Downregulated
24742	Exônico	83541	C20orf55	chromosome 20 open reading frame 55	Downregulated
24757	Exônico	55192	-	-	Downregulated
24758	Exônico	55011	-	-	Downregulated
24798	Exônico	6428	SFRS3	splicing factor, arginine/serine-rich 3	Downregulated
24809	Exônico	3831	KNS2	kinesin 2 60/70kDa	Downregulated
24810	Exônico	3607	FOXK2	forkhead box K2	Downregulated
24818	Exônico	1314	COPA	coatamer protein complex, subunit alpha	Downregulated
24839	Exônico	114799	ESCO1	establishment of cohesion 1 homolog 1 (S. cerevisiae)	Downregulated
24869	Exônico	26273	FBXO3	F-box protein 3	Downregulated
24877	Exônico	22919	MAPRE1	microtubule-associated protein, RP/EB family, member 1	Downregulated
24886	Exônico	9793	CKAP5	cytoskeleton associated protein 5	Downregulated
24887	Exônico	9665	-	-	Downregulated
24905	Exônico	5373	PMM2	phosphomannomutase 2	Downregulated
24956	Exônico	57614	KIAA1468	KIAA1468	Downregulated
24971	Exônico	28512	NKIRAS1	NFKB inhibitor interacting Ras-like 1	Downregulated
24984	Exônico	10724	MGEA5	meningioma expressed antigen 5 (hyaluronidase)	Downregulated
24985	Exônico	10523	CHERP	calcium homeostasis endoplasmic reticulum protein	Downregulated
24993	Exônico	9188	DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21	Downregulated
24995	Exônico	8761	PABPC4	poly(A) binding protein, cytoplasmic 4 (inducible form)	Downregulated
25008	Exônico	5519	PPP2R1B	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	Downregulated
25016	Exônico	3607	FOXK2	forkhead box K2	Downregulated
25022	Exônico	1859	DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A	Downregulated
25030	Exônico	84663	CYorf15B	chromosome Y open reading frame 15B	Downregulated
25054	Exônico	79706	PRKRIP1	PRKR interacting protein 1 (IL11 inducible)	Downregulated
25061	Exônico	56623	INPP5E	inositol polyphosphate-5-phosphatase, 72 kDa	Downregulated
25071	Exônico	51203	NUSAP1	nucleolar and spindle associated protein 1	Downregulated
25074	Exônico	27152	PDZK6	PDZ domain containing 6	Downregulated
25085	Exônico	11060	WWP2	WW domain containing E3 ubiquitin protein ligase 2	Downregulated
25103	Exônico	7132	TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	Downregulated
25112	Exônico	5164	PKD2	pyruvate dehydrogenase kinase, isoenzyme 2	Downregulated

25126	Exônico	1459	CSNK2A2	casein kinase 2, alpha prime polypeptide	Downregulated
25128	Exônico	901	CCNG2	cyclin G2	Downregulated
25150	Exônico	90956	ADCK2	aarF domain containing kinase 2	Downregulated
25152	Exônico	84552	PARD6G	par-6 partitioning defective 6 homolog gamma (C. elegans)	Downregulated
25158	Exônico	79577	HRPT2	hyperparathyroidism 2 (with jaw tumor)	Downregulated
25169	Exônico	55193	-	-	Downregulated
25175	Exônico	51101	-	-	Downregulated
25193	Exônico	10155	TRIM28	tripartite motif-containing 28	Downregulated
25198	Exônico	9439	CRSP3	cofactor required for Sp1 transcriptional activation, subunit 3, 130kDa	Downregulated
25204	Exônico	7936	RDBP	RD RNA binding protein	Downregulated
25213	Exônico	5690	PSMB2	proteasome (prosome, macropain) subunit, beta type, 2	Downregulated
25217	Exônico	4898	NRD1	nardilysin (N-arginine dibasic convertase)	Downregulated
25220	Exônico	4067	LYN	v-yes-1 Yamaguchi sarcoma viral related oncogene homolog	Downregulated
25231	Exônico	1062	CENPE	centromere protein E, 312kDa	Downregulated
25233	Exônico	604	BCL6	B-cell CLL/lymphoma 6 (zinc finger protein 51)	Downregulated
25241	Exônico	399979	SNX19	sorting nexin 19	Downregulated
25244	Exônico	222658	C6orf69	chromosome 6 open reading frame 69	Downregulated
25246	Exônico	166647	GPR125	G protein-coupled receptor 125	Downregulated
25259	Exônico	79932	KIAA0319L	KIAA0319-like	Downregulated
25289	Exônico	22919	MAPRE1	microtubule-associated protein, RP/EB family, member 1	Downregulated
25291	Exônico	11060	WWP2	WW domain containing E3 ubiquitin protein ligase 2	Downregulated
25292	Exônico	10809	STARD10	START domain containing 10	Downregulated
25297	Exônico	9931	HELZ	helicase with zinc finger	Downregulated
25299	Exônico	9663	LPIN2	lipin 2	Downregulated
25308	Exônico	7357	UGCG	UDP-glucose ceramide glucosyltransferase	Downregulated
25311	Exônico	6728	SRP19	signal recognition particle 19kDa	Downregulated
25334	Exônico	899	CCNF	cyclin F	Downregulated
25335	Exônico	673	BRAF	v-raf murine sarcoma viral oncogene homolog B1	Downregulated
25339	Exônico	79962	-	-	Downregulated
25383	Exônico	28512	NKIRAS1	NFKB inhibitor interacting Ras-like 1	Downregulated
25397	Exônico	10524	HTATIP	HIV-1 Tat interacting protein, 60kDa	Downregulated
25405	Exônico	9188	DDX21	DEAD (Asp-Glu-Ala-Asp) box polypeptide 21	Downregulated
25416	Exônico	6429	SFRS4	splicing factor, arginine/serine-rich 4	Downregulated
25418	Exônico	5829	PXN	paxillin	Downregulated
25420	Exônico	5519	PPP2R1B	protein phosphatase 2 (formerly 2A), regulatory subunit A (PR 65), beta isoform	Downregulated
25427	Exônico	3831	KNS2	kinesin 2 60/70kDa	Downregulated

25428	Exônico	3608	ILF2	interleukin enhancer binding factor 2, 45kDa	Downregulated
25437	Exônico	1062	CENPE	centromere protein E, 312kDa	Downregulated
25456	Exônico	121260	SLC15A4	solute carrier family 15, member 4	Downregulated
25457	Exônico	114793	FMNL2	formin-like 2	Downregulated
25470	Exônico	57706	KIAA1608	KIAA1608	Downregulated
25478	Exônico	55090	MED9	mediator of RNA polymerase II transcription, subunit 9 homolog (yeast)	Downregulated
25490	Exônico	23514	-	-	Downregulated
25507	Exônico	9267	PSCD1	pleckstrin homology, Sec7 and coiled-coil domains 1(cytohesin 1)	Downregulated
25519	Exônico	6241	RRM2	ribonucleotide reductase M2 polypeptide	Downregulated
25539	Exônico	1173	AP2M1	adaptor-related protein complex 2, mu 1 subunit	Downregulated
25542	Exônico	471	ATIC	5-aminoimidazole-4-carboxamide ribonucleotide formyltransferase/IMP cyclohydrolase	Downregulated
25553	Exônico	256126	-	-	Downregulated
25566	Exônico	83543	C9orf58	chromosome 9 open reading frame 58	Downregulated
25572	Exônico	64215	DNAJC1	DnaJ (Hsp40) homolog, subfamily C, member 1	Downregulated
25613	Exônico	8763	CD164	CD164 antigen, sialomucin	Downregulated
25641	Exônico	1611	DAP	death-associated protein	Downregulated
25655	Exônico	283349	RASSF3	Ras association (RalGDS/AF-6) domain family 3	Downregulated
25681	Exônico	55706	TMEM48	transmembrane protein 48	Downregulated
25693	Exônico	26271	FBXO5	F-box protein 5	Downregulated
25697	Exônico	23358	USP24	ubiquitin specific protease 24	Downregulated
25698	Exônico	23256	SCFD1	sec1 family domain containing 1	Downregulated
25700	Exônico	23054	NCOA6	nuclear receptor coactivator 6	Downregulated
25701	Exônico	22916	NCBP2	nuclear cap binding protein subunit 2, 20kDa	Downregulated
25708	Exônico	10077	PHEMX	pan-hematopoietic expression	Downregulated
25716	Exônico	8634	RTCD1	RNA terminal phosphate cyclase domain 1	Downregulated
25720	Exônico	7353	UFD1L	ubiquitin fusion degradation 1-like	Downregulated
25721	Exônico	7132	TNFRSF1A	tumor necrosis factor receptor superfamily, member 1A	Downregulated
25724	Exônico	6522	SLC4A2	solute carrier family 4, anion exchanger, member 2 (erythrocyte membrane protein band 3-like 1)	Downregulated
25752	Exônico	23671	TMEFF2	transmembrane protein with EGF-like and two follistatin-like domains 2	Downregulated
25758	Exônico	284695	ZNF326	zinc finger protein 326	Downregulated
25763	Exônico	148789	B3GALNT2	UDP-GalNAc:betaGlcNAc beta 1,3-galactosaminyltransferase, polypeptide 2	Downregulated

25775	Exônico	79831	-	-	Downregulated
25776	Exônico	79581	GPR172A	G protein-coupled receptor 172A	Downregulated
25805	Exônico	22856	CHSY1	carbohydrate (chondroitin) synthase 1	Downregulated
25806	Exônico	11176	BAZ2A	bromodomain adjacent to zinc finger domain, 2A	Downregulated
25811	Exônico	10155	TRIM28	tripartite motif-containing 28	Downregulated
25824	Exônico	7227	TRPS1	trichorhinophalangeal syndrome I	Downregulated
25840	Exônico	3608	ILF2	interleukin enhancer binding factor 2, 45kDa	Downregulated
25846	Exônico	1869	E2F1	E2F transcription factor 1	Downregulated
25870	Exônico	92259	MRPS36	mitochondrial ribosomal protein S36	Downregulated
25887	Exônico	55706	TMEM48	transmembrane protein 48	Downregulated
25896	Exônico	50717	WDR42A	WD repeat domain 42A	Downregulated
25901	Exônico	25792	CIZ1	CDKN1A interacting zinc finger protein 1	Downregulated
25907	Exônico	22913	RALY	RNA binding protein (autoantigenic, hnRNP-associated with lethal yellow)	Downregulated
25909	Exônico	11059	WWP1	WW domain containing E3 ubiquitin protein ligase 1	Downregulated
25923	Exônico	8470	-	-	Downregulated
25925	Exônico	7644	ZNF91	zinc finger protein 91 (HPF7, HTF10)	Downregulated
25931	Exônico	6241	RRM2	ribonucleotide reductase M2 polypeptide	Downregulated
25965	Exônico	256126	-	-	Downregulated
25974	Exônico	90957	DHX57	DEAH (Asp-Glu-Ala-Asp/His) box polypeptide 57	Downregulated
25991	Exônico	55656	-	-	Downregulated
26005	Exônico	23649	POLA2	polymerase (DNA directed), alpha 2 (70kD subunit)	Downregulated
26013	Exônico	10985	GCN1L1	GCN1 general control of amino-acid synthesis 1-like 1 (yeast)	Downregulated
26043	Exônico	4289	MKLN1	muskelin 1, intracellular mediator containing kelch motifs	Downregulated
26060	Exônico	91663	MYADM	myeloid-associated differentiation marker	Downregulated
26070	Exônico	166614	DCAMKL2	doublecortin and CaM kinase-like 2	Downregulated
26080	Exônico	83752	-	-	Downregulated
26085	Exônico	66036	MTMR9	myotubularin related protein 9	Downregulated
26087	Exônico	60685	TEX27	testis expressed sequence 27	Downregulated
26089	Exônico	57521	-	-	Downregulated
26101	Exônico	51201	ZDHHC2	zinc finger, DHHC-type containing 2	Downregulated
26102	Exônico	50717	WDR42A	WD repeat domain 42A	Downregulated
26105	Exônico	26271	FBXO5	F-box protein 5	Downregulated
26116	Exônico	10808	HSPH1	heat shock 105kDa/110kDa protein 1	Downregulated
26120	Exônico	10075	HUWE1	HECT, UBA and WWE domain containing 1	Downregulated

26124	Exônico	9510	ADAMTS1	a disintegrin-like and metalloprotease (repolysin type) with thrombospondin type 1 motif, 1	Downregulated
26162	Exônico	31	ACACA	acetyl-Coenzyme A carboxylase alpha	Downregulated
26190	Exônico	64216	TFB2M	transcription factor B2, mitochondrial	Downregulated
26213	Exônico	23313	C22orf9	chromosome 22 open reading frame 9	Downregulated
26221	Exônico	10524	HTATIP	HIV-1 Tat interacting protein, 60kDa	Downregulated
26223	Exônico	10155	TRIM28	tripartite motif-containing 28	Downregulated
26225	Exônico	9885	OSBPL2	oxysterol binding protein-like 2	Downregulated
26240	Exônico	6430	SFRS5	splicing factor, arginine/serine-rich 5	Downregulated
26247	Exônico	4898	NRD1	nardilysin (N-arginine dibasic convertase)	Downregulated
26266	Exônico	25898	RCHY1	ring finger and CHY zinc finger domain containing 1	Downregulated
26305	Exônico	51762	RAB8B	RAB8B, member RAS oncogene family	Downregulated
26306	Exônico	51479	ANKFY1	ankyrin repeat and FYVE domain containing 1	Downregulated
26308	Exônico	50717	WDR42A	WD repeat domain 42A	Downregulated
26309	Exônico	29102	RNASEN	ribonuclease III, nuclear	Downregulated
26317	Exônico	23161	SNX13	sorting nexin 13	Downregulated
26341	Exônico	6721	SREBF2	sterol regulatory element binding transcription factor 2	Downregulated
26357	Exônico	2781	GNAZ	guanine nucleotide binding protein (G protein), alpha z polypeptide	Downregulated
26362	Exônico	1457	CSNK2A1	casein kinase 2, alpha 1 polypeptide	Downregulated
26377	Exônico	256281	NUDT14	nudix (nucleoside diphosphate linked moiety X)-type motif 14	Downregulated
26384	Exônico	116835	HSPA12B	heat shock 70kD protein 12B	Downregulated
26389	Exônico	84181	CHD6	chromodomain helicase DNA binding protein 6	Downregulated
26403	Exônico	55656	-	-	Downregulated
26410	Exônico	51366	-	-	Downregulated
26424	Exônico	11177	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	Downregulated
26434	Exônico	9441	CRSP7	cofactor required for Sp1 transcriptional activation, subunit 7, 70kDa	Downregulated
26440	Exônico	7936	RDBP	RD RNA binding protein	Downregulated
26448	Exônico	5832	ALDH18A1	aldehyde dehydrogenase 18 family, member A1	Downregulated
26449	Exônico	5695	PSMB7	proteasome (prosome, macropain) subunit, beta type, 7	Downregulated
26455	Exônico	4292	MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	Downregulated
26456	Exônico	4072	TACSTD1	tumor-associated calcium signal transducer 1	Downregulated
26459	Exônico	3181	HNRPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	Downregulated

26466	Exônico	1315	COPB	coatamer protein complex, subunit beta	Downregulated
26483	Exônico	152100	-	-	Downregulated
26488	Exônico	92140	-	-	Downregulated
26492	Exônico	83752	-	-	Downregulated
26497	Exônico	66008	ALS2CR3	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 3	Downregulated
26502	Exônico	57175	CORO1B	coronin, actin binding protein, 1B	Downregulated
26509	Exônico	54872	-	-	Downregulated
26517	Exônico	26268	FBXO9	F-box protein 9	Downregulated
26542	Exônico	8065	CUL5	cullin 5	Downregulated
26547	Exônico	6721	SREBF2	sterol regulatory element binding transcription factor 2	Downregulated
26549	Exônico	6241	RRM2	ribonucleotide reductase M2 polypeptide	Downregulated
26558	Exônico	4154	MBNL1	muscleblind-like (Drosophila)	Downregulated
26573	Exônico	215	ABCD1	ATP-binding cassette, sub-family D (ALD), member 1	Downregulated
26575	Exônico	5539	PPY	pancreatic polypeptide	Downregulated
26588	Exônico	137964	-	-	Downregulated
26594	Exônico	84553	C6orf168	chromosome 6 open reading frame 168	Downregulated
26614	Exônico	54453	RIN2	Ras and Rab interactor 2 polymerase (DNA directed), alpha 2 (70kD subunit)	Downregulated
26623	Exônico	23649	POLA2	chromosome 22 open reading frame 9	Downregulated
26625	Exônico	23313	C22orf9	HIV-1 Tat interacting protein, 60kDa	Downregulated
26633	Exônico	10524	HTATIP	splicing factor, arginine/serine-rich 7, 35kDa	Downregulated
26652	Exônico	6432	SFRS7	E2F transcription factor 5, p130-binding	Downregulated
26670	Exônico	1875	E2F5	adaptor-related protein complex 1, sigma 3 subunit	Downregulated
26691	Exônico	130340	AP1S3	membrane associated DNA binding protein	Downregulated
26716	Exônico	54542	MNAB	ankyrin repeat and FYVE domain containing 1	Downregulated
26718	Exônico	51479	ANKFY1	zinc finger protein 84 (HPF2)	Downregulated
26749	Exônico	7637	ZNF84	solute carrier family 4, anion exchanger, member 1 (erythrocyte membrane protein band 3, Diego blood group)	Downregulated
26754	Exônico	6521	SLC4A1	host cell factor C1 (VP16-accessory protein)	Downregulated
26768	Exônico	3054	HCFC1	casein kinase 1, delta	Downregulated
26774	Exônico	1453	CSNK1D	cyclin E1	Downregulated
26776	Exônico	898	CCNE1	lymphocyte antigen 86	Downregulated
26781	Exônico	9450	LY86	cofactor required for Sp1 transcriptional activation, subunit 9, 33kDa	Downregulated
26782	Exônico	9443	CRSP9	tumor protein p53 inducible nuclear protein 1	Downregulated
26797	Exônico	94241	TP53INP1	chromosome 1 open reading frame 91	Downregulated
26813	Exônico	56063	C1orf91		

26863	Exônimo	5297	PIK4CA	phosphatidylinositol 4-kinase, catalytic, alpha polypeptide	Downregulated
26864	Exônimo	5108	PCM1	pericentriolar material 1	Downregulated
26870	Exônimo	3613	IMPA2	inositol(myo)-1(or 4)-monophosphatase 2	Downregulated
26880	Exônimo	816	CAMK2B	calcium/calmodulin-dependent protein kinase (CaM kinase) II beta	Downregulated
26900	Exônimo	92140	-	-	Downregulated
26901	Exônimo	89766	UMODL1	uromodulin-like 1	Downregulated
26902	Exônimo	84790	-	-	Downregulated
26914	Exônimo	57161	PELI2	pellino homolog 2 (Drosophila)	Downregulated
26920	Exônimo	55088	C10orf118	chromosome 10 open reading frame 118	Downregulated
26950	Exônimo	9063	PIAS2	protein inhibitor of activated STAT, 2	Downregulated
26956	Exônimo	7342	UBP1	upstream binding protein 1 (LBP-1a)	Downregulated
26957	Exônimo	7126	TNFAIP1	tumor necrosis factor, alpha induced protein 1 (endothelial)	Downregulated
26978	Exônimo	1998	ELF2	E74-like factor 2 (ets domain transcription factor)	Downregulated
26988	Exônimo				Downregulated
27027	Exônimo	51667	-	-	Downregulated
27041	Exônimo	22859	LPHN1	latrophilin 1	Downregulated
27042	Exônimo	11177	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	Downregulated
27057	Exônimo	8405	SPOP	speckle-type POZ protein	Downregulated
27064	Exônimo	6432	SFRS7	splicing factor, arginine/serine-rich 7, 35kDa	Downregulated
27072	Exônimo	4676	NAP1L4	nucleosome assembly protein 1-like 4	Downregulated
27087	Exônimo	605	BCL7A	B-cell CLL/lymphoma 7A	Downregulated
27110	Exônimo	83746	L3MBTL2	l(3)mbt-like 2 (Drosophila)	Downregulated
27114	Exônimo	79705	LRRK1	leucine-rich repeat kinase 1	Downregulated
27137	Exônimo	25778	RIPK5	receptor interacting protein kinase 5	Downregulated
27142	Exônimo	23053	KIAA0913	KIAA0913	Downregulated
27146	Exônimo	10806	SDCCAG8	serologically defined colon cancer antigen 8	Downregulated
27154	Exônimo	9501	RPH3AL	rabphilin 3A-like (without C2 domains)	Downregulated
27167	Exônimo	6241	RRM2	ribonucleotide reductase M2 polypeptide	Downregulated
27187	Exônimo	1164	CKS2	CDC28 protein kinase regulatory subunit 2	Downregulated
27208	Exônimo	116841	-	-	Downregulated
27211	Exônimo	85007	-	-	Downregulated
27219	Exônimo	64854	USP46	ubiquitin specific protease 46	Downregulated
27224	Exônimo	57002	C7orf36	chromosome 7 open reading frame 36	Downregulated
27225	Exônimo	56063	C1orf91	chromosome 1 open reading frame 91	Downregulated
27226	Exônimo	55784	-	-	Downregulated
27236	Exônimo	29992	PILRA	paired immunoglobulin-like type 2 receptor alpha	Downregulated
27240	Exônimo	25909	-	-	Downregulated
27255	Exônimo	9885	OSBPL2	oxysterol binding protein-like 2	Downregulated
27262	Exônimo	8567	MADD	MAP-kinase activating death domain	Downregulated

27274	Exônico	5520	PPP2R2A	protein phosphatase 2 (formerly 2A), regulatory subunit B (PR 52), alpha isoform	Downregulated
27280	Exônico	4074	M6PR	mannose-6-phosphate receptor (cation dependent)	Downregulated
27301	Exônico	399665	C9orf132	chromosome 9 open reading frame 132	Downregulated
27319	Exônico	79932	KIAA0319L	KIAA0319-like	Downregulated
27322	Exônico	64478	CSMD1	CUB and Sushi multiple domains 1	Downregulated
27347	Exônico	23160	WDR43	WD repeat domain 43	Downregulated
27363	Exônico	8834	C17orf35	chromosome 17 open reading frame 35	Downregulated
27389	Exônico	2146	EZH2	enhancer of zeste homolog 2 (Drosophila)	Downregulated
27393	Exônico	1164	CKS2	CDC28 protein kinase regulatory subunit 2	Downregulated
27415	Exônico	96764	NCOA6IP	nuclear receptor coactivator 6 interacting protein	Downregulated
27439	Exônico	51668	C1orf41	chromosome 1 open reading frame 41	Downregulated
27463	Exônico	9617	MTRF1	mitochondrial translational release factor 1	Downregulated
27484	Exônico	4678	NASP	nuclear autoantigenic sperm protein (histone-binding)	Downregulated
27514	Exônico	143384	C10orf46	chromosome 10 open reading frame 46	Downregulated
27520	Exônico	84779	-	-	Downregulated
27537	Exônico	55254	TMEM39A	transmembrane protein 39A	Downregulated
27538	Exônico	55088	C10orf118	chromosome 10 open reading frame 118	Downregulated
27545	Exônico	29102	RNASEN	ribonuclease III, nuclear	Downregulated
27546	Exônico	27146	-	-	Downregulated
27549	Exônico	25778	RIPK5	receptor interacting protein kinase 5	Downregulated
27551	Exônico	23357	KIAA0759	KIAA0759	Downregulated
27555	Exônico	22913	RALY	RNA binding protein (autoantigenic, hnRNP-associated with lethal yellow)	Downregulated
27563	Exônico	9927	MFN2	mitofusin 2	Downregulated
27573	Exônico	7633	ZNF79	zinc finger protein 79 (pT7)	Downregulated
27574	Exônico	7337	UBE3A	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	Downregulated
27600	Exônico	892	CCNC	cyclin C	Downregulated
27636	Exônico	57002	C7orf36	chromosome 7 open reading frame 36	Downregulated
27646	Exônico	51366	-	-	Downregulated
27652	Exônico	25911	-	-	Downregulated
27656	Exônico	23230	VPS13A	vacuolar protein sorting 13A (yeast)	Downregulated
27661	Exônico	10988	METAP2	methionyl aminopeptidase 2	Downregulated
27669	Exônico	9618	TRAF4	TNF receptor-associated factor 4	Downregulated
27670	Exônico	9442	CRSP8	cofactor required for Sp1 transcriptional activation, subunit 8, 34kDa	Downregulated
27677	Exônico	7468	WHSC1	Wolf-Hirschhorn syndrome candidate 1	Downregulated

27682	Exônico	6433	SFRS8	splicing factor, arginine/serine-rich 8 (suppressor-of-white-apricot homolog, Drosophila)	Downregulated
27687	Exônico	5298	PIK4CB	phosphatidylinositol 4-kinase, catalytic, beta polypeptide	Downregulated
27691	Exônico	4292	MLH1	mutL homolog 1, colon cancer, nonpolyposis type 2 (E. coli)	Downregulated
27694	Exônico	3613	IMPA2	inositol(myo)-1(or 4)-monophosphatase 2	Downregulated
27714	Exônico	286354	C9orf130	chromosome 9 open reading frame 130	Downregulated
27749	Exônico	51199	NIN	ninein (GSK3B interacting protein)	Downregulated
27759	Exônico	23160	WDR43	WD repeat domain 43	Downregulated
27766	Exônico	10425	ARIH2	ariadne homolog 2 (Drosophila)	Downregulated
27780	Exônico	7337	UBE3A	ubiquitin protein ligase E3A (human papilloma virus E6-associated protein, Angelman syndrome)	Downregulated
27791	Exônico	5018	OXA1L	oxidase (cytochrome c) assembly 1-like	Downregulated
27795	Exônico	3915	LAMC1	laminin, gamma 1 (formerly LAMB2)	Downregulated
27829	Exônico	85007	-	-	Downregulated
27835	Exônico	79832	-	-	Downregulated
27847	Exônico	55193	-	-	Downregulated
27851	Exônico	51668	C1orf41	chromosome 1 open reading frame 41	Downregulated
27857	Exônico	26115	-	-	Downregulated
27859	Exônico	23657	SLC7A11	solute carrier family 7, (cationic amino acid transporter, y+ system) member 11	Downregulated
27860	Exônico	23451	SF3B1	splicing factor 3b, subunit 1 155kDa	Downregulated
27874	Exônico	9738	-	-	Downregulated
27883	Exônico	7468	WHSC1	Wolf-Hirschhorn syndrome candidate 1	Downregulated
27901	Exônico	3181	HNRPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	Downregulated
27922	Exônico	222256	-	-	Downregulated
27925	Exônico	152006	RNF38	ring finger protein 38	Downregulated
27930	Exônico	92106	-	-	Downregulated
27934	Exônico	83737	ITCH	itchy homolog E3 ubiquitin protein ligase (mouse)	Downregulated
27935	Exônico	81502	HM13	histocompatibility (minor) 13	Downregulated
27949	Exônico	55254	TMEM39A	transmembrane protein 39A	Downregulated
27969	Exônico	11057	ABHD2	abhydrolase domain containing 2	Downregulated
27981	Exônico	8833	GMPS	guanine monphosphate synthetase	Downregulated
27987	Exônico	7113	TMPRSS2	transmembrane protease, serine 2	Downregulated
27988	Exônico	6926	TBX3	T-box 3 (ulnar mammary syndrome)	Downregulated
27991	Exônico	6240	RRM1	ribonucleotide reductase M1 polypeptide	Downregulated
27992	Exônico	5908	RAP1B	RAP1B, member of RAS oncogene family	Downregulated
28007	Exônico	2140	EYA3	eyes absent homolog 3 (Drosophila)	Downregulated

28017	Exônico				Downregulated
28025	Exônico	256435	ST6GALNAC3	ST6 (alpha-N-acetyl-neuraminy-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3	Downregulated
28026	Exônico	221037	JMJD1C	jumonji domain containing 1C	Downregulated
28028	Exônico	158866	ZDHC15	zinc finger, DHHC-type containing 15	Downregulated
28029	Exônico	148867	SLC30A7	solute carrier family 30 (zinc transporter), member 7	Downregulated
28047	Exônico	57466	SFRS15	splicing factor, arginine/serine-rich 15	Downregulated
28058	Exônico	51367	POP5	processing of precursor 5, ribonuclease P/MRP subunit (S. cerevisiae)	Downregulated
28064	Exônico	25911	-	-	Downregulated
28067	Exônico	23315	SLC9A8	solute carrier family 9 (sodium/hydrogen exchanger), isoform 8	Downregulated
28089	Exônico	7468	WHSC1	Wolf-Hirschhorn syndrome candidate 1	Downregulated
28097	Exônico	5700	PSMC1	proteasome (prosome, macropain) 26S subunit, ATPase, 1	Downregulated
28106	Exônico	3615	IMPDH2	IMP (inosine monophosphate) dehydrogenase 2	Downregulated
28132	Exônico	143279	HECTD2	HECT domain containing 2	Downregulated
28143	Exônico	79932	KIAA0319L	KIAA0319-like	Downregulated
28157	Exônico	54862	-	-	Downregulated
28159	Exônico	51755	-	-	Downregulated
28161	Exônico	51199	NIN	ninein (GSK3B interacting protein)	Downregulated
28171	Exônico	23158	-	-	Downregulated
28177	Exônico	10579	TACC2	transforming, acidic coiled-coil containing protein 2	Downregulated
28203	Exônico	5018	OXA1L	oxidase (cytochrome c) assembly 1-like	Downregulated
28216	Exônico	1452	CSNK1A1	casein kinase 1, alpha 1	Downregulated
28237	Exônico	126308	MOBK2A	MOB1, Mps One Binder kinase activator-like 2A (yeast)	Downregulated
28240	Exônico	91039	DPP9	dipeptidylpeptidase 9	Downregulated
28250	Exônico	64225	ARL6IP2	ADP-ribosylation factor-like 6 interacting protein 2	Downregulated
28265	Exônico	51106	TFB1M	transcription factor B1, mitochondrial	Downregulated
28278	Exônico	11177	BAZ1A	bromodomain adjacent to zinc finger domain, 1A	Downregulated
28295	Exônico	7468	WHSC1	Wolf-Hirschhorn syndrome candidate 1	Downregulated
28300	Exônico	6433	SFRS8	splicing factor, arginine/serine-rich 8 (suppressor-of-white-apricot homolog, Drosophila)	Downregulated
28308	Exônico	4678	NASP	nuclear autoantigenic sperm protein (histone-binding)	Downregulated
28311	Exônico	3832	KIF11	kinesin family member 11	Downregulated
28331	Exônico	399512	SLC25A35	solute carrier family 25, member 35	Downregulated

28333	Exônico	283219	-	-	Downregulated
28352	Exônico	64431	ACTR6	ARP6 actin-related protein 6 homolog (yeast)	Downregulated
28373	Exônico	25776	PGEA1	PKD2 interactor, golgi and endoplasmic reticulum associated 1	Downregulated
28376	Exônico	23256	SCFD1	sec1 family domain containing 1	Downregulated
28382	Exônico	10806	SDCCAG8	serologically defined colon cancer antigen 8	Downregulated
28385	Exônico	10213	PSMD14	proteasome (prosome, macropain) 26S subunit, non-ATPase, 14	Downregulated
28388	Exônico	9787	DLG7	discs, large homolog 7 (Drosophila)	Downregulated
28422	Exônico	1452	CSNK1A1	casein kinase 1, alpha 1	Downregulated
28424	Exônico	892	CCNC	cyclin C	Downregulated
28437	Exônico	256471	-	-	Downregulated
28448	Exônico	84612	PARD6B	par-6 partitioning defective 6 homolog beta (C. elegans)	Downregulated
28462	Exônico	55788	C6orf209	chromosome 6 open reading frame 209	Downregulated
28463	Exônico	55660	FNBP3	formin binding protein 3	Downregulated
28468	Exônico	54454	-	-	Downregulated
28496	Exônico	8493	PPM1D	protein phosphatase 1D magnesium-dependent, delta isoform	Downregulated
28497	Exônico	8774	NAPG	N-ethylmaleimide-sensitive factor attachment protein, gamma	Downregulated
28509	Exônico	5700	PSMC1	proteasome (prosome, macropain) 26S subunit, ATPase, 1	Downregulated
28516	Exônico	4076	M11S1	membrane component, chromosome 11, surface marker 1	Downregulated
28519	Exônico	3181	HNRPA2B1	heterogeneous nuclear ribonucleoprotein A2/B1	Downregulated
28525	Exônico	1629	DBT	dihydrolipoamide branched chain transacylase E2	Downregulated
28543	Exônico	152002	C3orf21	chromosome 3 open reading frame 21	Downregulated
28575	Exônico	29097	-	-	Downregulated
28584	Exônico	23049	-	-	Downregulated
28586	Exônico	11252	PACSN2	protein kinase C and casein kinase substrate in neurons 2	Downregulated
28594	Exônico	9786	KIAA0586	KIAA0586	Downregulated
28605	Exônico	7110	TMF1	TATA element modulatory factor 1	Downregulated
28606	Exônico	6926	TBX3	T-box 3 (ulnar mammary syndrome)	Downregulated
28610	Exônico	5906	RAP1A	RAP1A, member of RAS oncogene family	Downregulated
28631	Exônico	672	BRCA1	breast cancer 1, early onset	Downregulated
28641	Exônico	375260	-	-	Downregulated
28649	Exônico	126321	C19orf28	chromosome 19 open reading frame 28	Downregulated
28651	Exônico	112399	EGLN3	egl nine homolog 3 (C. elegans)	Downregulated
28660	Exônico	79596	C13orf7	chromosome 13 open reading frame 7	Downregulated
28666	Exônico	57018	CCNL1	cyclin L1	Downregulated

28676	Exônimo	51367	POP5	processing of precursor 5, ribonuclease P/MRP subunit (<i>S. cerevisiae</i>)	Downregulated
28701	Exônimo	9202	ZNF262	zinc finger protein 262	Downregulated
28724	Exônimo	3615	IMPDH2	IMP (inosine monophosphate) dehydrogenase 2	Downregulated
28735	Exônimo	613	BCR	breakpoint cluster region	Downregulated
28751	Exônimo	130074	-	-	Downregulated
28758	Exônimo	83737	ITCH	itchy homolog E3 ubiquitin protein ligase (mouse)	Downregulated
28761	Exônimo	79932	KIAA0319L	KIAA0319-like	Downregulated
28771	Exônimo	55704	KIAA1212	KIAA1212	Downregulated
28785	Exônimo	25771	TBC1D22A	TBC1 domain family, member 22A	Downregulated
28810	Exônimo	7332	UBE2L3	ubiquitin-conjugating enzyme E2L 3	Downregulated
28833	Exônimo	1739	DLG1	discs, large homolog 1 (<i>Drosophila</i>)	Downregulated
28841	Exônimo				Downregulated
28849	Exônimo	256471	-	-	Downregulated
28869	Exônimo	59343	SEN2	SUMO1/sentrin/SMT3 specific protease 2	Downregulated
28873	Exônimo	56171	DNAH7	dynein, axonemal, heavy polypeptide 7	Downregulated
28875	Exônimo	55661	DDX27	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	Downregulated
28883	Exônimo	51108	-	-	Downregulated
28921	Exônimo	5704	PSMC4	proteasome (prosome, macropain) 26S subunit, ATPase, 4	Downregulated
28931	Exônimo	3182	HNRPAB	heterogeneous nuclear ribonucleoprotein A/B	Downregulated
28961	Exônimo	89122	TRIM4	tripartite motif-containing 4	Downregulated
28972	Exônimo	57704	GBA2	glucosidase, beta (bile acid) 2	Downregulated
28975	Exônimo	56478	EIF4ENIF1	eukaryotic translation initiation factor 4E nuclear import factor 1	Downregulated
28976	Exônimo	55835	CENPJ	centromere protein J	Downregulated
28984	Exônimo	51477	-	-	Downregulated
28987	Exônimo	29097	-	-	Downregulated
29011	Exônimo	8833	GMPS	guanine monophosphate synthetase	Downregulated
29016	Exônimo	7332	UBE2L3	ubiquitin-conjugating enzyme E2L 3	Downregulated
29017	Exônimo	7110	TMF1	TATA element modulatory factor 1	Downregulated
29039	Exônimo	1739	DLG1	discs, large homolog 1 (<i>Drosophila</i>)	Downregulated
29042	Exônimo	892	CCNC	cyclin C	Downregulated
29043	Exônimo	672	BRCA1	breast cancer 1, early onset	Downregulated
29060	Exônimo	137994	LETM2	leucine zipper-EF-hand containing transmembrane protein 2	Downregulated
29063	Exônimo	112495	C6orf51	chromosome 6 open reading frame 51	Downregulated
29083	Exônimo	55196	-	-	Downregulated
29088	Exônimo	51371	C13orf12	chromosome 13 open reading frame 12	Downregulated
29120	Exônimo	7264	TSTA3	tissue specific transplantation antigen P35B	Downregulated
29134	Exônimo	4076	M11S1	membrane component, chromosome 11, surface marker 1	Downregulated

29135	Exônico	3833	KIFC1	kinesin family member C1	Downregulated
29136	Exônico	3615	IMPDH2	IMP (inosine monophosphate) dehydrogenase 2	Downregulated
29149	Exônico	132	ADK	adenosine kinase polymerase (DNA-directed), delta interacting protein 3	Downregulated
29169	Exônico	84271	POLDIP3	ARP6 actin-related protein 6 homolog (yeast)	Downregulated
29176	Exônico	64431	ACTR6	polymerase (RNA) III (DNA directed) polypeptide B	Downregulated
29183	Exônico	55703	POLR3B	radical S-adenosyl methionine and flavodoxin domains 1	Downregulated
29185	Exônico	55253	RSAFD1	-	Downregulated
29198	Exônico	23514	-	-	Downregulated
29211	Exônico	9922	IQSEC1	IQ motif and Sec7 domain 1	Downregulated
29226	Exônico	6513	SLC2A1	solute carrier family 2 (facilitated glucose transporter), member 1	Downregulated
29227	Exônico	6240	RRM1	ribonucleotide reductase M1 polypeptide	Downregulated
29228	Exônico	5903	RANBP2	RAN binding protein 2	Downregulated
29233	Exônico	5002	SLC22A18	solute carrier family 22 (organic cation transporter), member 18	Downregulated
29268	Exônico	116985	CENTD2	centaurin, delta 2	Downregulated
29274	Exônico	83549	UCK1	uridine-cytidine kinase 1	Downregulated
29300	Exônico	25920	-	-	Downregulated
29336	Exônico	5108	PCM1	pericentriolar material 1	Downregulated
29384	Exônico	57700	KIAA1600	KIAA1600	Downregulated
29399	Exônico	29087	-	-	Downregulated
29400	Exônico	27131	SNX5	sorting nexin 5	Downregulated
29412	Exônico	10802	SEC24A	SEC24 related gene family, member A (S. cerevisiae)	Downregulated
29438	Exônico	5162	PDHB	pyruvate dehydrogenase (lipoamide) beta	Downregulated
29454	Exônico	891	CCNB1	cyclin B1	Downregulated
29459	Exônico	118881	COMTD1	catechol-O-methyltransferase domain containing 1	Downregulated
29466	Exônico	284996	RNF149	ring finger protein 149	Downregulated
29473	Exônico	126321	C19orf28	chromosome 19 open reading frame 28	Downregulated
29493	Exônico	55661	DDX27	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	Downregulated
29503	Exônico	28957	MRPS28	mitochondrial ribosomal protein S28	Downregulated
29507	Exônico	23670	TMEM2	transmembrane protein 2	Downregulated
29513	Exônico	22862	FNDC3A	fibronectin type III domain containing 3A	Downregulated
29514	Exônico	11180	WDR6	WD repeat domain 6	Downregulated
29528	Exônico	8569	MKNK1	MAP kinase interacting serine/threonine kinase 1	Downregulated
29539	Exônico	5704	PSMC4	proteasome (prosome, macropain) 26S subunit, ATPase, 4	Downregulated
29549	Exônico	3182	HNRPAB	heterogeneous nuclear ribonucleoprotein A/B	Downregulated
29579	Exônico	88745	C6orf153	chromosome 6 open reading frame 153	Downregulated
29584	Exônico	80204	FBXO11	F-box protein 11	Downregulated
29588	Exônico	64430	C14orf135	chromosome 14 open reading frame 135	Downregulated

29591	Exônico	57519	STARD9	START domain containing 9	Downregulated
29594	Exônico	55835	CENPJ	centromere protein J	Downregulated
29611	Exônico	23352	-	-	Downregulated
29645	Exônico	5001	ORC5L	origin recognition complex, subunit 5-like (yeast)	Downregulated
29660	Exônico	890	CCNA2	cyclin A2	Downregulated
29675	Exônico	196515	-	-	Downregulated
29681	Exônico	112495	C6orf51	chromosome 6 open reading frame 51	Downregulated
29682	Exônico	91057	-	-	Downregulated
29699	Exônico	55661	DDX27	DEAD (Asp-Glu-Ala-Asp) box polypeptide 27	Downregulated
29712	Exônico	25921	ZDHHC5	zinc finger, DHHC-type containing 5	Downregulated
29714	Exônico	23456	ABCB10	ATP-binding cassette, subfamily B (MDR/TAP), member 10	Downregulated
29716	Exônico	23232	TBC1D12	TBC1 domain family, member 12	Downregulated
29727	Exônico	9887	C1orf16	chromosome 1 open reading frame 16	Downregulated
29737	Exônico	7486	WRN	Werner syndrome	Downregulated
29739	Exônico	7072	TIA1	TIA1 cytotoxic granule-associated RNA binding protein	Downregulated
29750	Exônico	4682	NUBP1	nucleotide binding protein 1 (MinD homolog, E. coli)	Downregulated
29767	Exônico	136	ADORA2B	adenosine A2b receptor	Downregulated
29776	Exônico	222223	KIAA1324L	KIAA1324-like	Downregulated
29783	Exônico	114791	TUBGCP5	tubulin, gamma complex associated protein 5	Downregulated
29784	Exônico	92017	-	-	Downregulated
29788	Exônico	83734	APG10L	APG10 autophagy 10-like (S. cerevisiae)	Downregulated
29795	Exônico	60684	-	-	Downregulated
29799	Exônico	56478	EIF4ENIF1	eukaryotic translation initiation factor 4E nuclear import factor 1	Downregulated
29803	Exônico	55252	ASXL2	additional sex combs like 2 (Drosophila)	Downregulated
29811	Exônico	29087	-	-	Downregulated
29812	Exônico	27131	SNX5	sorting nexin 5	Downregulated
29816	Exônico	23514	-	-	Downregulated
29824	Exônico	10802	SEC24A	SEC24 related gene family, member A (S. cerevisiae)	Downregulated
29830	Exônico	9785	DHX38	DEAH (Asp-Glu-Ala-His) box polypeptide 38	Downregulated
29847	Exônico	5747	PTK2	PTK2 protein tyrosine kinase 2	Downregulated
29850	Exônico	5162	PDHB	pyruvate dehydrogenase (lipoamide) beta	Downregulated
29851	Exônico	5001	ORC5L	origin recognition complex, subunit 5-like (yeast)	Downregulated
29857	Exônico	3337	DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1	Downregulated
29883	Exônico	149041	-	-	Downregulated
29889	Exônico	85019	C18orf45	chromosome 18 open reading frame 45	Downregulated
29900	Exônico	57623	ZNF406	zinc finger protein 406	Downregulated
29912	Exônico	51371	C13orf12	chromosome 13 open reading frame 12	Downregulated
29917	Exônico	26122	EPC2	enhancer of polycomb homolog 2 (Drosophila)	Downregulated

29951	Exônico	5707	PSMD1	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1	Downregulated
29967	Exônico	1633	DCK	deoxycytidine kinase	Downregulated
29974	Exônico	11072	DUSP14	dual specificity phosphatase 14	Downregulated
29984	Exônico	165545	DQX1	DEAQ box polypeptide 1 (RNA-dependent ATPase)	Downregulated
29991	Exônico	88745	C6orf153	chromosome 6 open reading frame 153	Downregulated
29995	Exônico	81490	PTDSS2	phosphatidylserine synthase 2	Downregulated
30001	Exônico	60684	-	-	Downregulated
30002	Exônico	57700	KIAA1600	KIAA1600	Downregulated
30009	Exônico	55252	ASXL2	additional sex combs like 2 (Drosophila)	Downregulated
30011	Exônico	54859	-	-	Downregulated
30016	Exônico	50640	-	-	Downregulated
30087	Exônico	196515	-	-	Downregulated
30102	Exônico	79607	-	-	Downregulated
30110	Exônico	55788	C6orf209	chromosome 6 open reading frame 209	Downregulated
30127	Exônico	23318	ZCCHC11	zinc finger, CCHC domain containing 11	Downregulated
30128	Exônico	23232	TBC1D12	TBC1 domain family, member 12	Downregulated
30150	Exônico	7265	TTC1	tetratricopeptide repeat domain 1	Downregulated
30162	Exônico	4682	NUBP1	nucleotide binding protein 1 (MinD homolog, E. coli)	Downregulated
30164	Exônico	4076	M11S1	membrane component, chromosome 11, surface marker 1	Downregulated
30167	Exônico	3182	HNRPAB	heterogeneous nuclear ribonucleoprotein A/B	Downregulated
30180	Exônico				Downregulated
30189	Exônico	201176	ARHGAP27	Rho GTPase activating protein 27	Downregulated
30203	Exônico	79915	C17orf41	chromosome 17 open reading frame 41	Downregulated
30219	Exônico	51742	ARID4B	AT rich interactive domain 4B (RBP1- like)	Downregulated
30233	Exônico	22909	-	-	Downregulated
30244	Exônico	9497	SLC4A7	solute carrier family 4, sodium bicarbonate cotransporter, member 7	Downregulated
30258	Exônico	5898	RALA	v-ral simian leukemia viral oncogene homolog A (ras related)	Downregulated
30282	Exônico	25	ABL1	v-abl Abelson murine leukemia viral oncogene homolog 1	Downregulated
30283	Exônico				Downregulated
30289	Exônico	375484	-	-	Downregulated
30301	Exônico	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
30307	Exônico	79846	-	-	Downregulated
30308	Exônico	79607	-	-	Downregulated
30317	Exônico	55664	CDC37L1	cell division cycle 37 homolog (S. cerevisiae)-like 1	Downregulated
30320	Exônico	55027	-	-	Downregulated
30329	Exônico	26123	C10orf61	chromosome 10 open reading frame 61	Downregulated
30330	Exônico	25921	ZDHC5	zinc finger, DHHC-type containing 5	Downregulated
30333	Exônico	23325	KIAA1033	KIAA1033	Downregulated

30339	Exônico	10994	ILVBL	ilvB (bacterial acetolactate synthase)-like	Downregulated
30342	Exônico	10390	CEPT1	choline/ethanolamine phosphotransferase 1	Downregulated
30359	Exônico	6651	SON	SON DNA binding protein	Downregulated
30408	Exônico	80201	HKDC1	hexokinase domain containing 1	Downregulated
30411	Exônico	65986	ZBTB10	zinc finger and BTB domain containing 10	Downregulated
30412	Exônico	64425	-	-	Downregulated
30423	Exônico	54859	-	-	Downregulated
30424	Exônico	54535	CCHCR1	coiled-coil alpha-helical rod protein 1	Downregulated
30432	Exônico	25972	UNC50	unc-50 homolog (C. elegans)	Downregulated
30438	Exônico	23048	FNBP1	formin binding protein 1	Downregulated
30447	Exônico	9919	KIAA0310	KIAA0310	Downregulated
30458	Exônico	7329	UBE2I	ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)	Downregulated
30470	Exônico	4751	NEK2	NIMA (never in mitosis gene a)-related kinase 2	Downregulated
30475	Exônico	3336	HSPE1	heat shock 10kDa protein 1 (chaperonin 10)	Downregulated
30483	Exônico	1152	CKB	creatine kinase, brain	Downregulated
30512	Exônico	80045	GPR157	G protein-coupled receptor 157	Downregulated
30517	Exônico	59349	KLHL12	kelch-like 12 (Drosophila)	Downregulated
30532	Exônico	29995	LMCD1	LIM and cysteine-rich domains 1	Downregulated
30546	Exônico	10730	YME1L1	YME1-like 1 (S. cerevisiae)	Downregulated
30551	Exônico	9889	ZBED4	zinc finger, BED-type containing 4	Downregulated
30576	Exônico	4076	M11S1	membrane component, chromosome 11, surface marker 1	Downregulated
30598	Exônico	286053	-	-	Downregulated
30603	Exônico	151903	CCDC12	coiled-coil domain containing 12	Downregulated
30634	Exônico	50618	ITSN2	intersectin 2	Downregulated
30645	Exônico	22907	DHX30	DEAH (Asp-Glu-Ala-His) box polypeptide 30	Downregulated
30661	Exônico	8464	SUPT3H	suppressor of Ty 3 homolog (S. cerevisiae)	Downregulated
30672	Exônico	5573	PRKAR1A	protein kinase, cAMP-dependent, regulatory, type I, alpha (tissue specific extinguisher 1)	Downregulated
30681	Exônico	3329	HSPD1	heat shock 60kDa protein 1 (chaperonin)	Downregulated
30704	Exônico	221037	JMJD1C	jumonji domain containing 1C	Downregulated
30718	Exônico	80045	GPR157	G protein-coupled receptor 157	Downregulated
30730	Exônico	55374	-	-	Downregulated
30751	Exônico	10994	ILVBL	ilvB (bacterial acetolactate synthase)-like	Downregulated
30756	Exônico	10010	TANK	TRAF family member-associated NFKB activator	Downregulated
30766	Exônico	7975	MAFK	v-maf musculoaponeurotic fibrosarcoma oncogene homolog K (avian)	Downregulated

30768	Exônico	7266	DNAJC7	DnaJ (Hsp40) homolog, subfamily C, member 7	Downregulated
30772	Exônico	6433	SFRS8	splicing factor, arginine/serine-rich 8 (suppressor-of-white-apricot homolog, Drosophila)	Downregulated
30789	Exônico	2101	ESRRA	estrogen-related receptor alpha	Downregulated
30792	Exônico	1353	COX11	COX11 homolog, cytochrome c oxidase assembly protein (yeast)	Downregulated
30798	Exônico	84272	-	-	Downregulated
30808	Exônico	165055	-	-	Downregulated
30814	Exônico	92002	FAM58A	family with sequence similarity 58, member A	Downregulated
30816	Exônico	84750	FUT10	fucosyltransferase 10 (alpha (1,3) fucosyltransferase)	Downregulated
30817	Exônico	84269	CHCHD5	coiled-coil-helix-coiled-coil-helix domain containing 5	Downregulated
30820	Exônico	80199	-	-	Downregulated
30829	Exônico	56474	CTPS2	CTP synthase II	Downregulated
30835	Exônico	54853	WDR55	WD repeat domain 55	Downregulated
30836	Exônico	54535	CCHCR1	coiled-coil alpha-helical rod protein 1	Downregulated
30844	Exônico	25966	C21orf25	chromosome 21 open reading frame 25	Downregulated
30851	Exônico	22907	DHX30	DEAH (Asp-Glu-Ala-His) box polypeptide 30	Downregulated
30860	Exônico	9785	DHX38	DEAH (Asp-Glu-Ala-His) box polypeptide 38	Downregulated
30870	Exônico	7329	UBE2I	ubiquitin-conjugating enzyme E2I (UBC9 homolog, yeast)	Downregulated
30881	Exônico	4999	ORC2L	origin recognition complex, subunit 2-like (yeast)	Downregulated
30887	Exônico	3321	IGSF3	immunoglobulin superfamily, member 3	Downregulated
30901	Exônico	162966	ZNF600	zinc finger protein 600	Downregulated
30907	Exônico	375743	PTAR1	protein prenyltransferase alpha subunit repeat containing 1	Downregulated
30919	Exônico	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
30932	Exônico	57020	-	-	Downregulated
30935	Exônico	55665	-	-	Downregulated
30947	Exônico	26123	C10orf61	chromosome 10 open reading frame 61	Downregulated
30986	Exônico	4683	NBN	nibrin	Downregulated
30989	Exônico	3837	KPNB1	karyopherin (importin) beta 1	Downregulated
31013	Exônico	201163	FLCN	folliculin	Downregulated
31019	Exônico	114791	TUBGCP5	tubulin, gamma complex associated protein 5	Downregulated
31041	Exônico	54853	WDR55	WD repeat domain 55	Downregulated
31089	Exônico	4439	MSH5	mutS homolog 5 (E. coli)	Downregulated
31099	Exônico	1737	DLAT	dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydrogenase complex)	Downregulated
31107	Exônico				Downregulated
31147	Exônico	51686	OAZ3	ornithine decarboxylase antizyme 3	Downregulated
31173	Exônico	9208	LRRFIP1	leucine rich repeat (in FLII) interacting protein 1	Downregulated

31181	Exônico	7072	TIA1	TIA1 cytotoxic granule-associated RNA binding protein	Downregulated
31188	Exônico	5523	PPP2R3A	protein phosphatase 2 (formerly 2A), regulatory subunit B", alpha	Downregulated
31230	Exônico	83732	RIOK1	RIO kinase 1 (yeast)	Downregulated
31242	Exônico	55833	UBAP2	ubiquitin associated protein 2	Downregulated
31248	Exônico	54532	USP53	ubiquitin specific protease 53	Downregulated
31263	Exônico	22907	DHX30	DEAH (Asp-Glu-Ala-His) box polypeptide 30	Downregulated
31271	Exônico	9918	-	-	Downregulated
31281	Exônico	7626	ZNF75	zinc finger protein 75 (D8C6)	Downregulated
31282	Exônico	7328	UBE2H	ubiquitin-conjugating enzyme E2H (UBC8 homolog, yeast)	Downregulated
31288	Exônico	5894	RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	Downregulated
31300	Exônico	3015	H2AFZ	H2A histone family, member Z	Downregulated
31305	Exônico	1737	DLAT	dihydrolipoamide S-acetyltransferase (E2 component of pyruvate dehydrogenase complex)	Downregulated
31325	Exônico	149076	-	-	Downregulated
31326	Exônico	138046	-	-	Downregulated
31328	Exônico	116985	CENTD2	centaurin, delta 2	Downregulated
31329	Exônico	112597	-	-	Downregulated
31341	Exônico	59349	KLHL12	kelch-like 12 (Drosophila)	Downregulated
31356	Exônico	29997	GLTSCR2	glioma tumor suppressor candidate region gene 2	Downregulated
31393	Exônico	5707	PSMD1	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1	Downregulated
31398	Exônico	4683	NBN	nibrin	Downregulated
31402	Exônico	3632	INPP5A	inositol polyphosphate-5-phosphatase, 40kDa	Downregulated
31435	Exônico	84268	-	-	Downregulated
31438	Exônico	80199	-	-	Downregulated
31442	Exônico	64422	APG3L	APG3 autophagy 3-like (S. cerevisiae)	Downregulated
31444	Exônico	57696	DDX55	DEAD (Asp-Glu-Ala-Asp) box polypeptide 55	Downregulated
31449	Exônico	55699	IARS2	isoleucine-tRNA synthetase 2, mitochondrial	Downregulated
31451	Exônico	55250	STATIP1	signal transducer and activator of transcription 3 interacting protein 1	Downregulated
31453	Exônico	54853	WDR55	WD repeat domain 55	Downregulated
31459	Exônico	29081	METTL5	methyltransferase like 5 methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 2,	Downregulated
31472	Exônico	10797	MTHFD2	methenyltetrahydrofolate cyclohydrolase	Downregulated
31482	Exônico	9057	SLC7A6	solute carrier family 7 (cationic amino acid transporter, y+ system), member 6	Downregulated
31502	Exônico	4152	MBD1	methyl-CpG binding domain protein 1	Downregulated
31508	Exônico	2335	FN1	fibronectin 1	Downregulated
31528	Exônico	221061	C10orf38	chromosome 10 open reading frame 38	Downregulated

31537	Exônico	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
31540	Exônico	83606	C22orf13	chromosome 22 open reading frame 13	Downregulated
31541	Exônico	80725	-	-	Downregulated
31552	Exônico	55789	DEPDC1B	DEP domain containing 1B	Downregulated
31560	Exônico	51375	SNX7	sorting nexin 7	Downregulated
31568	Exônico	23460	ABCA6	ATP-binding cassette, sub-family A (ABC1), member 6	Downregulated
31571	Exônico	23112	TNRC6B	trinucleotide repeat containing 6B	Downregulated
31598	Exônico	5859	QARS	glutaminyl-tRNA synthetase	Downregulated
31607	Exônico	3837	KPNB1	karyopherin (importin) beta 1	Downregulated
31610	Exônico	2932	GSK3B	glycogen synthase kinase 3 beta	Downregulated
31647	Exônico	65983	-	-	Downregulated
31674	Exônico	23045	KIAA0826	KIAA0826	Downregulated
31682	Exônico	10059	DNM1L	dynammin 1-like	Downregulated
31695	Exônico	7095	TLOC1	translocation protein 1	Downregulated
31700	Exônico	5894	RAF1	v-raf-1 murine leukemia viral oncogene homolog 1	Downregulated
31711	Exônico	3313	HSPA9B	heat shock 70kDa protein 9B (mortalin-2)	Downregulated
31718	Exônico	1445	CSK	c-src tyrosine kinase	Downregulated
31725	Exônico				Downregulated
31737	Exônico	149095	DCST1	DC-STAMP domain containing 1	Downregulated
31749	Exônico	79848	-	-	Downregulated
31757	Exônico	56172	ANKH	ankylosis, progressive homolog (mouse)	Downregulated
31762	Exônico	55028	-	-	Downregulated
31766	Exônico	51377	UCHL5	ubiquitin carboxyl-terminal hydrolase L5	Downregulated
31771	Exônico	26123	C10orf61	chromosome 10 open reading frame 61	Downregulated
31778	Exônico	23028	AOF2	amine oxidase (flavin containing) domain 2	Downregulated
31781	Exônico	11004	KIF2C	kinesin family member 2C	Downregulated
31797	Exônico	7494	XBP1	X-box binding protein 1	Downregulated
31799	Exônico	7073	TIAL1	TIA1 cytotoxic granule-associated RNA binding protein-like 1	Downregulated
31802	Exônico	6434	SFRS10	splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)	Downregulated
31812	Exônico	4085	MAD2L1	MAD2 mitotic arrest deficient-like 1 (yeast)	Downregulated
31821	Exônico	1635	DCTD	dCMP deaminase	Downregulated
31836	Exônico	221955	-	-	Downregulated
31850	Exônico	80196	RNF34	ring finger protein 34	Downregulated
31855	Exônico	60678	EEFSEC	eukaryotic elongation factor, selenocysteine-tRNA specific	Downregulated
31856	Exônico	57696	DDX55	DEAD (Asp-Glu-Ala-Asp) box polypeptide 55	Downregulated
31863	Exônico	55250	STATIP1	signal transducer and activator of transcription 3 interacting protein 1	Downregulated
31871	Exônico	29079	MED4	mediator of RNA polymerase II transcription, subunit 4 homolog (yeast)	Downregulated
31880	Exônico	23045	KIAA0826	KIAA0826	Downregulated

31888	Exônico	10059	DNM1L	dynamin 1-like	Downregulated
31925	Exônico	1145	CHRNE	cholinergic receptor, nicotinic, epsilon polypeptide	Downregulated
31961	Exônico	57473	-	-	Downregulated
31970	Exônico	54464	XRN1	5'-3' exoribonuclease 1	Downregulated
31971	Exônico	51692	CPSF3	cleavage and polyadenylation specific factor 3, 73kDa	Downregulated
31991	Exônico	10161	P2RY5	purinergic receptor P2Y, G- protein coupled, 5	Downregulated
32001	Exônico	8412	BCAR3	breast cancer anti-estrogen resistance 3	Downregulated
32046	Exônico	143162	FRMPD2	FERM and PDZ domain containing 2	Downregulated
32050	Exônico	91949	COG7	component of oligomeric golgi complex 7	Downregulated
32056	Exônico	80196	RNF34	ring finger protein 34	Downregulated
32082	Exônico	23512	SUZ12	suppressor of zeste 12 homolog (Drosophila)	Downregulated
32088	Exônico	11235	PDCD10	programmed cell death 10	Downregulated
32089	Exônico	11051	CPSF5	cleavage and polyadenylation specific factor 5, 25 kDa	Downregulated
32105	Exônico	7589	ZNF38	zinc finger protein 38	Downregulated
32129	Exônico	1736	DKC1	dyskeratosis congenita 1, dyskerin	Downregulated
32146	Exônico	221074	SLC39A12	solute carrier family 39 (zinc transporter), member 12	Downregulated
32148	Exônico	159195	USP54	ubiquitin specific protease 54	Downregulated
32174	Exônico	55028	-	-	Downregulated
32181	Exônico	28964	GIT1	G protein-coupled receptor kinase interactor 1	Downregulated
32188	Exônico	23233	SEC15L2	SEC15-like 2 (S. cerevisiae)	Downregulated
32194	Exônico	10732	TCFL5	transcription factor-like 5 (basic helix-loop-helix)	Downregulated
32216	Exônico	5859	QARS	glutaminyl-tRNA synthetase	Downregulated
32227	Exônico	3184	HNRPD	heterogeneous nuclear ribonucleoprotein D (AU- rich element RNA binding protein 1, 37kDa)	Downregulated
32250	Exônico	164684	-	-	Downregulated
32264	Exônico	79693	-	-	Downregulated
32268	Exônico	57695	USP37	ubiquitin specific protease 37	Downregulated
32272	Exônico	55832	CAND1	cullin-associated and neddylation-dissociated 1	Downregulated
32275	Exônico	55249	YY1AP1	YY1 associated protein 1	Downregulated
32300	Exônico	10059	DNM1L	dynamin 1-like	Downregulated
32305	Exônico	9258	MFHAS1	malignant fibrous histiocytoma amplified sequence 1	Downregulated
32337	Exônico	1138	CHRNA5	cholinergic receptor, nicotinic, alpha polypeptide 5	Downregulated
32342	Exônico	22	ABCB7	ATP-binding cassette, sub- family B (MDR/TAP), member 7	Downregulated
32356	Exônico	138050	-	-	Downregulated

32373	Exônico	57475	PLEKHH1	pleckstrin homology domain containing, family H (with MyTH4 domain) member 1	Downregulated
32381	Exônico	54815	GATAD2A	GATA zinc finger domain containing 2A	Downregulated
32384	Exônico	51379	CRLF3	cytokine receptor-like factor 3	Downregulated
32390	Exônico	25924	MYRIP	myosin VIIA and Rab interacting protein	Downregulated
32399	Exônico	11004	KIF2C	kinesin family member 2C	Downregulated
32408	Exônico	9446	GSTO1	glutathione S-transferase omega 1	Downregulated
32417	Exônico	7073	TIAL1	TIA1 cytotoxic granule-associated RNA binding protein-like 1	Downregulated
32418	Exônico	6867	TACC1	transforming, acidic coiled-coil containing protein 1	Downregulated
32423	Exônico	5707	PSMD1	proteasome (prosome, macropain) 26S subunit, non-ATPase, 1	Downregulated
32430	Exônico	4087	SMAD2	SMAD, mothers against DPP homolog 2 (Drosophila)	Downregulated
32435	Exônico	2729	GCLC	glutamate-cysteine ligase, catalytic subunit	Downregulated
32460	Exônico	120224	TMEM45B	transmembrane protein 45B	Downregulated
32468	Exônico	80196	RNF34	ring finger protein 34	Downregulated
32478	Exônico	55832	CAND1	cullin-associated and neddylation-dissociated 1	Downregulated
32500	Exônico	11235	PDCD10	programmed cell death 10	Downregulated
32501	Exônico	11051	CPSF5	cleavage and polyadenylation specific factor 5, 25 kDa	Downregulated
32509	Exônico	9657	IQCB1	IQ motif containing B1	Downregulated
32512	Exônico	9055	PRC1	protein regulator of cytokinesis 1	Downregulated
32542	Exônico	1434	CSE1L	CSE1 chromosome segregation 1-like (yeast)	Downregulated
32565	Exônico	113115	FAM54A	family with sequence similarity 54, member A	Downregulated
32576	Exônico	64282	PAPD5	PAP associated domain containing 5	Downregulated
32585	Exônico	55204	GOLPH3L	golgi phosphoprotein 3-like	Downregulated
32589	Exônico	51692	CPSF3	cleavage and polyadenylation specific factor 3, 73kDa	Downregulated
32593	Exônico	28968	SLC6A16	solute carrier family 6, member 16	Downregulated
32602	Exônico	23028	AOF2	amine oxidase (flavin containing) domain 2	Downregulated
32609	Exônico	10162	-	-	Downregulated
32612	Exônico	9748	SLK	STE20-like kinase (yeast)	Downregulated
32626	Exônico	6434	SFRS10	splicing factor, arginine/serine-rich 10 (transformer 2 homolog, Drosophila)	Downregulated
32630	Exônico	5525	PPP2R5A	protein phosphatase 2, regulatory subunit B (B56), alpha isoform	Downregulated

32637	Exônico	3838	KPNA2	karyopherin alpha 2 (RAG cohort 1, importin alpha 1)	Downregulated
32647	Exônico	1105	CHD1	chromodomain helicase DNA binding protein 1	Downregulated
32670	Exônico	84735	CNDP1	carnosine dipeptidase 1 (metallopeptidase M20 family)	Downregulated
32695	Exônico	29079	MED4	mediator of RNA polymerase II transcription, subunit 4 homolog (yeast)	Downregulated
32705	Exônico	22906	-	-	Downregulated
32711	Exônico	10207	INADL	InaD-like (Drosophila)	Downregulated
32718	Exônico	9053	MAP7	microtubule-associated protein 7	Downregulated
32741	Exônico	3309	HSPA5	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)	Downregulated
32749	Exônico	1138	CHRNA5	cholinergic receptor, nicotinic, alpha polypeptide 5	Downregulated
32765	Exônico	196527	TMEM16F	transmembrane protein 16F	Downregulated
32769	Exônico	126823	-	-	Downregulated
32773	Exônico	85021	REPS1	RALBP1 associated Eps domain containing 1	Downregulated
32829	Exônico	7073	TIAL1	TIA1 cytotoxic granule-associated RNA binding protein-like 1	Downregulated
32854	Exônico	818	CAMK2G	calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma	Downregulated
32868	Exônico	163720	-	-	Downregulated
32872	Exônico	120103	SLC36A4	solute carrier family 36 (proton/amino acid symporter), member 4	Downregulated
32880	Exônico	80185	-	-	Downregulated
32882	Exônico	79693	-	-	Downregulated
32893	Exônico	55247	NEIL3	nei endonuclease VIII-like 3 (E. coli)	Downregulated
32923	Exônico	9255	SCYE1	small inducible cytokine subfamily E, member 1 (endothelial monocyte-activating)	Downregulated
32944	Exônico	4149	MAX	MYC associated factor X	Downregulated
32997	Exônico	55204	GOLPH3L	golgi phosphoprotein 3-like	Downregulated
32999	Exônico	54819	ZCCHC10	zinc finger, CCHC domain containing 10	Downregulated
33000	Exônico	54467	ANKIB1	ankyrin repeat and IBR domain containing 1	Downregulated
33016	Exônico	11186	RASSF1	Ras association (RalGDS/AF-6) domain family 1	Downregulated
33017	Exônico	11004	KIF2C	kinesin family member 2C	Downregulated
33041	Exônico	5708	PSMD2	proteasome (prosome, macropain) 26S subunit, non-ATPase, 2	Downregulated
33048	Exônico	4087	SMAD2	SMAD, mothers against DPP homolog 2 (Drosophila)	Downregulated
33073	Exônico	200894	ARL2L1	ADP-ribosylation factor-like 2-like 1	Downregulated
33092	Exônico	57688	ZSWIM6	zinc finger, SWIM-type containing 6	Downregulated
33107	Exônico	29072	-	-	Downregulated

33128	Exônico	9491	PSMF1	proteasome (prosome, macropain) inhibitor subunit 1 (PI31)	Downregulated
33150	Exônico	4149	MAX	MYC associated factor X	Downregulated
33161	Exônico	1130	LYST	lysosomal trafficking regulator	Downregulated
33174	Exônico	285315	-	-	Downregulated
33181	Exônico	126823	-	-	Downregulated
33183	Exônico	113174	-	-	Downregulated
33191	Exônico	79866	-	-	Downregulated
33193	Exônico	64921	-	-	Downregulated
33199	Exônico	56204	-	-	Downregulated
33213	Exônico	26130	GAPVD1	GTPase activating protein and VPS9 domains 1	Downregulated
33246	Exônico	5861	RAB1A	RAB1A, member RAS oncogene family	Downregulated
33257	Exônico	3184	HNRPD	heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	Downregulated
33258	Exônico	2932	GSK3B	glycogen synthase kinase 3 beta	Downregulated
33269	Exônico	157	ADRBK2	adrenergic, beta, receptor kinase 2	Downregulated
33270	Exônico				Downregulated
33283	Exônico	130013	ACMSD	aminocarboxymuconate semialdehyde decarboxylase	Downregulated
33290	Exônico	83715	ESPN	espin	Downregulated
33294	Exônico	79685	-	-	Downregulated
33303	Exônico	55692	LUC7L	LUC7-like (S. cerevisiae)	Downregulated
33306	Exônico	55081	ESRRBL1	estrogen-related receptor beta like 1	Downregulated
33316	Exônico	25962	-	-	Downregulated
33325	Exônico	11051	CPSF5	cleavage and polyadenylation specific factor 5, 25 kDa	Downregulated
33343	Exônico	7095	TLOC1	translocation protein 1	Downregulated
33350	Exônico	5558	PRIM2A	primase, polypeptide 2A, 58kDa	Downregulated
33366	Exônico	1434	CSE1L	CSE1 chromosome segregation 1-like (yeast)	Downregulated
33395	Exônico	80745	THUMPD2	THUMP domain containing 2	Downregulated
33402	Exônico	57634	EP400	E1A binding protein p400	Downregulated
33407	Exônico	55666	-	-	Downregulated
33419	Exônico	26130	GAPVD1	GTPase activating protein and VPS9 domains 1	Downregulated
33429	Exônico	11004	KIF2C	kinesin family member 2C	Downregulated
33435	Exônico	9894	-	-	Downregulated
33439	Exônico	9208	LRRFIP1	leucine rich repeat (in FLII) interacting protein 1	Downregulated
33445	Exônico	7508	XPC	xeroderma pigmentosum, complementation group C	Downregulated
33463	Exônico	3184	HNRPD	heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	Downregulated
33502	Exônico	64419	-	-	Downregulated
33505	Exônico	57509	MTUS1	mitochondrial tumor suppressor 1	Downregulated
33516	Exônico	51460	SFMBT1	Scm-like with four mbt domains 1	Downregulated
33517	Exônico	51185	CRBN	cereblon	Downregulated

33525	Exônico	23352	-	-	Downregulated
33533	Exônico	10576	CCT2	chaperonin containing TCP1, subunit 2 (beta)	Downregulated
33537	Exônico	9913	-	-	Downregulated
33544	Exônico	8621	CDC2L5	cell division cycle 2-like 5 (cholinesterase-related cell division controller)	Downregulated
33548	Exônico	7326	UBE2G1	ubiquitin-conjugating enzyme E2G 1 (UBC7 homolog, C. elegans)	Downregulated
33553	Exônico	6197	RPS6KA3	ribosomal protein S6 kinase, 90kDa, polypeptide 3	Downregulated
33578	Exônico	21	ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	Downregulated
33589	Exônico	196528	ARID2	AT rich interactive domain 2 (ARID, RFX-like)	Downregulated
33590	Exônico	160065	-	-	Downregulated
33606	Exônico	64285	RHBDF1	rhomboid family 1 (Drosophila)	Downregulated
33607	Exônico	60313	-	-	Downregulated
33615	Exônico	55205	ZNF532	zinc finger protein 532	Downregulated
33621	Exônico	51115	-	-	Downregulated
33622	Exônico	30001	ERO1L	ERO1-like (S. cerevisiae)	Downregulated
33632	Exônico	23032	USP33	ubiquitin specific protease 33	Downregulated
33643	Exônico	9637	FEZ2	fasciculation and elongation protein zeta 2 (zygin II)	Downregulated
33644	Exônico	9448	MAP4K4	mitogen-activated protein kinase kinase kinase 4	Downregulated
33659	Exônico	5708	PSMD2	proteasome (prosome, macropain) 26S subunit, non-ATPase, 2	Downregulated
33681	Exônico	158	ADSL	adenylosuccinate lyase	Downregulated
33689	Exônico	266655	-	-	Downregulated
33694	Exônico	142686	ASB14	ankyrin repeat and SOCS box-containing 14	Downregulated
33696	Exônico	120071	-	-	Downregulated
33709	Exônico	60672	-	-	Downregulated
33717	Exônico	55246	-	-	Downregulated
33718	Exônico	55076	TMEM45A	transmembrane protein 45A	Downregulated
33737	Exônico	11047	ADRM1	adhesion regulating molecule 1	Downregulated
33740	Exônico	10419	SKB1	SKB1 homolog (S. pombe)	Downregulated
33759	Exônico	6197	RPS6KA3	ribosomal protein S6 kinase, 90kDa, polypeptide 3	Downregulated
33775	Exônico	2131	EXT1	exostoses (multiple) 1	Downregulated
33802	Exônico	91151	TIGD7	tigger transposable element derived 7	Downregulated
33834	Exônico	23468	CBX5	chromobox homolog 5 (HP1 alpha homolog, Drosophila)	Downregulated
33836	Exônico	23235	SNF1LK2	SNF1-like kinase 2	Downregulated
33867	Exônico	5306	PITPNA	phosphatidylinositol transfer protein, alpha	Downregulated
33878	Exônico	2305	FOXM1	forkhead box M1	Downregulated
33908	Exônico	83707	-	-	Downregulated
33926	Exônico	54529	-	-	Downregulated
33929	Exônico	51185	CRBN	cereblon	Downregulated
33936	Exônico	23503	ZFYVE26	zinc finger, FYVE domain containing 26	Downregulated
33949	Exônico	9913	-	-	Downregulated

33957	Exônico	8458	TTF2	transcription termination factor, RNA polymerase II	Downregulated
33962	Exônico	6907	TBL1X	transducin (beta)-like 1X-linked	Downregulated
33968	Exônico	5557	PRIM1	primase, polypeptide 1, 49kDa	Downregulated
33982	Exônico	1994	ELAVL1	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	Downregulated
33984	Exônico	1434	CSE1L	CSE1 chromosome segregation 1-like (yeast)	Downregulated
34002	Exônico	160065	-	-	Downregulated
34011	Exônico	84187	-	-	Downregulated
34015	Exônico	79867	-	-	Downregulated
34017	Exônico	64924	SLC30A5	solute carrier family 30 (zinc transporter), member 5	Downregulated
34018	Exônico	64318	C10orf117	chromosome 10 open reading frame 117	Downregulated
34027	Exônico	55206	SBNO1	sno, strawberry notch homolog 1 (Drosophila)	Downregulated
34029	Exônico	54820	NDE1	nudE nuclear distribution gene E homolog 1 (A. nidulans)	Downregulated
34054	Exônico	9750	C6orf32	chromosome 6 open reading frame 32	Downregulated
34055	Exônico	9637	FEZ2	fasciculation and elongation protein zeta 2 (zyglin II)	Downregulated
34057	Exônico	9209	LRRFIP2	leucine rich repeat (in FLII) interacting protein 2	Downregulated
34062	Exônico	7982	ST7	suppression of tumorigenicity 7	Downregulated
34093	Exônico	158	ADSL	adenylosuccinate lyase	Downregulated
34120	Exônico	64417	-	-	Downregulated
34141	Exônico	25771	TBC1D22A	TBC1 domain family, member 22A	Downregulated
34148	Exônico	11234	HPS5	Hermansky-Pudlak syndrome 5	Downregulated
34151	Exônico	10576	CCT2	chaperonin containing TCP1, subunit 2 (beta)	Downregulated
34152	Exônico	10419	SKB1	SKB1 homolog (S. pombe)	Downregulated
34183	Exônico	3295	HSD17B4	hydroxysteroid (17-beta) dehydrogenase 4	Downregulated
34207	Exônico	196528	ARID2	AT rich interactive domain 2 (ARID, RFX-like)	Downregulated
34211	Exônico	127018	LYPLAL1	lysophospholipase-like 1	Downregulated
34222	Exônico	79621	-	-	Downregulated
34239	Exônico	51118	-	-	Downregulated
34249	Exônico	23116	KIAA0423	KIAA0423	Downregulated
34250	Exônico	23032	USP33	ubiquitin specific protease 33	Downregulated
34251	Exônico	22871	NLGN1	neuroligin 1	Downregulated
34252	Exônico	11188	NISCH	nischarin	Downregulated
34254	Exônico	10734	STAG3	stromal antigen 3	Downregulated
34262	Exônico	9448	MAP4K4	mitogen-activated protein kinase kinase kinase 4	Downregulated
34268	Exônico	7982	ST7	suppression of tumorigenicity 7	Downregulated
34272	Exônico	6867	TACC1	transforming, acidic coiled-coil containing protein 1	Downregulated
34289	Exônico	2729	GCLC	glutamate-cysteine ligase, catalytic subunit	Downregulated

34322	Exônico	80184	-	-	Downregulated
34323	Exônico	79903	-	-	Downregulated
34332	Exônico	55831	-	-	Downregulated
34333	Exônico	55692	LUC7L	LUC7-like (S. cerevisiae)	Downregulated
34338	Exônico	54522	ANKRD16	ankyrin repeat domain 16	Downregulated
34355	Exônico	11045	UPK1A	uroplakin 1A	Downregulated
34363	Exônico	9656	MDC1	mediator of DNA damage checkpoint 1	Downregulated
34369	Exônico	8455	ATRN	attractin	Downregulated
34372	Exônico	7326	UBE2G1	ubiquitin-conjugating enzyme E2G 1 (UBC7 homolog, C. elegans)	Downregulated
34392	Exônico	2332	FMR1	fragile X mental retardation 1	Downregulated
34412	Exônico	221120	-	-	Downregulated
34427	Exônico	79869	-	-	Downregulated
34452	Exônico	23469	PHF3	PHD finger protein 3	Downregulated
34482	Exônico	5861	RAB1A	RAB1A, member RAS oncogene family	Downregulated
34495	Exônico	2729	GCLC	glutamate-cysteine ligase, catalytic subunit	Downregulated
34516	Exônico	163223	ZNF676	zinc finger protein 676	Downregulated
34524	Exônico	84727	-	-	Downregulated
34533	Exônico	60625	DHX35	DEAH (Asp-Glu-Ala-His) box polypeptide 35	Downregulated
34537	Exônico	56474	CTPS2	CTP synthase II polymerase (DNA directed), gamma 2, accessory subunit	Downregulated
34560	Exônico	11232	POLG2	chaperonin containing TCP1, subunit 4 (delta)	Downregulated
34563	Exônico	10575	CCT4	-	Downregulated
34566	Exônico	10054	-	-	Downregulated
34567	Exônico	9912	-	-	Downregulated
34589	Exônico	4976	OPA1	optic atrophy 1 (autosomal dominant)	Downregulated
34602	Exônico	1432	MAPK14	mitogen-activated protein kinase 14	Downregulated
34616	Exônico	285331	-	-	Downregulated
34634	Exônico	79622	C16orf33	chromosome 16 open reading frame 33	Downregulated
34637	Exônico	60313	-	-	Downregulated
34639	Exônico	57478	USP31	ubiquitin specific protease 31	Downregulated
34651	Exônico	51123	-	-	Downregulated
34656	Exônico	25927	C2orf32	chromosome 2 open reading frame 32	Downregulated
34663	Exônico	22872	SEC31L1	SEC31-like 1 (S. cerevisiae)	Downregulated
34664	Exônico	11190	CEP2	centrosomal protein 2	Downregulated
34665	Exônico	11010	GLIPR1	GLI pathogenesis-related 1 (glioma)	Downregulated
34670	Exônico	10016	PDCD6	programmed cell death 6	Downregulated
34684	Exônico	6868	ADAM17	a disintegrin and metalloproteinase domain 17 (tumor necrosis factor, alpha, converting enzyme)	Downregulated
34699	Exônico	3187	HNRPH1	heterogeneous nuclear ribonucleoprotein H1 (H)	Downregulated
34701	Exônico	2730	GCLM	glutamate-cysteine ligase, modifier subunit	Downregulated
34719	Exônico	266629	SEC14L3	SEC14-like 3 (S. cerevisiae)	Downregulated
34723	Exônico	151525	WDSUB1	WD repeat, SAM and U-box domain containing 1	Downregulated
34726	Exônico	119710	-	-	Downregulated
34729	Exônico	85476	GFM1	G elongation factor, mitochondrial 1	Downregulated
34736	Exônico	79685	-	-	Downregulated

34748	Exônico	55075	UACA	uveal autoantigen with coiled-coil domains and ankyrin repeats	Downregulated
34767	Exônico	11044	POLS	polymerase (DNA directed) sigma	Downregulated
34784	Exônico	7325	UBE2E2	ubiquitin-conjugating enzyme E2E 2 (UBC4/5 homolog, yeast)	Downregulated
34806	Exônico	1994	ELAVL1	ELAV (embryonic lethal, abnormal vision, Drosophila)-like 1 (Hu antigen R)	Downregulated
34816	Exônico	22931	RAB18	RAB18, member RAS oncogene family	Downregulated
34833	Exônico	85363	TRIM5	tripartite motif-containing 5	Downregulated
34834	Exônico	84641	-	-	Downregulated
34850	Exônico	55422	ZNF331	zinc finger protein 331	Downregulated
34853	Exônico	54822	TRPM7	transient receptor potential cation channel, subfamily M, member 7	Downregulated
34863	Exônico	23683	PRKD3	protein kinase D3	Downregulated
34865	Exônico	23327	NEDD4L	neural precursor cell expressed, developmentally down-regulated 4-like	Downregulated
34903	Exônico	3840	KPNA4	karyopherin alpha 4 (importin alpha 3)	Downregulated
34933	Exônico	114785	MBD6	methyl-CpG binding domain protein 6	Downregulated
34936	Exônico	84727	-	-	Downregulated
34948	Exônico	57153	-	-	Downregulated
34968	Exônico	23253	ANKRD12	ankyrin repeat domain 12	Downregulated
34975	Exônico	10575	CCT4	chaperonin containing TCP1, subunit 4 (delta)	Downregulated
34976	Exônico	10419	SKB1	SKB1 homolog (S. pombe)	Downregulated
34992	Exônico	6905	TBCE	tubulin-specific chaperone e	Downregulated
35037	Exônico	113251	-	-	Downregulated
35042	Exônico	83636	C19orf12	chromosome 19 open reading frame 12	Downregulated
35046	Exônico	79622	C16orf33	chromosome 16 open reading frame 33	Downregulated
35051	Exônico	57478	USP31	ubiquitin specific protease 31	Downregulated
35094	Exônico	7286	TUFT1	tuftelin 1	Downregulated
35100	Exônico	5862	RAB2	RAB2, member RAS oncogene family	Downregulated
35106	Exônico	4685	NCAM2	neural cell adhesion molecule 2	Downregulated
35123	Exônico	159	ADSS	adenylosuccinate synthase	Downregulated
35131	Exônico	261734	NPHP4	nephronophthisis 4	Downregulated
35195	Exônico	7580	ZNF32	zinc finger protein 32 (KOX 30)	Downregulated
35202	Exônico	5889	RAD51C	RAD51 homolog C (S. cerevisiae)	Downregulated
35204	Exônico	5537	PPP6C	protein phosphatase 6, catalytic subunit	Downregulated
35205	Exônico	5356	PLRG1	pleiotropic regulator 1 (PRL1 homolog, Arabidopsis)	Downregulated
35240	Exônico	138639	PTPDC1	protein tyrosine phosphatase domain containing 1	Downregulated
35245	Exônico	85363	TRIM5	tripartite motif-containing 5	Downregulated

35257	Exônico	57478	USP31	ubiquitin specific protease 31	Downregulated
35264	Exônico	55030	FBXO34	F-box protein 34	Downregulated
35266	Exônico	54477	PLEKHA5	pleckstrin homology domain containing, family A member 5	Downregulated
35268	Exônico	51390	AIG1	androgen-induced 1	Downregulated
35269	Exônico	51123	-	-	Downregulated
35273	Exônico	26133	TRPC4AP	transient receptor potential cation channel, subfamily C member 4 associated protein	Downregulated
35279	Exônico	23118	MAP3K7IP2	mitogen-activated protein kinase kinase kinase 7 interacting protein 2	Downregulated
35286	Exônico	10393	ANAPC10	anaphase promoting complex subunit 10	Downregulated
35299	Exônico	7515	XRCC1	X-ray repair complementing defective repair in Chinese hamster cells 1	Downregulated
35312	Exônico	4690	NCK1	NCK adaptor protein 1	Downregulated
35316	Exônico	3636	INPPL1	inositol polyphosphate phosphatase-like 1	Downregulated
35325	Exônico	1105	CHD1	chromodomain helicase DNA binding protein 1	Downregulated
35348	Exônico	84726	KIAA0515	KIAA0515	Downregulated
35359	Exônico	57508	-	-	Downregulated
35366	Exônico	55071	C9orf40	chromosome 9 open reading frame 40	Downregulated
35368	Exônico	54521	WDR44	WD repeat domain 44	Downregulated
35370	Exônico	51454	GULP1	GULP, engulfment adaptor PTB domain containing 1	Downregulated
35374	Exônico	27113	BBC3	BCL2 binding component 3	Downregulated
35381	Exônico	23142	-	-	Downregulated
35386	Exônico	10781	ZNF266	zinc finger protein 266	Downregulated
35399	Exônico	8454	CUL1	cullin 1	Downregulated
35404	Exônico	6905	TBCE	tubulin-specific chaperone e	Downregulated
35408	Exônico	5889	RAD51C	RAD51 homolog C (S. cerevisiae)	Downregulated
35431	Exônico	208	AKT2	v-akt murine thymoma viral oncogene homolog 2	Downregulated
35433	Exônico				Downregulated
35462	Exônico	57643	ZSWIM5	zinc finger, SWIM-type containing 5	Downregulated
35469	Exônico	55206	SBNO1	sno, strawberry notch homolog 1 (Drosophila)	Downregulated
35488	Exônico	11190	CEP2	centrosomal protein 2	Downregulated
35502	Exônico	8604	SLC25A12	solute carrier family 25 (mitochondrial carrier, Aralar), member 12	Downregulated
35510	Exônico	6448	SGSH	N-sulfoglucosamine sulfohydrolase (sulfamidase)	Downregulated
35535	Exônico	161	AP2A2	adaptor-related protein complex 2, alpha 2 subunit	Downregulated
35554	Exônico	84726	KIAA0515	KIAA0515	Downregulated
35557	Exônico	80853	-	-	Downregulated
35566	Exônico	57153	-	-	Downregulated
35567	Exônico	56474	CTPS2	CTP synthase II	Downregulated
35594	Exônico	10419	SKB1	SKB1 homolog (S. pombe)	Downregulated
35599	Exônico	9654	TTLL4	tubulin tyrosine ligase-like family, member 4	Downregulated

35602	Exônico	9049	AIP	aryl hydrocarbon receptor interacting protein	Downregulated
35616	Exônico	5537	PPP6C	protein phosphatase 6, catalytic subunit	Downregulated
35661	Exônico	80746	TSEN2	tRNA splicing endonuclease 2 homolog (SEN2, <i>S. cerevisiae</i>)	Downregulated
35662	Exônico	80124	VCPIP1	valosin containing protein (p97)/p47 complex interacting protein 1	Downregulated
35663	Exônico	79869	-	-	Downregulated
35664	Exônico	79623	GALNT14	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 14 (GalNAc-T14)	Downregulated
35678	Exônico	54477	PLEKHA5	pleckstrin homology domain containing, family A member 5	Downregulated
35710	Exônico	7988	ZNF212	zinc finger protein 212	Downregulated
35726	Exônico	4094	MAF	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)	Downregulated
35737	Exônico	1106	CHD2	chromodomain helicase DNA binding protein 2	Downregulated
35740	Exônico	375	ARF1	ADP-ribosylation factor 1	Downregulated
35742	Exônico	51022	GLRX2	glutaredoxin 2	Downregulated
35765	Exônico	79902	PCNT1	pericentrin 1	Downregulated
35792	Exônico	23250	ATP11A	ATPase, Class VI, type 11A	Downregulated
35796	Exônico	11232	POLG2	polymerase (DNA directed), gamma 2, accessory subunit	Downregulated
35799	Exônico	10575	CCT4	chaperonin containing TCP1, subunit 4 (delta)	Downregulated
35811	Exônico	8454	CUL1	cullin 1	Downregulated
35812	Exônico	8038	ADAM12	a disintegrin and metalloproteinase domain 12 (meltrin alpha)	Downregulated
35813	Exônico	7562	ZNF15L1	zinc finger protein 15-like 1 (KOX 8)	Downregulated
35814	Exônico	7323	UBE2D3	ubiquitin-conjugating enzyme E2D 3 (UBC4/5 homolog, yeast)	Downregulated
35827	Exônico	4363	ABCC1	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	Downregulated
35842	Exônico	439	ASNA1	arsA arsenite transporter, ATP-binding, homolog 1 (bacterial)	Downregulated
35846	Exônico	5987	RFP	ret finger protein	Downregulated
35865	Exônico	84191	-	-	Downregulated
35874	Exônico	57643	ZSWIM5	zinc finger, SWIM-type containing 5	Downregulated
35875	Exônico	57479	KIAA1205	KIAA1205	Downregulated
35901	Exônico	11010	GLIPR1	GLI pathogenesis-related 1 (glioma)	Downregulated
35912	Exônico	7874	USP7	ubiquitin specific protease 7 (herpes virus-associated)	Downregulated
35914	Exônico	8604	SLC25A12	solute carrier family 25 (mitochondrial carrier, Aralar), member 12	Downregulated
35922	Exônico	6453	ITSN1	intersectin 1 (SH3 domain protein)	Downregulated

35932	Exônico	4094	MAF	v-maf musculoaponeurotic fibrosarcoma oncogene homolog (avian)	Downregulated
35936	Exônico	2935	GSPT1	G1 to S phase transition 1	Downregulated
35947	Exônico	161	AP2A2	adaptor-related protein complex 2, alpha 2 subunit	Downregulated
35961	Exônico	129642	OACT2	O-acyltransferase (membrane bound) domain containing 2	Downregulated
35964	Exônico	91768	CABLES1	Cdk5 and Abl enzyme substrate 1	Downregulated
35972	Exônico	79682	MLF1IP	MLF1 interacting protein	Downregulated
35978	Exônico	57148	-	-	Downregulated
35979	Exônico	56300	IL1F9	interleukin 1 family, member 9	Downregulated
35995	Exônico	25766	-	-	Downregulated
35996	Exônico	23500	DAAM2	dishevelled associated activator of morphogenesis 2	Downregulated
36042	Exônico	1983	EIF5	eukaryotic translation initiation factor 5	Downregulated
36043	Exônico	1729	DIAPH1	diaphanous homolog 1 (Drosophila)	Downregulated
36064	Exônico	138639	PTPDC1	protein tyrosine phosphatase domain containing 1	Downregulated
36096	Exônico	27067	STAU2	staufen, RNA binding protein, homolog 2 (Drosophila)	Downregulated
36098	Exônico	25929	GEMIN5	gem (nuclear organelle) associated protein 5	Downregulated
36104	Exônico	23032	USP33	ubiquitin specific protease 33	Downregulated
36131	Exônico	5709	PSMD3	proteasome (prosome, macropain) 26S subunit, non-ATPase, 3	Downregulated
36132	Exônico	5527	PPP2R5C	protein phosphatase 2, regulatory subunit B (B56), gamma isoform	Downregulated
36153	Exônico	161	AP2A2	adaptor-related protein complex 2, alpha 2 subunit	Downregulated
36161	Exônico	260425	MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3	Downregulated
36190	Exônico	55069	-	-	Downregulated
36205	Exônico	23142	-	-	Downregulated
36213	Exônico	10199	MPHOSPH10	M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein)	Downregulated
36216	Exônico	9780	FAM38A	family with sequence similarity 38, member A	Downregulated
36217	Exônico	9652	KIAA0372	KIAA0372	Downregulated
36221	Exônico	8826	IQGAP1	IQ motif containing GTPase activating protein 1	Downregulated
36244	Exônico	2976	GTF3C2	general transcription factor IIIC, polypeptide 2, beta 110kDa	Downregulated
36257	Exônico	81561	C2orf31	chromosome 2 open reading frame 31	Downregulated
36263	Exônico	379025	-	-	Downregulated
36272	Exônico	116987	CENTG2	centaurin, gamma 2	Downregulated
36273	Exônico	113251	-	-	Downregulated
36275	Exônico	85369	FAM40A	family with sequence similarity 40, member A	Downregulated

36296	Exônico	54487	DGCR8	DiGeorge syndrome critical region gene 8	Downregulated
36307	Exônico	23328	SASH1	SAM and SH3 domain containing 1	Downregulated
36322	Exônico	9451	EIF2AK3	eukaryotic translation initiation factor 2-alpha kinase 3	Downregulated
36323	Exônico	130535	KCTD18	potassium channel tetramerisation domain containing 18	Downregulated
36339	Exônico	5311	PKD2	polycystic kidney disease 2 (autosomal dominant)	Downregulated
36355	Exônico	1106	CHD2	chromodomain helicase DNA binding protein 2	Downregulated
36367	Exônico	260425	MAGI3	membrane associated guanylate kinase, WW and PDZ domain containing 3	Downregulated
36388	Exônico	57680	CHD8	chromodomain helicase DNA binding protein 8	Downregulated
36395	Exônico	55244	-	-	Downregulated
36400	Exônico	51447	IHPK2	inositol hexaphosphate kinase 2	Downregulated
36407	Exônico	24148	C20orf14	chromosome 20 open reading frame 14	Downregulated
36413	Exônico	22905	EPN2	epsin 2	Downregulated
36415	Exônico	11044	POLS	polymerase (DNA directed) sigma	Downregulated
36428	Exônico	8621	CDC2L5	cell division cycle 2-like 5 (cholinesterase-related cell division controller)	Downregulated
36431	Exônico	7556	ZNF10	zinc finger protein 10	Downregulated
36438	Exônico	5889	RAD51C	RAD51 homolog C (S. cerevisiae)	Downregulated
36472	Exônico	221303	C6orf189	chromosome 6 open reading frame 189	Downregulated
36485	Exônico	80755	-	-	Downregulated
36497	Exônico	55676	SLC30A6	solute carrier family 30 (zinc transporter), member 6	Downregulated
36500	Exônico	55034	MOCOS	molybdenum cofactor sulfurase	Downregulated
36529	Exônico	9209	LRRFIP2	leucine rich repeat (in FLII) interacting protein 2	Downregulated
36531	Exônico	8798	DYRK4	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 4	Downregulated
36550	Exônico	4117	MAK	male germ cell-associated kinase	Downregulated
36584	Exônico	84725	PLEKHA8	pleckstrin homology domain containing, family A (phosphoinositide binding specific) member 8	Downregulated
36585	Exônico	84259	-	-	Downregulated
36588	Exônico	80179	MYOHD1	myosin head domain containing 1	Downregulated
36598	Exônico	55827	IQWD1	IQ motif and WD repeats 1	Downregulated
36602	Exônico	55069	-	-	Downregulated
36604	Exônico	54517	-	-	Downregulated
36615	Exônico	23351	KIAA0323	KIAA0323	Downregulated
36617	Exônico	23141	-	-	Downregulated
36620	Exônico	11231	SEC63	SEC63-like (S. cerevisiae)	Downregulated
36629	Exônico	9652	KIAA0372	KIAA0372	Downregulated
36667	Exônico	204	AK2	adenylate kinase 2	Downregulated
36677	Exônico	257160	-	-	Downregulated

36692	Exônimo	80135	BXDC5	brix domain containing 5	Downregulated
36704	Exônimo	55450	-	-	Downregulated
36709	Exônimo	51701	NLK	nemo like kinase	Downregulated
36724	Exônimo	11193	WBP4	WW domain binding protein 4 (formin binding protein 21)	Downregulated
36734	Exônimo	9453	GGPS1	geranylgeranyl diphosphate synthase 1	Downregulated
36735	Exônimo	9209	LRRFIP2	leucine rich repeat (in FLII) interacting protein 2	Downregulated
36750	Exônimo	5527	PPP2R5C	protein phosphatase 2, regulatory subunit B (B56), gamma isoform	Downregulated
36761	Exônimo	2734	GLG1	golgi apparatus protein 1	Downregulated
36765	Exônimo	1653	DDX1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	Downregulated
36768	Exônimo	831	CAST	calpastatin	Downregulated
36772	Exônimo				Downregulated
36788	Exônimo	91746	-	-	Downregulated
36796	Exônimo	79677	SMC6L1	SMC6 structural maintenance of chromosomes 6-like 1 (yeast)	Downregulated
36813	Exônimo	51172	NAGPA	N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase	Downregulated
36825	Exônimo	22903	BTBD3	BTB (POZ) domain containing 3	Downregulated
36829	Exônimo	10574	CCT7	chaperonin containing TCP1, subunit 7 (eta)	Downregulated
36837	Exônimo	9238	TBRG4	transforming growth factor beta regulator 4	Downregulated
36841	Exônimo	8453	CUL2	cullin 2	Downregulated
36842	Exônimo	8036	SHOC2	soc-2 suppressor of clear homolog (C. elegans)	Downregulated
36850	Exônimo	5887	RAD23B	RAD23 homolog B (S. cerevisiae)	Downregulated
36866	Exônimo	1983	EIF5	eukaryotic translation initiation factor 5	Downregulated
36873	Exônimo	204	AK2	adenylate kinase 2	Downregulated
36909	Exônimo	55676	SLC30A6	solute carrier family 30 (zinc transporter), member 6	Downregulated
36910	Exônimo	55450	-	-	Downregulated
36947	Exônimo	7520	XRCC5	X-ray repair complementing defective repair in Chinese hamster cells 5 (double-strand-break rejoining; Ku autoantigen, 80kDa)	Downregulated
36972	Exônimo	1374	CPT1A	carnitine palmitoyltransferase 1A (liver)	Downregulated
36973	Exônimo	1106	CHD2	chromodomain helicase DNA binding protein 2	Downregulated
36984	Exônimo	285855	RPL7L1	ribosomal protein L7-like 1	Downregulated
36995	Exônimo	85461	-	-	Downregulated
36997	Exônimo	84255	SLC37A3	solute carrier family 37 (glycerol-3-phosphate transporter), member 3	Downregulated
37001	Exônimo	79902	PCNT1	pericentrin 1	Downregulated
37014	Exônimo	55068	-	-	Downregulated
37019	Exônimo	51172	NAGPA	N-acetylglucosamine-1-phosphodiester alpha-N-acetylglucosaminidase	Downregulated
37021	Exônimo	29066	ZC3H7A	zinc finger CCCH-type containing 7A	Downregulated
37034	Exônimo	10776	-	-	Downregulated

37041	Exônimo	9650	-	-	Downregulated
37050	Exônimo	7322	UBE2D2	ubiquitin-conjugating enzyme E2D 2 (UBC4/5 homolog, yeast)	Downregulated
37051	Exônimo	7090	TLE3	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	Downregulated
37073	Exônimo	1717	DHCR7	7-dehydrocholesterol reductase	Downregulated
37089	Exônimo	257160	-	-	Downregulated
37097	Exônimo	113251	-	-	Downregulated
37101	Exônimo	84193	C14orf154	chromosome 14 open reading frame 154	Downregulated
37104	Exônimo	80135	BXDC5	brix domain containing 5	Downregulated
37105	Exônimo	79873	-	-	Downregulated
37107	Exônimo	64968	MRPS6	mitochondrial ribosomal protein S6	Downregulated
37130	Exônimo	23473	CAPN7	calpain 7	Downregulated
37156	Exônimo	6868	ADAM17	a disintegrin and metalloproteinase domain 17 (tumor necrosis factor, alpha, converting enzyme)	Downregulated
37171	Exônimo	3190	HNRPK	heterogeneous nuclear ribonucleoprotein K	Downregulated
37197	Exônimo	129531	-	-	Downregulated
37198	Exônimo	118812	C10orf83	chromosome 10 open reading frame 83	Downregulated
37206	Exônimo	80179	MYOHD1	myosin head domain containing 1	Downregulated
37229	Exônimo	26224	FBXL3	F-box and leucine-rich repeat protein 3	Downregulated
37241	Exônimo	10570	DPYSL4	dihydropyrimidinase-like 4	Downregulated
37243	Exônimo	10198	MPHOSPH9	M-phase phosphoprotein 9	Downregulated
37257	Exônimo	7090	TLE3	transducin-like enhancer of split 3 (E(sp1) homolog, Drosophila)	Downregulated
37260	Exônimo	6502	SKP2	S-phase kinase-associated protein 2 (p45)	Downregulated
37269	Exônimo	4363	ABCC1	ATP-binding cassette, sub-family C (CFTR/MRP), member 1	Downregulated
37273	Exônimo	3276	HRMT1L2	HMT1 hnRNP methyltransferase-like 2 (S. cerevisiae)	Downregulated
37278	Exônimo	1982	EIF4G2	eukaryotic translation initiation factor 4 gamma, 2	Downregulated
37285	Exônimo	204	AK2	adenylate kinase 2	Downregulated
37288	Exônimo				Downregulated
37301	Exônimo	127700	-	-	Downregulated
37303	Exônimo	113251	-	-	Downregulated
37305	Exônimo	85379	-	-	Downregulated
37312	Exônimo	79637	ARMC7	armadillo repeat containing 7	Downregulated
37321	Exônimo	55677	-	-	Downregulated
37350	Exônimo	9757	-	-	Downregulated
37353	Exônimo	9213	XPR1	xenotropic and polytropic retrovirus receptor	Downregulated
37359	Exônimo	7525	YES1	v-yes-1 Yamaguchi sarcoma viral oncogene homolog 1	Downregulated
37377	Exônimo	3190	HNRPK	heterogeneous nuclear ribonucleoprotein K	Downregulated
37383	Exônimo	1653	DDX1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	Downregulated

37389	Exônico	163	AP2B1	adaptor-related protein complex 2, beta 1 subunit	Downregulated
37396	Exônico	285855	RPL7L1	ribosomal protein L7-like 1	Downregulated
37408	Exônico	84722	-	-	Downregulated
37417	Exônico	60561	-	-	Downregulated
37424	Exônico	55536	CDCA7L	cell division cycle associated 7-like	Downregulated
37425	Exônico	55243	KIRREL	kin of IRRE like (Drosophila)	Downregulated
37428	Exônico	54517	-	-	Downregulated
37430	Exônico	51447	IHPK2	inositol hexaphosphate kinase 2	Downregulated
37449	Exônico	10197	PSME3	proteasome (prosome, macropain) activator subunit 3 (PA28 gamma; Ki)	Downregulated
37509	Exônico	113251	-	-	Downregulated
37513	Exônico	84196	USP48	ubiquitin specific protease 48	Downregulated
37519	Exônico	64978	MRPL38	mitochondrial ribosomal protein L38	Downregulated
37532	Exônico	54496	-	-	Downregulated
37542	Exônico	23476	BRD4	bromodomain containing 4	Downregulated
37548	Exônico	11196	SEC23IP	SEC23 interacting protein	Downregulated
37551	Exônico	10533	APG7L	APG7 autophagy 7-like (S. cerevisiae)	Downregulated
37558	Exônico	9453	GGPS1	geranylgeranyl diphosphate synthase 1	Downregulated
37571	Exônico	6124	RPL4	ribosomal protein L4	Downregulated
37573	Exônico	5709	PSMD3	proteasome (prosome, macropain) 26S subunit, non-ATPase, 3	Downregulated
37577	Exônico	4926	NUMA1	nuclear mitotic apparatus protein 1	Downregulated
37594	Exônico	378	ARF4	ADP-ribosylation factor 4	Downregulated
37601	Exônico	388677	NOTCH2NL	Notch homolog 2 (Drosophila) N-terminal like	Downregulated
37619	Exônico	79902	PCNT1	pericentrin 1	Downregulated
37633	Exônico	54840	APTAX	apataxin	Downregulated
37648	Exônico	23042	-	-	Downregulated
37677	Exônico	5347	PLK1	polo-like kinase 1 (Drosophila)	Downregulated
37685	Exônico	3276	HRMT1L2	HMT1 hnRNP methyltransferase-like 2 (S. cerevisiae)	Downregulated
37699	Exônico	3728	JUP	junction plakoglobin	Downregulated
37752	Exônico	23033	KIAA1117	KIAA1117	Downregulated
37757	Exônico	10533	APG7L	APG7 autophagy 7-like (S. cerevisiae)	Downregulated
37761	Exônico	9897	-	-	Downregulated
37772	Exônico	7290	HIRA	HIR histone cell cycle regulation defective homolog A (S. cerevisiae)	Downregulated
37795	Exônico	1653	DDX1	DEAD (Asp-Glu-Ala-Asp) box polypeptide 1	Downregulated
37829	Exônico	60559	SPCS3	signal peptidase complex subunit 3 homolog (S. cerevisiae)	Downregulated
37830	Exônico	57674	C17orf27	chromosome 17 open reading frame 27	Downregulated
37842	Exônico	51444	RNF138	ring finger protein 138	Downregulated
37849	Exônico	24139	EML2	echinoderm microtubule associated protein like 2	Downregulated

37873	Exônico	7555	ZNF9	zinc finger protein 9 (a cellular retroviral nucleic acid binding protein)	Downregulated
37874	Exônico	7320	UBE2B	ubiquitin-conjugating enzyme E2B (RAD6 homolog)	Downregulated
37899	Exônico	1119	CHKA	choline kinase alpha	Downregulated
37905	Exônico	3925	STMN1	stathmin 1/oncoprotein 18	Downregulated
37928	Exônico	80143	-	-	Downregulated
37940	Exônico	55454	-	-	Downregulated
37969	Exônico	9639	ARHGEF10	Rho guanine nucleotide exchange factor (GEF) 10	Downregulated
37973	Exônico	8800	PEX11A	peroxisomal biogenesis factor 11A	Downregulated
37997	Exônico	2734	GLG1	golgi apparatus protein 1	Downregulated
38006	Exônico	384	ARG2	arginase, type II	Downregulated
38013	Exônico	388536	-	-	Downregulated
38051	Exônico	29057	TMEM29	transmembrane protein 29	Downregulated
38089	Exônico	5347	PLK1	polo-like kinase 1 (Drosophila)	Downregulated
38098	Exônico	2975	GTF3C1	general transcription factor IIIC, polypeptide 1, alpha 220kDa	Downregulated
38110	Exônico	16	AARS	alanyl-tRNA synthetase	Downregulated
38111	Exônico	5511	PPP1R8	protein phosphatase 1, regulatory (inhibitor) subunit 8	Downregulated
38125	Exônico	127700	-	-	Downregulated
38150	Exônico	54496	-	-	Downregulated
38169	Exônico	10535	RNASEH2A	ribonuclease H2, large subunit	Downregulated
38171	Exônico	10165	SLC25A13	solute carrier family 25, member 13 (citrin)	Downregulated
38184	Exônico	7290	HIRA	HIR histone cell cycle regulation defective homolog A (S. cerevisiae)	Downregulated
38188	Exônico	6455	SH3GL1	SH3-domain GRB2-like 1	Downregulated
38190	Exônico	5862	RAB2	RAB2, member RAS oncogene family	Downregulated
38192	Exônico	5528	PPP2R5D	protein phosphatase 2, regulatory subunit B (B56), delta isoform	Downregulated
38202	Exônico	2956	MSH6	mutS homolog 6 (E. coli)	Downregulated
38207	Exônico	1654	DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	Downregulated
38225	Exônico	151195	-	-	Downregulated
38241	Exônico	60559	SPCS3	signal peptidase complex subunit 3 homolog (S. cerevisiae)	Downregulated
38258	Exônico	27102	EIF2AK1	eukaryotic translation initiation factor 2-alpha kinase 1	Downregulated
38279	Exônico	9238	TBRG4	transforming growth factor beta regulator 4	Downregulated
38285	Exônico	7553	ZNF7	zinc finger protein 7 (KOX 4, clone HF.16)	Downregulated
38308	Exônico	1982	EIF4G2	eukaryotic translation initiation factor 4 gamma, 2	Downregulated
38309	Exônico	1716	DGUOK	deoxyguanosine kinase	Downregulated
38318	Exônico				Downregulated
38323	Exônico	387338	NSUN4	NOL1/NOP2/Sun domain family, member 4	Downregulated

38346	Exônico	57645	POGK	pogo transposable element with KRAB domain	Downregulated
38352	Exônico	55471	-	-	Downregulated
38354	Exônico	55034	MOCOS	molybdenum cofactor sulfurase	Downregulated
38369	Exônico	23122	CLASP2	cytoplasmic linker associated protein 2	Downregulated
38372	Exônico	11198	SUPT16H	suppressor of Ty 16 homolog (S. cerevisiae)	Downregulated
38379	Exônico	9898	UBAP2L	ubiquitin associated protein 2-like	Downregulated
38385	Exônico	8800	PEX11A	peroxisomal biogenesis factor 11A	Downregulated
38391	Exônico	7082	TJP1	tight junction protein 1 (zonula occludens 1)	Downregulated
38418	Exônico	384	ARG2	arginase, type II	Downregulated
38446	Exônico	64398	MPP5	membrane protein, palmitoylated 5 (MAGUK p55 subfamily member 5)	Downregulated
38464	Exônico	27102	EIF2AK1	eukaryotic translation initiation factor 2-alpha kinase 1	Downregulated
38471	Exônico	23141	-	-	Downregulated
38480	Exônico	10051	SMC4L1	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)	Downregulated
38487	Exônico	8821	INPP4B	inositol polyphosphate-4-phosphatase, type II, 105kDa	Downregulated
38499	Exônico	5728	PTEN	phosphatase and tensin homolog (mutated in multiple advanced cancers 1)	Downregulated
38500	Exônico	5533	PPP3CC	protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcineurin A gamma)	Downregulated
38506	Exônico	4134	MAP4	microtubule-associated protein 4	Downregulated
38513	Exônico	2121	EVC	Ellis van Creveld syndrome	Downregulated
38537	Exônico	128308	MRPL55	mitochondrial ribosomal protein L55	Downregulated
38554	Exônico	57099	AVEN	apoptosis, caspase activation inhibitor	Downregulated
38565	Exônico	51126	NAT5	N-acetyltransferase 5 (ARD1 homolog, S. cerevisiae)	Downregulated
38574	Exônico	23241	PACS2	phosphofurin acidic cluster sorting protein 2	Downregulated
38583	Exônico	10166	SLC25A15	solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 15	Downregulated
38589	Exônico	9218	VAPA	VAMP (vesicle-associated membrane protein)-associated protein A, 33kDa	Downregulated
38595	Exônico	7528	YY1	YY1 transcription factor	Downregulated
38604	Exônico	5529	PPP2R5E	protein phosphatase 2, regulatory subunit B (B56), epsilon isoform	Downregulated
38607	Exônico	4927	NUP88	nucleoporin 88kDa	Downregulated
38626	Exônico				Downregulated
38642	Exônico	91746	-	-	Downregulated

38645	Exônimo	84254	CAMKK1	calcium/calmodulin-dependent protein kinase kinase 1, alpha	Downregulated
38650	Exônimo	79675	-	-	Downregulated
38653	Exônimo	60528	ELAC2	elaC homolog 2 (E. coli)	Downregulated
38663	Exônimo	54839	LRRC49	leucine rich repeat containing 49	Downregulated
38666	Exônimo	51441	YTHDF2	YTH domain family, member 2	Downregulated
38673	Exônimo	24137	KIF4A	kinesin family member 4A	Downregulated
38685	Exônimo	10196	HRMT1L3	HMT1 hnRNP methyltransferase-like 3 (S. cerevisiae)	Downregulated
38715	Exônimo	3268	HRBL	HIV-1 Rev binding protein-like	Downregulated
38720	Exônimo	1981	EIF4G1	eukaryotic translation initiation factor 4 gamma, 1	Downregulated
38789	Exônimo	10166	SLC25A15	solute carrier family 25 (mitochondrial carrier; ornithine transporter) member 15	Downregulated
38800	Exônimo	7994	MYST3	MYST histone acetyltransferase (monocytic leukemia) 3	Downregulated
38818	Exônimo	3655	ITGA6	integrin, alpha 6	Downregulated
38837	Exônimo	388536	-	-	Downregulated
38876	Exônimo	27102	EIF2AK1	eukaryotic translation initiation factor 2-alpha kinase 1	Downregulated
38902	Exônimo	8031	NCOA4	nuclear receptor coactivator 4	Downregulated
38912	Exônimo	5532	PPP3CB	protein phosphatase 3 (formerly 2B), catalytic subunit, beta isoform (calcineurin A beta)	Downregulated
38926	Exônimo	1981	EIF4G1	eukaryotic translation initiation factor 4 gamma, 1	Downregulated
38949	Exônimo	128308	MRPL55	mitochondrial ribosomal protein L55	Downregulated
38969	Exônimo	55677	-	-	Downregulated
38986	Exônimo	23241	PACS2	phosphofurin acidic cluster sorting protein 2	Downregulated
38993	Exônimo	10535	RNASEH2A	ribonuclease H2, large subunit	Downregulated
39026	Exônimo	2956	MSH6	mutS homolog 6 (E. coli)	Downregulated
39054	Exônimo	91694	-	-	Downregulated
39084	Exônimo	25948	KBTBD2	kelch repeat and BTB (POZ) domain containing 2	Downregulated
39107	Exônimo	8451	CUL4A	cullin 4A	Downregulated
39109	Exônimo	7546	ZIC2	Zic family member 2 (odd-paired homolog, Drosophila)	Downregulated
39116	Exônimo	5886	RAD23A	RAD23 homolog A (S. cerevisiae)	Downregulated
39140	Exônimo	14	AAMP	angio-associated, migratory cell protein	Downregulated
39141	Exônimo	63979	FIGNL1	fidgetin-like 1	Downregulated
39149	Exônimo	257218	SHPRH	SNF2 histone linker PHD RING helicase	Downregulated
39159	Exônimo	85440	DOCK7	dedicator of cytokinesis 7	Downregulated
39178	Exônimo	55035	NOL8	nucleolar protein 8	Downregulated
39183	Exônimo	51126	NAT5	N-acetyltransferase 5 (ARD1 homolog, S. cerevisiae)	Downregulated
39185	Exônimo	28982	-	-	Downregulated

39187	Exônico	26146	TRAF3IP1	TNF receptor-associated factor 3 interacting protein 1	Downregulated
39199	Exônico	10535	RNASEH2A	ribonuclease H2, large subunit	Downregulated
39201	Exônico	10171	RCL1	RNA terminal phosphate cyclase-like 1	Downregulated
39202	Exônico	10038	PARP2	poly (ADP-ribose) polymerase family, member 2	Downregulated
39216	Exônico	6871	TADA2L	transcriptional adaptor 2 (ADA2 homolog, yeast)-like	Downregulated
39241	Exônico	636	BICD1	bicaudal D homolog 1 (Drosophila)	Downregulated
39273	Exônico	57504	MTA3	metastasis associated 1 family, member 3	Downregulated
39279	Exônico	55230	USP40	ubiquitin specific protease 40	Downregulated
39290	Exônico	25948	KBTBD2	kelch repeat and BTB (POZ) domain containing 2	Downregulated
39294	Exônico	23248	KIAA0460	KIAA0460	Downregulated
39303	Exônico	10190	TXNDC9	thioredoxin domain containing 9	Downregulated
39307	Exônico	9649	RALGPS1	Ral GEF with PH domain and SH3 binding motif 1	Downregulated
39309	Exônico	9232	PTTG1	pituitary tumor-transforming 1	Downregulated
39318	Exônico	6901	TAZ	tafazzin (cardiomyopathy, dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome)	Downregulated
39328	Exônico	4729	NDUFV2	NADH dehydrogenase (ubiquinone) flavoprotein 2, 24kDa	Downregulated
39338	Exônico	1979	EIF4EBP2	eukaryotic translation initiation factor 4E binding protein 2	Downregulated
39347	Exônico	79089	-	-	Downregulated
39358	Exônico	160518	-	-	Downregulated
39362	Exônico	116988	CENTG3	centaurin, gamma 3	Downregulated
39374	Exônico	64324	NSD1	nuclear receptor binding SET domain protein 1	Downregulated
39395	Exônico	23731	C9orf5	chromosome 9 open reading frame 5	Downregulated
39422	Exônico	6871	TADA2L	transcriptional adaptor 2 (ADA2 homolog, yeast)-like	Downregulated
39438	Exônico	2958	GTF2A2	general transcription factor IIA, 2, 12kDa	Downregulated
39443	Exônico	1654	DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	Downregulated
39456	Exônico	285704	RGMB	RGM domain family, member B	Downregulated
39457	Exônico	259282	FAM44A	family with sequence similarity 44, member A	Downregulated
39463	Exônico	129285	-	-	Downregulated
39467	Exônico	85459	-	-	Downregulated
39472	Exônico	80173	CCDC2	coiled-coil domain containing 2	Downregulated
39489	Exônico	51734	SEPX1	selenoprotein X, 1	Downregulated
39501	Exônico	23140	ZZEF1	zinc finger, ZZ-type with EF hand domain 1	Downregulated
39506	Exônico	10771	ZMYND11	zinc finger, MYND domain containing 11	Downregulated

39510	Exônico	10049	DNAJB6	DnaJ (Hsp40) homolog, subfamily B, member 6	Downregulated
39528	Exônico	5886	RAD23A	RAD23 homolog A (S. cerevisiae)	Downregulated
39554	Exônico				Downregulated
39569	Exônico	113510	-	-	Downregulated
39604	Exônico	23242	COBL	cordon-bleu homolog (mouse)	Downregulated
39608	Exônico	11200	CHEK2	CHK2 checkpoint homolog (S. pombe)	Downregulated
39614	Exônico	10038	PARP2	poly (ADP-ribose) polymerase family, member 2	Downregulated
39627	Exônico	7082	TJP1	tight junction protein 1 (zona occludens 1)	Downregulated
39649	Exônico	1654	DDX3X	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, X-linked	Downregulated
39650	Exônico	1385	CREB1	cAMP responsive element binding protein 1	Downregulated
39665	Exônico	200095	C1orf67	chromosome 1 open reading frame 67	Downregulated
39683	Exônico	60528	ELAC2	elaC homolog 2 (E. coli)	Downregulated
39685	Exônico	57501	KIAA1257	KIAA1257	Downregulated
39698	Exônico	50485	SMARCA1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1	Downregulated
39726	Exônico	8031	NCOA4	nuclear receptor coactivator 4	Downregulated
39730	Exônico	6901	TAZ	tafazzin (cardiomyopathy, dilated 3A (X-linked); endocardial fibroelastosis 2; Barth syndrome)	Downregulated
39755	Exônico	664	BNIP3	BCL2/adenovirus E1B 19kDa interacting protein 3	Downregulated
39824	Exônico	9456	HOMER1	homer homolog 1 (Drosophila)	Downregulated
39843	Exônico	4927	NUP88	nucleoporin 88kDa	Downregulated
39847	Exônico	3858	KRT10	keratin 10 (epidermolytic hyperkeratosis; keratosis palmaris et plantaris)	Downregulated
39883	Exônico	80820	-	-	Downregulated
39903	Exônico	51155	HN1	hematological and neurological expressed 1	Downregulated
39921	Exônico	10190	TXNDC9	thioredoxin domain containing 9	Downregulated
39933	Exônico	7543	ZFX	zinc finger protein, X-linked	Downregulated
39940	Exônico	5885	RAD21	RAD21 homolog (S. pombe)	Downregulated
39945	Exônico	4953	ODC1	ornithine decarboxylase 1	Downregulated
39956	Exônico	1978	EIF4EBP1	eukaryotic translation initiation factor 4E binding protein 1	Downregulated
39965	Exônico	1185	CLCN6	chloride channel 6	Downregulated
39975	Exônico	197131	UBR1	ubiquitin protein ligase E3 component n-recogin 1	Downregulated
39983	Exônico	85440	DOCK7	dedicator of cytokinesis 7	Downregulated
39992	Exônico	64326	RFWD2	ring finger and WD repeat domain 2	Downregulated

40005	Exônico	51704	GPRC5B	G protein-coupled receptor, family C, group 5, member B	Downregulated
40007	Exônico	51127	TRIM17	tripartite motif-containing 17	Downregulated
40025	Exônico	10171	RCL1	RNA terminal phosphate cyclase-like 1	Downregulated
40026	Exônico	10040	TOM1L1	target of myb1-like 1 (chicken)	Downregulated
40027	Exônico	9898	UBAP2L	ubiquitin associated protein 2-like	Downregulated
40035	Exônico	8438	RAD54L	RAD54-like (S. cerevisiae)	Downregulated
40042	Exônico	6462	SHBG	sex hormone-binding globulin	Downregulated
40055	Exônico	3191	HNRPL	heterogeneous nuclear ribonucleoprotein L	Downregulated
40059	Exônico	2115	ETV1	ets variant gene 1	Downregulated
40061	Exônico	1655	DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	Downregulated
40076	Exônico	221785	ZNF498	zinc finger protein 498	Downregulated
40090	Exônico	80169	-	-	Downregulated
40097	Exônico	57501	KIAA1257	KIAA1257	Downregulated
40110	Exônico	50485	SMARCAL1	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a-like 1	Downregulated
40120	Exônico	23041	-	-	Downregulated
40138	Exônico	8030	CCDC6	coiled-coil domain containing 6	Downregulated
40146	Exônico	5884	RAD17	RAD17 homolog (S. pombe)	Downregulated
40161	Exônico	2120	ETV6	ets variant gene 6 (TEL oncogene)	Downregulated
40172	Exônico	648	PCGF4	polycomb group ring finger 4	Downregulated
40181	Exônico	197131	UBR1	ubiquitin protein ligase E3 component n-recognin 1	Downregulated
40192	Exônico	83642	-	-	Downregulated
40193	Exônico	80778	ZNF34	zinc finger protein 34 (KOX 32)	Downregulated
40231	Exônico	10172	ZNF256	zinc finger protein 256	Downregulated
40235	Exônico	9643	MORF4L2	mortality factor 4 like 2	Downregulated
40238	Exônico	9013	TAF1C	TATA box binding protein (TBP)-associated factor, RNA polymerase I, C, 110kDa	Downregulated
40311	Exônico	54834	GDAP2	ganglioside induced differentiation associated protein 2	Downregulated
40315	Exônico	51155	HN1	hematological and neurological expressed 1	Downregulated
40328	Exônico	11222	MRPL3	mitochondrial ribosomal protein L3	Downregulated
40337	Exônico	9648	GCC2	GRIP and coiled-coil domain containing 2	Downregulated
40344	Exônico	8030	CCDC6	coiled-coil domain containing 6	Downregulated
40373	Exônico	659	BMPR2	bone morphogenetic protein receptor, type II (serine/threonine kinase)	Downregulated
40412	Exônico	55486	PSARL	presenilin associated, rhomboid-like	Downregulated
40430	Exônico	23035	PHLPPL	PH domain and leucine rich repeat protein phosphatase like	Downregulated
40435	Exônico	10539	TXNL2	thioredoxin-like 2	Downregulated

40439	Exônico	9898	UBAP2L	ubiquitin associated protein 2-like	Downregulated
40456	Exônico	5868	RAB5A	RAB5A, member RAS oncogene family	Downregulated
40472	Exônico	1938	EEF2	eukaryotic translation elongation factor 2	Downregulated
40480	Exônico				Downregulated
40487	Exônico	259266	ASPM	asp (abnormal spindle)-like, microcephaly associated (Drosophila)	Downregulated
40497	Exônico	85459	-	-	Downregulated
40515	Exônico	55230	USP40	ubiquitin specific protease 40	Downregulated
40527	Exônico	23788	MTCH2	mitochondrial carrier homolog 2 (C. elegans)	Downregulated
40529	Exônico	23347	SMCHD1	structural maintenance of chromosomes flexible hinge domain containing 1	Downregulated
40532	Exônico	23041	-	-	Downregulated
40534	Exônico	11216	AKAP10	A kinase (PRKA) anchor protein 10	Downregulated
40565	Exônico	4351	MPI	mannose phosphate isomerase	Downregulated
40584	Exônico				Downregulated
40608	Exônico	79648	MCPH1	microcephaly, primary autosomal recessive 1	Downregulated
40612	Exônico	57650	KIAA1524	KIAA1524	Downregulated
40634	Exônico	23243	ANKRD28	ankyrin repeat domain 28	Downregulated
40637	Exônico	22878	KIAA1012	KIAA1012	Downregulated
40649	Exônico	9219	MTA2	metastasis associated 1 family, member 2	Downregulated
40665	Exônico	5316	PKNOX1	PBX/knotted 1 homeobox 1	Downregulated
40695	Exônico	200008	-	-	Downregulated
40724	Exônico	54505	DHX29	DEAH (Asp-Glu-Ala-His) box polypeptide 29	Downregulated
40726	Exônico	51439	FAM8A1	family with sequence similarity 8, member A1	Downregulated
40728	Exônico	49855	ZNF291	zinc finger protein 291	Downregulated
40729	Exônico	29028	ATAD2	ATPase family, AAA domain containing 2	Downregulated
40735	Exônico	23347	SMCHD1	structural maintenance of chromosomes flexible hinge domain containing 1	Downregulated
40758	Exônico	7319	UBE2A	ubiquitin-conjugating enzyme E2A (RAD6 homolog)	Downregulated
40762	Exônico	6497	SKI	v-ski sarcoma viral oncogene homolog (avian)	Downregulated
40780	Exônico	1977	EIF4E	eukaryotic translation initiation factor 4E	Downregulated
40789	Exônico	7428	VHL	von Hippel-Lindau tumor suppressor	Downregulated
40816	Exônico	64328	XPO4	exportin 4	Downregulated
40838	Exônico	23476	BRD4	bromodomain containing 4	Downregulated
40880	Exônico	2965	GTF2H1	general transcription factor IIH, polypeptide 1, 62kDa	Downregulated
40885	Exônico	1655	DDX5	DEAD (Asp-Glu-Ala-Asp) box polypeptide 5	Downregulated
40905	Exônico	128869	CDC91L1	CDC91 cell division cycle 91-like 1 (S. cerevisiae)	Downregulated
40913	Exônico	80820	-	-	Downregulated

40925	Exônico	55689	YEATS2	YEATS domain containing 2	Downregulated
40959	Exônico	8819	SAP30	sin3-associated polypeptide, 30kDa	Downregulated
40966	Exônico	6899	TBX1	T-box 1	Downregulated
40969	Exônico	6183	MRPS12	mitochondrial ribosomal protein S12	Downregulated
40972	Exônico	5531	PPP4C	protein phosphatase 4 (formerly X), catalytic subunit	Downregulated
41027	Exônico	56243	KIAA1217	KIAA1217	Downregulated
41043	Exônico	23760	PITPNB	phosphatidylinositol transfer protein, beta	Downregulated
41050	Exônico	11214	AKAP13	A kinase (PRKA) anchor protein 13	Downregulated
41061	Exônico	9221	NOLC1	nucleolar and coiled-body phosphoprotein 1	Downregulated
41067	Exônico	7528	YY1	YY1 transcription factor	Downregulated
41070	Exônico	6874	TAF4	TAF4 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 135kDa	Downregulated
41079	Exônico	4927	NUP88	nucleoporin 88kDa	Downregulated
41123	Exônico	65117	-	-	Downregulated
41133	Exônico	55230	USP40	ubiquitin specific protease 40	Downregulated
41135	Exônico	54834	GDAP2	ganglioside induced differentiation associated protein 2	Downregulated
41136	Exônico	54505	DHX29	DEAH (Asp-Glu-Ala-His) box polypeptide 29	Downregulated
41151	Exônico	22891	ZNF365	zinc finger protein 365	Downregulated
41201	Exônico	11236	RNF139	ring finger protein 139	Downregulated
41202	Exônico				Downregulated
41234	Exônico	55802	DCP1A	decapping enzyme	Downregulated
41242	Exônico	51426	POLK	polymerase (DNA directed) kappa	Downregulated
41244	Exônico	30827	CXXC1	CXXC finger 1 (PHD domain)	Downregulated
41264	Exônico	9765	ZFYVE16	zinc finger, FYVE domain containing 16	Downregulated
41266	Exônico	9462	RASAL2	RAS protein activator like 2	Downregulated
41282	Exônico	5529	PPP2R5E	protein phosphatase 2, regulatory subunit B (B56), epsilon isoform	Downregulated
41292	Exônico	2968	GTF2H4	general transcription factor IIH, polypeptide 4, 52kDa	Downregulated
41301	Exônico	641	BLM	Bloom syndrome	Downregulated
41319	Exônico	114294	LACTB	lactamase, beta	Downregulated
41332	Exônico	57669	EPB41L5	erythrocyte membrane protein band 4.1 like 5	Downregulated
41340	Exônico	55057	AIM1L	absent in melanoma 1-like	Downregulated
41358	Exônico	11215	AKAP11	A kinase (PRKA) anchor protein 11	Downregulated
41366	Exônico	9774	BCLAF1	BCL2-associated transcription factor 1	Downregulated
41371	Exônico	8819	SAP30	sin3-associated polypeptide, 30kDa	Downregulated
41384	Exônico	5531	PPP4C	protein phosphatase 4 (formerly X), catalytic subunit	Downregulated
41422	Exônico	117143	-	-	Downregulated
41432	Exônico	79651	RHBDL6	rhomboid, veinlet-like 6 (Drosophila)	Downregulated
41477	Exônico	8438	RAD54L	RAD54-like (S. cerevisiae)	Downregulated

41479	Exônico	7529	YWHAB	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, beta polypeptide	Downregulated
41491	Exônico	4927	NUP88	nucleoporin 88kDa	Downregulated
41506	Exônico	833	CARS	cysteinyl-tRNA synthetase	Downregulated
41529	Exônico	84246	-	-	Downregulated
41544	Exônico	55508	SLC35E3	solute carrier family 35, member E3	Downregulated
41545	Exônico	55226	-	-	Downregulated
41553	Exônico	29028	ATAD2	ATPase family, AAA domain containing 2 SMC5 structural maintenance of chromosomes 5-like 1 (yeast)	Downregulated
41561	Exônico	23137	SMC5L1	BCL2-associated transcription factor 1	Downregulated
41572	Exônico	9774	BCLAF1	zinc finger protein 161 homolog (mouse)	Downregulated
41581	Exônico	7541	ZFP161	threonyl-tRNA synthetase	Downregulated
41584	Exônico	6897	TARS	DEAH (Asp-Glu-Ala-His) box polypeptide 15	Downregulated
41605	Exônico	1665	DHX15	bone morphogenetic protein receptor, type IA	Downregulated
41609	Exônico	657	BMPRI1A	-	Downregulated
41614	Exônico	-	-	-	Downregulated
41633	Exônico	84224	-	-	Downregulated
41656	Exônico	30827	CXXC1	CXXC finger 1 (PHD domain)	Downregulated
41657	Exônico	28990	-	-	Downregulated
41668	Exônico	11214	AKAP13	A kinase (PRKA) anchor protein 13	Downregulated
41679	Exônico	9221	NOLC1	nucleolar and coiled-body phosphoprotein 1	Downregulated
41680	Exônico	9014	TAF1B	TATA box binding protein (TBP)-associated factor, RNA polymerase I, B, 63kDa	Downregulated
41687	Exônico	7082	TJP1	tight junction protein 1 (zonula occludens 1)	Downregulated
41690	Exônico	6477	SIAH1	seven in absentia homolog 1 (Drosophila)	Downregulated
41704	Exônico	2968	GTF2H4	general transcription factor IIH, polypeptide 4, 52kDa	Downregulated
41716	Exônico	-	-	-	Downregulated
41735	Exônico	84246	-	-	Downregulated
41743	Exônico	60509	-	-	Downregulated
41753	Exônico	54832	VPS13C	vacuolar protein sorting 13C (yeast)	Downregulated
41755	Exônico	51719	CAB39	calcium binding protein 39	Downregulated
41757	Exônico	51143	DNCL11	dynein, cytoplasmic, light intermediate polypeptide 1	Downregulated
41791	Exônico	6687	SPG7	spastic paraplegia 7, paraplegin (pure and complicated autosomal recessive)	Downregulated
41796	Exônico	5530	PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)	Downregulated
41856	Exônico	55037	-	-	Downregulated
41860	Exônico	51426	POLK	polymerase (DNA directed) kappa	Downregulated

41870	Exônico	23243	ANKRD28	ankyrin repeat domain 28	Downregulated
41882	Exônico	9765	ZFYVE16	zinc finger, FYVE domain containing 16	Downregulated
41889	Exônico	8439	NSMAF	neutral sphingomyelinase (N-SMase) activation associated factor	Downregulated
41900	Exônico	5530	PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)	Downregulated
41907	Exônico	3875	KRT18	keratin 18	Downregulated
41960	Exônico	54505	DHX29	DEAH (Asp-Glu-Ala-His) box polypeptide 29	Downregulated
41962	Exônico	51434	ANAPC7	anaphase promoting complex subunit 7	Downregulated
41981	Exônico	10181	RBM5	RNA binding motif protein 5	Downregulated
41984	Exônico	9774	BCLAF1	BCL2-associated transcription factor 1	Downregulated
41986	Exônico	9474	APG5L	APG5 autophagy 5-like (S. cerevisiae)	Downregulated
41991	Exônico	8450	CUL4B	cullin 4B	Downregulated
42022	Exônico	406	ARNTL	aryl hydrocarbon receptor nuclear translocator-like	Downregulated
42061	Exônico	55212	BBS7	Bardet-Biedl syndrome 7	Downregulated
42068	Exônico	30833	NT5C	5', 3'-nucleotidase, cytosolic	Downregulated
42124	Exônico	833	CARS	cysteinyl-tRNA synthetase	Downregulated
42125	Exônico	641	BLM	Bloom syndrome	Downregulated
42135	Exônico	259230	TMEM23	transmembrane protein 23	Downregulated
42152	Exônico	79668	PARP8	poly (ADP-ribose) polymerase family, member 8	Downregulated
42155	Exônico	60492	-	-	Downregulated
42178	Exônico	23245	ASTN2	astrotactin 2	Downregulated
42195	Exônico	8816	WDR22	WD repeat domain 22	Downregulated
42197	Exônico	8450	CUL4B	cullin 4B	Downregulated
42212	Exônico	4720	NDUFS2	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49kDa (NADH-coenzyme Q reductase)	Downregulated
42231	Exônico	160760	-	-	Downregulated
42257	Exônico	65056	-	-	Downregulated
42292	Exônico	10042	HMG2L1	high-mobility group protein 2-like 1	Downregulated
42301	Exônico	8439	NSMAF	neutral sphingomyelinase (N-SMase) activation associated factor	Downregulated
42311	Exônico	5718	PSMD12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	Downregulated
42339	Exônico	387921	-	-	Downregulated
42367	Exônico	55689	YEATS2	YEATS domain containing 2	Downregulated
42373	Exônico	51715	RAB23	RAB23, member RAS oncogene family	Downregulated
42375	Exônico	51141	INSIG2	insulin induced gene 2	Downregulated
42379	Exônico	26190	FBXW2	F-box and WD-40 domain protein 2	Downregulated
42385	Exônico	23137	SMC5L1	SMC5 structural maintenance of chromosomes 5-like 1 (yeast)	Downregulated

42390	Exônico	10768	AHCYL1	S-adenosylhomocysteine hydrolase-like 1	Downregulated
42401	Exônico	8816	WDR22	WD repeat domain 22	Downregulated
42438	Exônico	7421	VDR	vitamin D (1,25-dihydroxyvitamin D3) receptor	Downregulated
42464	Exônico	64343	AZI2	5-azacytidine induced 2	Downregulated
42465	Exônico	60481	ELOVL5	ELOVL family member 5, elongation of long chain fatty acids (FEN1/Elo2, SUR4/Elo3-like, yeast)	Downregulated
42469	Exônico	56259	CTNBL1	catenin, beta like 1	Downregulated
42480	Exônico	30833	NT5C	5', 3'-nucleotidase, cytosolic	Downregulated
42484	Exônico	25938	C14orf125	chromosome 14 open reading frame 125	Downregulated
42489	Exônico	23131	-	-	Downregulated
42500	Exônico	9766	KIAA0247	KIAA0247	Downregulated
42513	Exônico	6675	UAP1	UDP-N-acetylglucosamine pyrophosphorylase 1	Downregulated
42518	Exônico	5530	PPP3CA	protein phosphatase 3 (formerly 2B), catalytic subunit, alpha isoform (calcineurin A alpha)	Downregulated
42521	Exônico	4928	NUP98	nucleoporin 98kDa	Downregulated
42535	Exônico	1108	CHD4	chromodomain helicase DNA binding protein 4	Downregulated
42567	Exônico	60492	-	-	Downregulated
42578	Exônico	54505	DHX29	DEAH (Asp-Glu-Ala-His) box polypeptide 29	Downregulated
42593	Exônico	22889	-	-	Downregulated
42618	Exônico	5877	RABIF	RAB interacting factor	Downregulated
42621	Exônico	5335	PLCG1	phospholipase C, gamma 1	Downregulated
42625	Exônico	4331	MNAT1	menage a trois 1 (CAK assembly factor)	Downregulated
42634	Exônico	1974	EIF4A2	eukaryotic translation initiation factor 4A, isoform 2	Downregulated
42635	Exônico	1662	DDX10	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	Downregulated
42655	Exônico	150274	-	-	Downregulated
42697	Exônico	22884	WDR37	WD repeat domain 37	Downregulated
42717	Exônico	7086	TKT	transketolase (Wernicke-Korsakoff syndrome)	Downregulated
42761	Exônico	114294	LACTB	lactamase, beta	Downregulated
42771	Exônico	65109	UPF3B	UPF3 regulator of nonsense transcripts homolog B (yeast)	Downregulated
42779	Exônico	55687	TRMT1	tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase 1	Downregulated
42791	Exônico	26190	FBXW2	F-box and WD-40 domain protein 2	Downregulated
42876	Exônico	64343	AZI2	5-azacytidine induced 2	Downregulated
42883	Exônico	55681	SCYL2	SCY1-like 2 (S. cerevisiae)	Downregulated
42893	Exônico	28991	COMMD5	COMM domain containing 5	Downregulated
42895	Exônico	26152	ZNF337	zinc finger protein 337	Downregulated
42929	Exônico	5718	PSMD12	proteasome (prosome, macropain) 26S subunit, non-ATPase, 12	Downregulated
42947	Exônico	1108	CHD4	chromodomain helicase DNA binding protein 4	Downregulated
42993	Exônico	51141	INSIG2	insulin induced gene 2	Downregulated

43004	Exônico	23039	XPO7	exportin 7	Downregulated
43036	Exônico	4720	NDUFS2	NADH dehydrogenase (ubiquinone) Fe-S protein 2, 49kDa (NADH-coenzyme Q reductase)	Downregulated
43090	Exônico	55503	TRPV6	transient receptor potential cation channel, subfamily V, member 6	Downregulated
43183	Exônico	65109	UPF3B	UPF3 regulator of nonsense transcripts homolog B (yeast)	Downregulated
43203	Exônico	26168	SEN3	SUMO1/sentrin/SMT3 specific protease 3	Downregulated
43206	Exônico	23484	LEPROTL1	leptin receptor overlapping transcript-like 1	Downregulated
43220	Exônico	9774	BCLAF1	BCL2-associated transcription factor 1	Downregulated
43234	Exônico	6484	ST3GAL4	ST3 beta-galactoside alpha 2,3-sialyltransferase 4	Downregulated
43253	Exônico	1660	DHX9	DEAH (Asp-Glu-Ala-His) box polypeptide 9	Downregulated
43259	Exônico	191	AHCY	S-adenosylhomocysteine hydrolase	Downregulated
43262	Exônico	84233	-	-	Downregulated
43281	Exônico				Downregulated
43329	Exônico	8813	DPM1	dolichyl-phosphate mannosyltransferase polypeptide 1, catalytic subunit	Downregulated
43360	Exônico	837	CASP4	caspase 4, apoptosis-related cysteine protease	Downregulated
43378	Exônico	117247	SLC16A10	solute carrier family 16 (monocarboxylic acid transporters), member 10	Downregulated
43383	Exônico	84236	-	-	Downregulated
43400	Exônico	55051	C14orf102	chromosome 14 open reading frame 102	Downregulated
43409	Exônico	26168	SEN3	SUMO1/sentrin/SMT3 specific protease 3	Downregulated
43419	Exônico	11020	RABL4	RAB, member of RAS oncogene family-like 4	Downregulated
43425	Exônico	9908	-	-	Downregulated
43432	Exônico	8614	STC2	stanniocalcin 2	Downregulated
43445	Exônico	5332	PLCB4	phospholipase C, beta 4	Downregulated
43465	Exônico	182	JAG1	jagged 1 (Alagille syndrome)	Downregulated
43500	Exônico	55809	TRERF1	transcriptional regulating factor 1	Downregulated
43510	Exônico	30844	EHD4	EH-domain containing 4	Downregulated
43516	Exônico	23484	LEPROTL1	leptin receptor overlapping transcript-like 1	Downregulated
43524	Exônico	10767	HBS1L	HBS1-like (S. cerevisiae)	Downregulated
43551	Exônico	4928	NUP98	nucleoporin 98kDa	Downregulated
43578	Exônico	221496	LEMD2	LEM domain containing 2	Downregulated
43599	Exônico	57492	ARID1B	AT rich interactive domain 1B (SWI1-like)	Downregulated
43605	Exônico	55215	-	-	Downregulated
43608	Exônico	54503	ZDHHC13	zinc finger, DHHC-type containing 13	Downregulated
43609	Exônico	51710	ZNF44	zinc finger protein 44 (KOX 7)	Downregulated
43631	Exônico	9908	-	-	Downregulated

43648	Exônico	5876	RABGGTB	Rab geranylgeranyltransferase, beta subunit	Downregulated
43683	Exônico	199745	THAP8	THAP domain containing 8	Downregulated
43691	Exônico	85456	TNKS1BP1	tankyrase 1 binding protein 1, 182kDa	Downregulated
43693	Exônico	84233	-	-	Downregulated
43694	Exônico	83667	SESN2	sestrin 2	Downregulated
43701	Exônico	60485	SAV1	salvador homolog 1 (Drosophila)	Downregulated
43704	Exônico	57122	NUP107	nucleoporin 107kDa	Downregulated
43710	Exônico	55041	PLEKHB2	pleckstrin homology domain containing, family B (evectins) member 2	Downregulated
43738	Exônico	9470	EIF4E2	eukaryotic translation initiation factor 4E member 2	Downregulated
43740	Exônico	9031	BAZ1B	bromodomain adjacent to zinc finger domain, 1B	Downregulated
43744	Exônico	8021	NUP214	nucleoporin 214kDa	Downregulated
43762	Exônico	3661	IRF3	interferon regulatory factor 3	Downregulated
43810	Exônico	55504	TNFRSF19	tumor necrosis factor receptor superfamily, member 19	Downregulated
43818	Exônico	30849	PIK3R4	phosphoinositide-3-kinase, regulatory subunit 4, p150	Downregulated
43821	Exônico	26168	SENP3	SUMO1/sentrin/SMT3 specific protease 3	Downregulated
43848	Exônico	7301	TYRO3	TYRO3 protein tyrosine kinase	Downregulated
43850	Exônico	6885	MAP3K7	mitogen-activated protein kinase kinase kinase 7	Downregulated
43859	Exônico	4929	NR4A2	nuclear receptor subfamily 4, group A, member 2	Downregulated
43910	Exônico	57122	NUP107	nucleoporin 107kDa	Downregulated
43921	Exônico	51133	KCTD3	potassium channel tetramerisation domain containing 3	Downregulated
43922	Exônico	30849	PIK3R4	phosphoinositide-3-kinase, regulatory subunit 4, p150	Downregulated
44000	Exônico	84669	USP32	ubiquitin specific protease 32	Downregulated
44009	Exônico	60487	C6orf75	chromosome 6 open reading frame 75	Downregulated
44014	Exônico	55813	C17orf40	chromosome 17 open reading frame 40	Downregulated
44018	Exônico	55048	-	-	Downregulated
44037	Exônico	11020	RABL4	RAB, member of RAS oncogene family-like 4	Downregulated
44043	Exônico	9908	-	-	Downregulated
44059	Exônico	6137	RPL13	ribosomal protein L13	Downregulated
44094	Exônico	221477	C6orf89	chromosome 6 open reading frame 89	Downregulated
44113	Exônico	60487	C6orf75	chromosome 6 open reading frame 75	Downregulated
44141	Exônico	11019	LIAS	lipoic acid synthetase	Downregulated
44147	Exônico	9904	RBM19	RNA binding motif protein 19	Downregulated
44174	Exônico	3661	IRF3	interferon regulatory factor 3	Downregulated
44182	Exônico	1387	CREBBP	CREB binding protein (Rubinstein-Taybi syndrome)	Downregulated
44221	Exônico	55683	-	-	Downregulated

44223	Exônico	55215	-	-	Downregulated
44238	Exônico	23244	-	-	Downregulated
44239	Exônico	23133	PHF8	PHD finger protein 8	Downregulated
44256	Exônico	8613	PPAP2B	phosphatidic acid phosphatase type 2B	Downregulated
44257	Exônico	8443	GNPAT	glyceronephosphate O- acyltransferase	Downregulated
44277	Exônico	3216	HOXB6	homeo box B6	Downregulated

Rodrigo Louro

Detalhes Pessoais

Nacionalidade: Brasileira

Data do nascimento: 06/01/1978

Estado Civil: solteiro

Endereço: Rua João Miguel Jarra, 75 - apto 14 – Vila Madalena

CEP: 05417-040 - São Paulo – Brasil

Telefone: +55-11-8155-3477

Detalhes Profissionais

Empregador atual: FAPESP – Fundação de Amparo a Pesquisa do Estado de São Paulo

Supervisor: Prof. Dr. Sergio Verjovski-Almeida

Endereço: Avenida Professor Lineu Prestes, 748 – Lab. 1200

CEP: 05508-900 - São Paulo – Brasil

Telefone: +55-11-3091-2173

Fax: +55-11-3091-21-86

e-mail: rolouro@iq.usp.br

Línguas

- Português – Língua Nativa
- Inglês – Avançado. Nota geral no teste de proficiência IELTS – 7,0 (10/2005)
- Francês - Básico

Pós-graduação

- 02/2002 – em andamento: Doutorado pelo Programa de Pós-graduação em Bioquímica – Instituto de Química da Universidade de São Paulo – Brasil.

Projeto: “Identificação e caracterização de transcritos humanos: novas famílias de pequenas GTPases e novos Longos RNAs Intrônicos não-codificantes”

Orientador: Prof. Dr. Sergio Verjovski-Almeida

Graduação

- 2001: Bacharelado em Ciências Biológicas, Instituto de Biociências da Universidade de São Paulo – Brasil.
- 2005: Licenciatura em Ciências Biológicas, Instituto de Biociências da Universidade de São Paulo – Brasil.

Projetos de Pesquisa desenvolvidos durante a graduação

- 08/1999 – 12/ 2001: Iniciação Científica no Instituto de Química, Departamento de Bioquímica, da Universidade de São Paulo – Brasil.

Projeto: “Expressão e caracterização de mutantes pontuais de triptofano da Ca^{2+} - ATPase do retículo sarcoplasmático”.

Metodologia: Expressão heteróloga em levedura, purificação de proteínas por cromatografia de troca iônica e ensaios enzimáticos.

Orientador: Prof. Dr. Sergio Verjovski-Almeida e Dr. Eduardo Moraes Rego Reis.

- 07/1998 – 07/1999: Iniciação Científica no Instituto de Química, Departamento de Bioquímica, da Universidade de São Paulo – Brasil.
- **Projeto:** “Construção de mutantes pontuais de triptofano da Ca^{2+} - ATPase do retículo sarcoplasmático por PCR mutagênico”.

Metodologia: Mutagênese sítio dirigida por PCR e procedimentos gerais de clonagem.

Orientador: Prof. Dr. Sergio Verjovski-Almeida e Dr. Eduardo Moraes Rego Reis

Bolsas de Pesquisa Concedidas

- Início programado para 12/2006: FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo). Bolsa de Pós-Doutorado (Processo nº 06/55889-9).
- 05/2002 – 08/2006: FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo). Bolsa de Doutorado Direto (Processo nº 02/02950-1).
- 02/2002 – 04/2002: CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior). Bolsa de Doutorado Direto concedida pelo desempenho na prova de ingresso ao programa de Pós-graduação em Bioquímica.
- 08/1999 – 12/2001 FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo). Bolsa de Iniciação Científica (Processo nº 99/07509-7).
- 07/1998 – 07/1999: CNPQ (Conselho Nacional de Desenvolvimento Científico e Tecnológico). Bolsa de Iniciação Científica.

Atividades Relacionadas ao Ensino

- 03/2002 – 07/2006 PAE (Programa de Aperfeiçoamento ao Ensino) – Monitor de aulas práticas e teóricas de Bioquímica e Biologia Molecular do Instituto de Química da Universidade de São Paulo – Brasil.

Publicações

- Louro R, Nakaya HI, Paquola AC, Martins EA, da Silva AM, Verjovski-Almeida S, Reis EM: **RASL11A, member of a novel small monomeric GTPase gene family, is down-regulated in prostate tumors.** *Biochem Biophys Res Commun* 2004, **316**(3):618-627.

- Reis EM, Nakaya HI, Louro R, Canavez FC, Flatschart AV, Almeida GT, Egidio CM, Paquola AC, Machado AA, Festa F *et al*: **Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer**. *Oncogene* 2004, **23**(39):6684-6692.
- Reis EM, Louro R, Nakaya HI, Verjovski-Almeida S: **As antisense RNA gets intronic**. *Omics* 2005, **9**(1):2-12.
- Louro R, Nakaya HI, Amaral PP, Festa F, Sogayar MC, da Silva AM, Verjovski-Almeida S, Reis EM: **Androgen responsive intronic non-coding RNAs**. *BMC Genomics* 2006 (**Submitted**).
- Nakaya HI, Amaral PP, Louro R, Lopes A, Fachel AA, Moreira YB, El-Jundi TA, da Silva AM, Verjovski-Almeida S, Reis EM: **Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription** *Genome Biol* 2006 (**Submitted**).