UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

Programa de Pós Graduação em Ciências Biológicas (Bioquímica)

Felipe Cesar Ferrarezi Beckedorff

Recrutamento do complexo repressivo *polycomb* 2 pelo RNA não codificador longo antissenso *ANRASSF1* modula a expressão do gene *RASSF1A* e a proliferação celular

Versão corrigida da Tese

SÃO PAULO

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

Beckedorff, Felipe Cesar Ferrarezi
B394r Recrutamento do complexo repressivo *polymcob* 2 pelo RNA não codificador longo antissenso *ANRASSF1* modula a expressão do gene *RASSF1A* e a proliferação celular / Felipe Cesar Ferrarezi Beckedorff. -- São Paulo, 2012.
93p.

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

1. Regulação gênica I. T. II Verjovski-Almeida, Sergio, orientador.

574.88 CDD

Felipe Cesar Ferrarezi Beckedorff

Recrutamento do complexo repressivo *polycomb* 2 pelo RNA não codificador longo antissenso *ANRASSF1* modula a expressão do gene *RASSF1A* e a proliferação celular

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

Área de Concentração: Bioquímica
Orientador: Dr. Sergio Verjovski-Almeida

São Paulo
2012

Felipe Cesar Ferrarezi Beckedorff

Recrutamento do complexo repressivo *polycomb* 2 pelo RNA não codificador longo antissenso *ANRASSF1* modula a expressão do gene *RASSF1A* e a proliferação celular

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

Aprovado em: _____

Banca Examinadora

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

Prof. Dr. _____

Instituição: _____

Assinatura: _____

AGRADECIMENTOS

Gostaria de agradecer, primeiramente, ao meu Deus (Jeová Jiré), que sempre esteve ao meu lado e nunca me desamparou e sempre me inspirou nos momentos mais difíceis.

Aos meus pais, Irzo e Celia, que sempre me ajudaram, apoiaram e incentivaram, nessa longa jornada do doutorado. Ganhei a aposta, vou conseguir o título primeiro !!!

Ao meu irmão, Frederico pela amizade constante.

A minha esposa, Ana Carolina, pelo companheirismo, paciência, suporte em todos os momentos, leitura e sugestões dos meus trabalhos. Você foi essencial para eu ter chegado até aqui.

Ao Antonio e a Angela, Patrícia e Armando, Reginaldo e Cristiane, Júnior e Adrina pelas orações e palavras de apoio.

Ao Dr. Sergio Verjovski-Almeida, pela orientação, críticas construtivas, ensinamentos e principalmente, pelas oportunidades e confiança deposita durante meu período de doutoramento.

Ao Dr. Eduardo Reis, pela participação durante a execução do meu projeto de doutorado, com críticas e sugestões.

Aos amigos de café, Fábio (Perere) e Thiago Venâncio, pelo companheirismo e bons momentos partilhados.

Ao amigo Helder Nakaya, por me receber aqui em São Paulo pelas inúmeras ajudas ao longo do trabalho, além claro pela amizade, futebol churrascos e cafés.

Aos amigos de república, Eliezer, Vinicius, Perere, Gianlucca e Valdir pelos bons momentos e risadas compartilhadas.

Ao amigo Murilo, pelos duros momentos compartilhados durante a padronização do sequenciamento com o 454, além dos bons momentos de convívio e pelas críticas e sugestões ao longo do meu trabalho.

A amiga Giulliana, pelo ombro amigo e pelas risadas de apoquentos.

Aos amigos de laboratório, Carlos, Lauren, Ana Tahira, Letícia, Santiago, Kleber, Dinar, Ângela, Katia, Mariana, Otto, Bianca, Esther e Jefferson por estarem sempre dispostos a ajudar, por tornarem o ambiente de trabalho tão prazeroso e pelos inúmeros bandeijões.

Ao, Yuri, Vinícius, Rodrigo por toda ajuda nas análises em bioinformática e em especial ao Vinícius para a introdução a POG e todo o companheirismo.

A Ana Paula, minha mãezona, que inúmeras vezes tornou minha vida muito mais fácil e até, algumas vezes, mais doce.

Ao Renan Crocci Souza, pela colaboração durante um período do desenvolvimento desta tese, em especial do trabalho RASSF1.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo suporte financeiro.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo suporte financeiro.

RESUMO

Beckedorff, F. C. F. **Recrutamento do complexo repressivo *polycomb* 2 pelo RNA não codificador longo antissenso *ANRASSF1* modula a expressão do gene *RASSF1A* e a proliferação celular.** 2012. 93 p. Tese de Doutorado - Programa de Pós-Graduação em Bioquímica. Instituto de Química, Universidade de São Paulo, São Paulo.

O gene supressor tumoral *RASSF1A* tem sido associado com redução da proliferação celular em diversos tumores. Sua expressão é regulada por eventos epigenéticos que envolvem o complexo repressivo *polycomb* (PRC2), no entanto os mecanismos moleculares da modulação do recrutamento deste modificador epigenético para este *locus* ainda são desconhecidos. Neste trabalho identificamos e caracterizamos *ANRASSF1*, um RNA não codificador longo (lncRNA) intrônico *unspliced*, que é transcrito na fita oposta do gene *RASSF1A*, em várias linhagem celulares e tecidos, e se liga a PRC2. *ANRASSF1* é transcrito pela RNAPII, possui cap-5´ e cauda poli-A, além de localizar-se no núcleo e possuir uma meia-vida em média quatro vezes menor comparada com outros lncRNAs ligados à PRC2. A super-expressão ectópica de *ANRASSF1* reduziu os níveis de *RASSF1A* e aumentou a taxa de proliferação em células HeLa, enquanto seu silenciamento provocou efeito oposto. Essas mudanças nos níveis de *ANRASSF1* não afetaram a abundância da isoforma *RASSF1C* em nenhuma das condições. A super-expressão de *ANRASSF1* provocou um grande aumento tanto da ocupação de PRC2 como da marca de histona repressiva H3K27me3 especificamente na região promotora *RASSF1A*. Nenhum efeito da super-expressão de *ANRASSF1* foi detectado na ocupação de PRC2 e na histona H3K27me3 nas regiões promotoras de *RASSF1C* e de outros quatro genes vizinhos, incluindo dois genes supressores tumorais bem caracterizados. Além disso, foi demonstrado que *ANRASSF1* forma um híbrido de RNA/DNA e recruta SUZ12, um componente do PRC2, para o promotor de *RASSF1A*. Notavelmente, foi detectado pelo ensaio de RNase-ChIP que a degradação de *ANRASSF1* diminui a ocupação de PRC2 neste promotor. Esses resultados demonstram um novo mecanismo de repressão epigenética do supressor tumoral *RASSF1A*, envolvendo um lncRNA *unspliced* antissenso, onde *ANRASSF1* reprime seletivamente a expressão da isoforma de *RASSF1* que sobrepõe o transcrito antissenso de modo local e específico. Considerando uma perspectiva mais ampla, nossos resultados sugerem que outros lncRNAs intrônicos *unspliced* não caracterizados no genoma humano podem contribuir para uma modulação epigenética local e específica de cada região em que os lncRNAs são transcritos.

Palavras-chaves: RNA não codificador longo antissenso, PRC2, *RASSF1A*, híbrido RNA-DNA, epigenética.

ABSTRACT

Beckedorff, F. C. F. **Recruitment of polycomb repressive complex 2 by intronic long noncoding RNA *ANRASSF1* modulates *RASSF1A* expression and cell proliferation.** 2012. 93 p. PhD Thesis – Graduate Program in Biochemistry. Instituto de Química, Universidade de São Paulo, São Paulo.

Tumor-suppressor RASSF1A gene down-regulation has been implicated in increasing cell proliferation in several tumors. Its expression is regulated by epigenetic events involving polycomb repressive complex 2 (PRC2), however the molecular mechanisms modulating recruitment of this epigenetic modifier to the locus remain largely unknown. Here, we identify and characterize *ANRASSF1*, an endogenous unspliced long noncoding RNA (lncRNA) that is transcribed from the opposite strand of *RASSF1* gene in several cell lines and tissues, and binds to PRC2. *ANRASSF1* is transcribed by RNA Polymerase II, 5'-capped, polyadenylated, displays nuclear localization, and has on average a four-fold shorter half-life compared to other lncRNAs that bind PRC2. *ANRASSF1* ectopic overexpression decreases *RASSF1A* abundance and increases the proliferation rate of HeLa cells, whereas its silencing causes opposite effects. These changes in *ANRASSF1* levels do not affect *RASSF1C* isoform abundance. *ANRASSF1* overexpression causes a marked increase both in PRC2 occupancy and in histone H3K27me3 repressive mark specifically at the *RASSF1A* promoter region. No effect of *ANRASSF1* overexpression is detected on PRC2 occupancy and on histone H3K27me3 at the promoter regions of *RASSF1C* and of four other neighbor genes, including two well-characterized tumor suppressor genes. Additionally, we demonstrate that *ANRASSF1* forms an RNA/DNA hybrid, and recruits SUZ12, a PRC2 component, to the *RASSF1A* promoter. Notably, depletion of *ANRASSF1* disrupts SUZ12 occupancy on *RASSF1A* promoter as measured by RNAse-ChIP assay. Together, these results show a new mechanism of epigenetic repression of *RASSF1A* tumor suppressor gene involving an antisense unspliced lncRNA, in which *ANRASSF1* selectively represses expression of the *RASSF1* isoform overlapping the antisense transcript in a location-specific manner. In a broader perspective, our findings suggest that other non-characterized unspliced intronic lncRNAs transcribed in the human genome may contribute to a location-specific epigenetic modulation of genes.

Keywords: antisense unspliced long noncoding RNA, PRC2; *RASSF1A*, RNA-DNA hybrid, epigenetics

Lista de abreviaturas e siglas

ANRASSF1:	*Antisense Noncoding RNA RASSF1*
Cap 5':	7-metilguanosina na extremidade 5' de RNAs
cDNA:	DNA complementar
ChIP:	Imunoprecipitação da cromatina
DNA:	*Deoxyribonucleic Acid*
DNMT3B:	*DNA (cytosine-5)-methyltransferase 3B*
EZH2:	*Histone-lysine N-methyltransferase*
H3K4me3:	Trimetilação da lisina 4 da histona 3
H3K9me3:	Trimetilação da lisina 9 da histona 3
H3K27me3:	Trimetilação da lisina 27 da histona 3
HOXA9:	*Homeobox protein Hox-A9*
HOXB3:	*Homeobox protein Hox-B3*
IgG:	Imunoglobulina G
Kb:	kilobases
lincRNA:	RNA não codificador de proteínas longo intergênico
lncRNA:	RNA não codificador de proteínas longo
MYC:	*Myc proto-oncogene protein*
ncRNA:	RNA não codificador de proteínas
Nt:	nucleotídeos
PCR:	Reação em Cadeia da Polimerase
PRC2:	Complexo repressivo *polycomb* 2
qPCR:	PCR quantitativa
RASSF1A:	*RAS association domain family 1A*
RIP:	Imunoprecipitação de complexos ribonucleoproteicos
RNA:	*Ribonucleic acid*
RNAi:	Interferência de RNA
RNA-Seq:	Sequenciamento em larga escala de RNA
RNAP II:	RNA polimerase II
RT-qPCR:	Transcrição reversa sequida de PCR quantitativa
siRNA:	*small interfering* RNA
SUZ12:	*Polycomb protein SUZ12*
TAP:	*Tobacco Acid Pyrophosphatase*
TSS:	Sítio de início de transcrição

ÍNDICE

1. INTRODUÇÃO 10
1.1 *RASSF1* 10
1.2 Regulação epigenética da expressão de *RASSF1A* 15
1.3 A complexidade do transcriptoma eucariótico 17
1.4 RNAs não-codificadores longos em eucariotos 21
2. OBJETIVOS 25
2.1 Objetivo geral 25
2.2 Objetivos específicos 25
3. MATERIAIS E MÉTODOS 26
3.1 Linhagem celular humana e condições de cultivo 26
3.2 Extração de RNA, purificação e tratamento com DNase 27
3.3 Avaliação da qualidade do RNA 28
3.4 Desenho de oligonucleotídeos iniciadores (*Primers*) 28
3.5 Transcrição reversa e síntese de cDNA 31
3.5.1 Transcrição reversa 31
3.5.2 Transcrição reversa fita específica 31
3.6 Reação da Polimerase em Cadeia (*Polymerase Chain Reaction* - PCR) 32
3.6.1 PCR 32
3.6.2 PCR quantitativa em tempo real 33
3.7 RACE e caminhada com oligonucleotídeos iniciadores (*primer walking*) 34
3.8 Sequenciamento em larga escala e análise de bioinformática 35
3.9 Análise in silico da co-expressão de *ANRASSF1* e do gene *RASSF1* 35
3.10 Tratamento enzimático para análise de presença da modificação cap 5' 36
3.11 Fracionamento celular 36
3.12 Clonagem 37
3.13 Ensaio de promotor 38
3.14 Super-expressão 40
3.15 RNAi 41
3.16 *Western blot* 42
3.17 Ensaio de proliferação 43

3.18 Imunoprecipitação de complexos ribonucleoproteicos - RIP 44
3.19 Imunoprecipitação da cromatina - ChIP 45
3.20 Ensaio de detecção de metilação do DNA 46
3.21 Permeabilização celular e tratamento com RNase 46
4. RESULTADOS 48
4.1 – Identificação do RNA não codificador antissenso ao gene *RASSF1* (*ANRASSF1-Antisense Noncoding RNA RASSF1)* 48
4.2 - Caracterização de *ANRASSF1* 53
4.3 – Abundância de *ANRASSF1* modula a expressão de *RASSF1A* e afeta a taxa de proliferação de células em cultura 60
4.4 – A trimetilação da histona H3K27 no promotor de *RASSF1A* é dependente do nível de *ANRASSF1* ligado à PRC2 63
4.5 – *ANRASSF1* se liga ao DNA e é necessário para a ocupação de SUZ12 no promotor de *RASSF1A* 68
5. DISCUSSÃO 74
6. CONCLUSÃO 80
7. REFERÊNCIAS 81

1. INTRODUÇÃO

1.1 *RASSF1*

RASSF1A (Família do domínio de associação a Ras isoforma 1A, *RAS association domain family 1A*) é classificado como um gene supressor de tumor (Donninger et al. 2007). A perda da expressão de *RASSF1A*, um dos eventos mais comuns em cânceres humanos, é causada principalmente pela hipermetilação aberrante de seu promotor (Agathanggelou et al. 2005; Avruch et al. 2009; Hoque et al. 2010; Amin and Banerjee 2012; Fernandes et al. 2012; Jiang et al. 2012), e este fato já foi descrito em pelo menos 37 tipos tumorais (Dammann et al. 2005; Hesson et al. 2007); esse evento raramente ocorre em tecidos normais. Isso faz com que *RASSF1A* seja um candidato a marcador molecular para o diagnóstico de tumores (Agathanggelou et al. 2005).

O *locus* do gene *RASSF1* abrange 11.151 pb do genoma humano e é composto por oito exons, que combinados, geram além de *RASSF1A*, mais seis transcritos diferentes *RASSF1B-G* (Figura 1) por meio do uso de dois promotores diferentes e de *splicing* alternativo (Dammann et al. 2000; Lerman and Minna 2000; Burbee et al. 2001). Duas ilhas CpG estão associadas com os promotores de *RASSF1*. A menor ilha CpG (737 pb) abrange a região promotora do *RASSF1A* (*RASSF1D*, *RASSF1E*, *RASSF1F*, e *RASSF1G*), enquanto a maior ilha CpG (1365 pb) engloba as regiões promotoras para *RASSF1B* e *RASSF1C* (Agathanggelou et al. 2005).

RASSF1A e *RASSF1C* são as isoformas mais predominantes, normalmente expressas em tecidos normais, e são transcritas a partir de dois promotores diferentes distantes aproximadamente 3.5 kb um do outro (van der Weyden and Adams 2007). De maneira geral, as demais isoformas são tecido-especificas (Agathanggelou et al. 2005). A isoforma *RASSF1B* é principalmente expressa em células do sistema hematopoiético, enquanto *RASSF1D* e *RASSF1E* são expressas especificamente em coração e pâncreas, respectivamente. Essas duas

isoformas são muito semelhantes a *RASSF1A*, com uma pequena diferença no padrão de *splicing* no exon 2αβ de *RASSF1D* e no exon 3 de *RASSF1E*, dando origem a proteínas com quatro aminoácidos (aa) extras (Figura 1). A isoforma *RASSF1F* pula o exon 2αβ e codifica um peptídeo truncado de 92 aa (Burbee et al. 2001), já a isoforma *RASSF1G* pula os exons 2αβ e 3 e codifica um peptídeo truncado de 152 aa (Burbee et al. 2001). Com exceção das isoformas A e C, o papel biológico dessas outras isoformas ainda é desconhecido (Donninger et al. 2007).

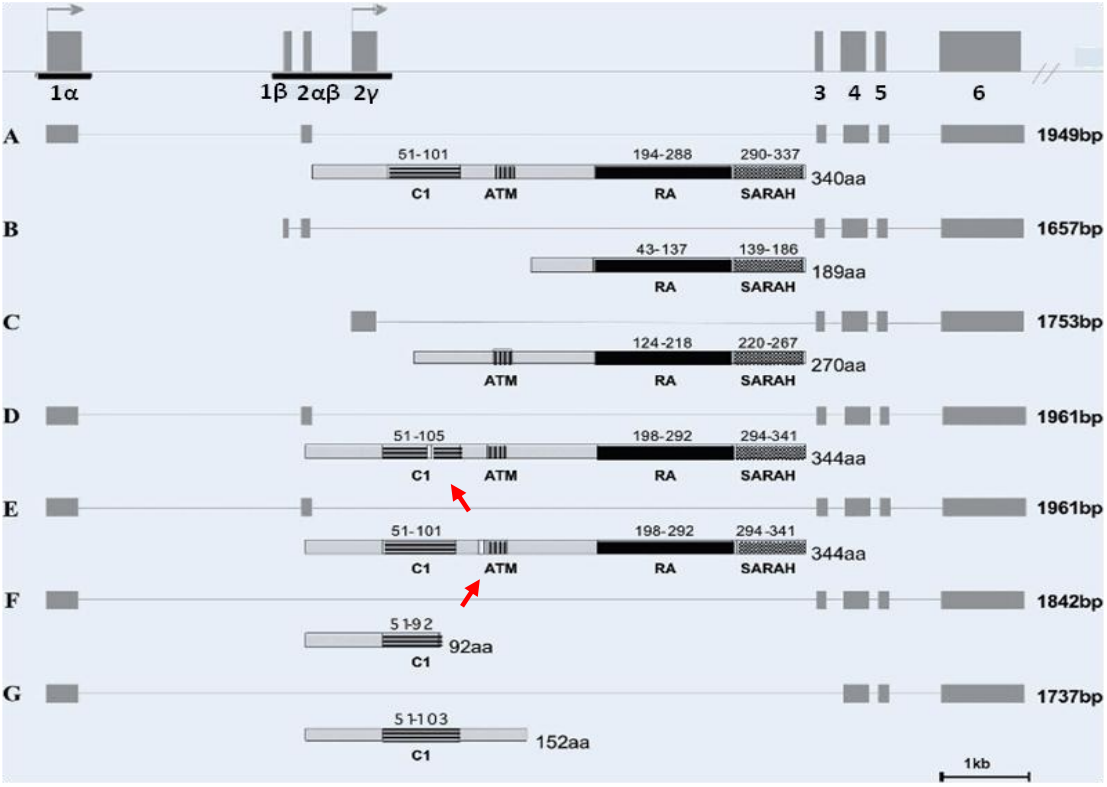

Figura 1. Mapa do *locus* gênico *RASSF1*. Diferentes isoformas *RASSF1A-G* são geradas a partir de dois promotores diferentes indicados por setas cinzas na parte superior da figura. Estão indicadas a posição dos promotores associados a ilhas CpGs (linha preta, parte superior), e a posição dos 4 aminoácidos extras das isoformas *RASSF1D* e *E* (caixa branca, as setas vermelhas apontam para estas regiões). Domínios estruturais conhecidos de *RASSF1*: C1, domínio conservado 1 da proteína quinase C; ATM, sequência consenso de fosforilação pela quinase ATM; RA, domínios de associação a Ras; SARAH, domínio de interação Sav/RASSF/Hpo. Figura adaptada de (Agathanggelou et al. 2005).

Apesar de ainda se saber pouco quanto à função das isoformas de *RASSF1*, visto que apenas duas das sete isoformas, *RASSF1A* e *RASSF1C*, possuem papel conhecido na célula, já se conhece bastante sobre os domínios estruturais constituintes de todas as isoformas. As isoformas *RASSF1A-E*, possuem dois domínios em comum, o domínio de associação a Ras (RA) e o domínio a Sav/RASSF/Hpo (SARAH) (Ponting and Benjamin 1996; Yamamoto et al. 1999), ambos importantes para a função de *RASSF1A*, como será discutido adiante. O domínio RA define o gene *RASSF*, e medeia as interações de *RASSF1* com Ras e outras pequenas GTPases. O domínio SARAH medeia as interações proteína-proteína que são cruciais para as funções de *RASSF1*, principalmente nas vias associadas com parada do ciclo celular (*cell cycle arrest*) e apoptose (Scheel and Hofmann 2003). Ainda foi identificada nas isoformas *RASSF1A* e *RASSF1C-E* uma sequência consenso de peptídeo que é substrato para a quinase *Ataxia Telagectasia Mutant* (ATM) (Agathanggelou et al. 2005). Por último, as isoformas *RASSF1A* e *RASSF1D-G* contêm o domínio C1, uma região rica em cisteína, também conhecida como região conservada 1 (C1) da proteína quinase C (Newton 1995).

A isoforma mais estudada desse gene é *RASSF1A*. Seu transcrito de 1,9 kb inicia-se no primeiro promotor, localizado na primeira ilha CpG, é formado por seis exons (Figura 1) e contém uma ORF que codifica para uma proteína de 340 aa, com massa molecular de 39 kDa (Donninger et al. 2007). Sua função de supressor tumoral ficou evidenciada em diversos estudos mostrando que a sua super-expressão em linhagens celulares tumorais resulta em células menos viáveis (com maior taxa apoptótica), com menores taxas proliferativas, menos invasivas e reduzida independência quanto ao ancoramento ao substrato (Dammann et al. 2000; Burbee et al. 2001; Dreijerink et al. 2001; Kuzmin et al. 2002; Chow et al. 2004; Hesson et al. 2004; Donninger et al. 2007).

RASSF1A exerce diversas funções na célula, regulando vários processos biológicos. Evidências recentes sugerem que a proteína RASSF1A age como uma plataforma (*scaffold*),

de modo que pode montar e modificar a função do complexo de proteínas com o qual está interagindo (Donninger et al. 2007; Oceandy et al. 2009). Essa capacidade de interagir com outras proteínas é muito importante para RASSF1A exercer a sua função regulatória dos processos biológicos, e seus domínios RA e SARAH são os responsáveis pela interação proteína-proteína (Avruch et al. 2009).

RASSF1A tem papel importante no controle do ciclo celular, sendo considerada uma nova guardiã da mitose (Mathe 2004) por meio de sua interação com Cdc 20 (Song et al. 2004; Song et al. 2005). Cdc 20 é um regulador essencial do ciclo celular, necessário para a conclusão da mitose (Yu 2007). Esse se liga e ativa o complexo promotor da anáfase (APC), promovendo a ubiquitinação e degradação das ciclinas A e B, o que resulta na saída da anáfase da mitose. Foi demonstrado que quando RASSF1A se associa a Cdc20, este complexo impede a ativação de APC e consequente progressão da mitose (Song et al. 2004) por meio de Cdc20, o que resulta em menores taxas de proliferação celular. Além disso, RASSF1A também controla a transição da fase G1 para S do ciclo celular (Fajas et al. 2001; Shivakumar et al. 2002; Fenton et al. 2004).

RASSF1A localiza-se nos microtúbulos e promove a estabilização dos mesmos (Dallol et al. 2004; Song et al. 2004) durante as mudanças dinâmicas sofridas pelo citoesqueleto ao longo das fases do ciclo celular. Essa estabilização ocorre por meio de interações com proteínas associadas a microtúbulos (MAPs) ou com o homólogo dessas proteínas, C190RF5 (Dallol et al. 2004). Também pode participar na manutenção da estabilidade genômica, por localizar-se nos centrossomos e nas fibras mitóticas, podendo modular a dinâmica da tubulina (Song et al. 2005).

Outra importante função exercida por RASSF1A é a modulação da apoptose, que ocorre por diversos mecanismos, envolvendo predominantemente a via de morte celular

extrínseca, requerendo sempre um estímulo externo por meio dos receptores de FAS ou de TNF-α (Baksh et al. 2005; Vos et al. 2006; Matallanas et al. 2007; Foley et al. 2008).

A segunda isoforma mais estudada e abundante do gene *RASSF1* é *RASSF1C*. Seu transcrito de 1,7 kb inicia-se em um promotor diferente, localizado na segunda ilha CpG, e é formado por cinco exons, sendo que esta isoforma possui quatro exons que codificam a região N-terminal em comum com *RASSF1A* (van der Weyden and Adams 2007). Em contraste com RASSF1A, a função de supressor tumoral de RASSF1C ainda não está clara. Apesar da maioria dos trabalhos indicarem que RASSF1C não possui esta função, existe uma constante discussão na literatura sobre esta questão (Reeves et al. 2010). Adicionalmente, *RASSF1C* não é alvo de inativação epigenética (Agathanggelou et al. 2005), como identificado em *RASSF1A*, no entanto a perda da expressão de RASSF1C foi observada em muitas das linhagens de células de carcinomas renais (Dreijerink et al. 2001) e em seis de nove linhagens de células ovarianas transformadas (Vos et al. 2000).

RASSF1C, aparentemente, compartilha algumas das funções biológicas de *RASSF1A*, como controle da proliferação e apoptose e estabilização de microtúbulos. Entretanto, o papel de *RASSF1C* ainda não está totalmente elucidado. Em relação à sua função, alguns estudos mostraram que a expressão ectópica de *RASSF1C* não apresentou efeitos significativos na proliferação e indução de apoptose em células H1299 e A549 NSCLC, *in vitro* e *in vivo* (Ji et al. 2002). Essa proteína também não apresentou efeito na atividade inibitória de proliferação de células U2020 SCLC *in vitro* (Li et al. 2004).

Por outro lado, RASSF1C apresenta atividade inibitória de proliferação *in vitro*, semelhante à observada para RASSF1A, em linhagem de células de próstata, LNCaP, e linhagem de carcinoma renal, KRC/Y, com consequente supressão da tumorigenicidade (Li et al. 2004). A proteína também se associa aos microtúbulos e como resultado causa uma estabilização dos mesmos, semelhante a ação de RASSF1A (Rong et al. 2004). Contudo, em

alguns tipos celulares esta associação não promove a estabilização (Song et al. 2004). Tendo em vista esses resultados contraditórios, possivelmente essa proteína possua uma função tecido específica (Li et al. 2004). RASSF1C pode regular a apoptose através da participação na ativação da via *SAPK/JNK* (Kitagawa et al. 2006).

1.2 Regulação epigenética da expressão de *RASSF1A*

Tendo em vista os importantes papéis de *RASSF1A* e *RASSF1C*, é necessária uma regulação fina da expressão dessas isoformas que são transcritas a partir de promotores independentes. O mecanismo de regulação de *RASSF1C* ainda é desconhecido, mas para *RASSF1A* sabe-se que o mecanismo de regulação é por meio de um silenciamento transcricional mediado pela hipermetilação do DNA de seu promotor (Dammann et al. 2000). Esse evento epigenético parece ser mais comum em câncer humano, de maneira que a metilação aberrante do promotor de *RASSF1A* e a consequente perda de sua expressão já foram detectadas em pelo menos 37 tipos de tumores (van der Weyden and Adams 2007).

Apesar de já existirem alguns trabalhos relacionados ao estudo de *RASSF1A*, ainda se faz necessário identificar os fatores regulatórios, bem como as vias regulatórias e os mecanismos subjacentes ao silenciamento epigenético de *RASSF1A.* No entanto, alguns trabalhos já têm avançado nesse estudo. Strunnikova e colaboradores (Strunnikova et al. 2005), inicialmente, sugeriram que a deacetilação da histona H3 e a trimetilação da histona H3K9 antecedem a metilação do DNA, durante a progressão da inativação de *RASSF1A* em células HMEC (Célula de epitélio mamário humano - *Human Mammary Epithelial Cell*). De maneira similar, Kawamoto e colaboradores (Kawamoto et al. 2007) encontraram que as linhagens celulares tumorais de próstata, LNCaP e PC3, apresentaram silenciamento de *RASSF1A* mediado pela hipermetilação e pelo enriquecimento da marca de histona H3K9me2 ao longo do seu promotor. Em contraste, em células de próstata normal RWPE-1, o promotor de *RASSF1A* não está metilado e também não existe enriquecimento da marca de histona

H3K9me2, e consequentemente há expressão de *RASSF1A*. Outro trabalho mostrou que a repressão transcricional do gene *RASSF1A* em glândulas mamárias, durante as fases iniciais da carcinogênese de mama, estava associada com um aumento na trimetilação das histonas H3K9 e H3K27 e metilação do promotor de *RASSF1A* (Starlard-Davenport et al. 2010).

Estes trabalhos sustentam a ideia de que o silenciamento gênico causado pela metilação do DNA está intimamente ligado com modificações repressivas de histonas no promotor do gene em células tumorais (Fahrner et al. 2002; Vire et al. 2006). Nesse sentido, um trabalho recente mostrou que o evento de silenciamento epigenético da ilha CpG da região promotora de *RASSF1A* requer o fator de transcrição HOXB3. Esse fator de transcrição induz um aumento da expressão da enzima DNA (citosina-5-)-metiltransferase 3 beta (DNMT3B), que é recrutada para o promotor de *RASSF1A* e silencia o gene pela hipermetilação desse promotor (Palakurthy et al. 2009). Ainda nesse mesmo trabalho, foi sugerido um modelo de regulação desse evento epigenético, o qual depende do recrutamento de DNMT3B para o promotor de *RASSF1A*. Este recrutamento requer a interação de DNMT3B com EZH2, uma histona metiltransferase componente do complexo repressivo polycomb 2 (*polycomb repressive complex 2* – PRC2). O complexo DNMT3B/PRC2 se liga a MYC, que já está associado à região promotora de *RASSF1A*, e esse evento facilita seu recrutamento para o promotor. Isto resulta em uma repressão epigenética tanto por aumentar a marca de histona H3K27me3, característica de PRC2, quanto pelo aumento da metilação do DNA no promotor (Palakurthy et al. 2009). Vale ressaltar que, embora MYC seja necessário para o recrutamento de PRC2 (Goodliffe et al. 2005), de maneira geral este não é suficiente para recrutar PRC2 e os fatores de regulação (Palakurthy et al. 2009).

Embora já seja conhecido que o silenciamento transcricional de *RASSF1A* por hipermetilação do DNA do promotor pode ser desencadeado por modificações da cromatina,

os mecanismos subjacentes ao recrutamento de PRC2 para o promotor deste gene e o silenciamento epigenético de *RASSF1A* ainda não foram totalmente identificados.

1.3 A complexidade do transcriptoma eucariótico

A programação genética dos humanos e de outros organismos complexos tem sido mal interpretada nos últimos 50 anos (Mattick 2011). Isso se deve à suposição de que o fluxo da informação genética se dava primordialmente no sentido DNA→RNA→proteína, o que era conhecido como o dogma central da biologia molecular descrito inicialmente por Crick (Crick 1958). Pensava-se que os RNAs tinham um papel limitado a elementos intermediários, enquanto as proteínas seriam responsáveis por exercer a maior parte das funções estruturais, catalíticas e regulatórias das células (Crick 1958).

Doze anos depois, Crick respondeu às críticas de que o dogma central era uma simplificação excessiva e esclareceu a sua intenção inicial. Nessa ocasião ele atualizou o conceito inicial, de modo que o RNA não deveria ser considerado como um elemento funcional de propósito único, mas teria um papel mais versátil, o que possibilitaria que a informação fluísse de volta para o genoma (Crick 1970). Embora o campo da genética tenha amadurecido, o papel funcional dos RNAs como produto dos genes permaneceu por muito tempo focado na produção de proteínas, com poucas exceções notáveis (Gingeras 2007).

Acreditava-se que o transcriptoma eucariótico era composto por RNAs mensageiros (mRNAs), elementos intermediários que transferem para o citoplasma a informação genética para a biossíntese de proteínas, e por RNAs transportadores (tRNAs), RNAs ribossomais (rRNAs), e pequenos RNAs nucleares (*small nuclear* RNAs, snRNA; *small nucleolar* RNAs, snoRNA), que exercem funções estruturais e catalíticas no processamento do mRNA ou na tradução proteica. Entretanto, apenas uma pequena fração do genoma humano, cerca de 2,2

%, corresponde a esses produtos funcionais alvos mais bem estudados até o momento, que são compreendidos principalmente pelos mRNAs (Frith et al. 2005).

O sequenciamento de diversos genomas e os avanços recentes de análises genômicas em larga escala têm revelado que, em organismos complexos, a complexidade biológica não é diretamente relacionada ao número de genes codificadores de proteínas de seu genoma (Taft et al. 2007). Por exemplo, o número de genes codificadores de proteínas estimado para eucariotos recentes mais complexos, como humanos e camundongo, é de 20,000-25,000 (Waterston et al. 2002; Collins et al. 2004), sendo próximo do número de genes de eucariotos ancestrais menos complexos como *Drosophila* e *Caenorhabditis elegans*, que possuem 13,000 e 19,000, respectivamente (Consortium 1998; Adams et al. 2000). Apesar das enormes diferenças na complexidade entre esses organismos, ainda não está totalmente claro onde estão armazenadas as informações que programam o desenvolvimento e como estas diferem entre esses organismos.

Parte desse questionamento pode ser respondido, por exemplo, pelo aumento das taxas de *splicing* alternativo, do uso de promotores e sítios de terminação alternativos e dos eventos de edição de RNAs em organismos mais complexos, permitindo assim a expressão de um conjunto maior de isoformas proteicas (Nagasaki et al. 2005; Gingeras 2007; Taft et al. 2007; Farajollahi and Maas 2010).

Outro aspecto que pode responder pela complexidade é o aumento da proporção de DNA não-codificador de proteínas presente no genoma dos organismos, que pode variar de 0,25 % do DNA total no genoma de procariotos até mais que 98 % do DNA total no genoma humano (Taft et al. 2007). Estas observações sugerem que novos RNAs não codificadores (ncRNAs) originados destas regiões possa fornecer um repertório maior de moléculas com funções regulatórias que permitiram o desenvolvimento dos eucariotos (Mattick 2007; Taft et al. 2007).

Diversas técnicas de estudo, baseadas em clonagem e sequenciamento de bibliotecas de cDNAs (Numata et al. 2003; Ota et al. 2004; Carninci et al. 2005; Katayama et al. 2005), *tilling array* (Kapranov et al. 2002; Rinn et al. 2003; Bertone et al. 2004; Kampa et al. 2004) e mais recentemente sequenciamento em larga escala (Guttman et al. 2010; Kapranov et al. 2010; Almeida et al. 2011; Mercer et al. 2012), têm sido utilizadas para investigar a transcrição do genoma eucariótico em larga escala, demonstrando que o transcriptoma de eucariotos é surpreendentemente complexo. Ficou evidente que a transcrição ocorre principalmente em regiões não codificadoras, indicando que esta acontece de forma generalizada no genoma eucariótico (Mattick 2011). Esses transcritos podem ser expressos em ambas às fitas, em regiões intrônicas de genes codificadores de proteínas ou intergênicas, podendo apresentar sítios alternativos de início de transcrição e de terminação, *splicing* alternativo e expressão tecido específica (Frith et al. 2005; Birney et al. 2007; Kapranov et al. 2007; Nakaya et al. 2007; Ponting et al. 2009).

O projeto ENCODE (*Encyclopedia of DNA Elements Consortium*), em 2007, analisou intensivamente 1 % do genoma humano através de várias técnicas, incluindo *tiling arrays* genômicos de alta densidade e sequenciamento com alta profundidade de bibliotecas de cDNA, demonstrando que ao menos 90 % do DNA genômico é transcrito em diferentes células humanas, sendo que desta fração apenas 1,2 % é capaz de codificar para proteínas (Birney et al. 2007). Mais recentemente essa análise foi estendida para todo o genoma e os autores observaram que mais de 50 % dos transcritos anotados não tem potencial de codificar proteínas (ENCODE, 2011). Do mesmo modo, a análise do transcriptoma de outros organismos eucarióticos mostrou que a transcrição generalizada do genoma não é restrita apenas a humanos, mas tem revelado, por exemplo, que pelo menos 70% do genoma de *Drosophila melanogaster*, *Caenorhabditis elegans*, levedura e camundongo são transcritos

(Carninci et al. 2005; David et al. 2006; Amaral and Mattick 2008; Gerstein et al. 2010; Roy et al. 2010).

Essas descobertas têm alterado e expandido a compreensão da organização dos genomas eucarióticos, e têm sido utilizadas como argumento para apoiar a ideia de que os ncRNAs contribuam com funções regulatórias essenciais para o ganho de complexidade, de modo análogo aos mecanismos já bem descritos de *splicing* alternativo, utilização de promotores e sítios de terminação alternativos e modificações pós-transcricionais (Taft et al. 2007). Dessa forma, vários estudos têm mostrado que os ncRNAs são regulados durante o desenvolvimento, possuem expressão tecido específica, localizam-se em compartimentos subcelulares, são regulados durante o ciclo celular, são associados com uma variedade de doenças humanas, incluindo o câncer, e se associam a complexos proteicos remodeladores da cromatina (Reis et al. 2004; Nakaya et al. 2007; Sunwoo et al. 2009; Taft et al. 2010; Guttman et al. 2011; Hung et al. 2011; Tahira et al. 2011). Além disso, evidências de seleção evolutiva têm sido observadas em alguns ncRNAs (Ponjavic et al. 2007; Guttman et al. 2009; Ponting et al. 2009).

Contudo, existe um debate intenso na literatura sobre o real significado biológico da transcrição generalizada do genoma, motivado pelo fato de que a maioria dos ncRNAs são expressos em baixos níveis na célula e de que a maior parte destes transcritos ainda não foi caracterizada funcionalmente. Essas questões têm levantado a hipótese que estes transcritos possam ser possíveis ruídos transcricionais ou artefatos técnicos (Struhl 2007; van Bakel et al. 2010). Recentemente, uma nova técnica denominada RNA *Capture-Seq*, que combina microarranjo de captura de cDNA e sequenciamento em larga escala para aumentar a sensibilidade de detecção do transcriptoma em *loci* específicos (Mercer et al. 2012), possibilitou uma rápida resposta aos questionamentos levantados principalmente por van Bakel e colaboradores. Esses dados confirmaram que as regiões intergênicas expressam uma

grande variedade de transcritos, e ainda revelaram novos transcritos, que até então não tinham sido detectadas por nenhuma técnica disponível, assim como identificaram novas isoformas desconhecidas de genes codificadores de proteína já bem caracterizados, tais como p53 e p73 (Mercer et al. 2012).

1.4 RNAs não-codificadores longos em eucariotos

Como apresentado no item anterior, os avanços nas análises em escala genômica do transcriptoma de eucariotos revelaram que a maior parte do genoma humano é transcrita, gerando um grande repertório de ncRNAs. Esses ncRNA são convencionalmente classificados em ncRNAs longos (lncRNAs), quando maiores que 200 nucleotídeos, e ncRNAs curtos, quando menores que 200 nucleotídeos. Essa classificação dos ncRNAs foi estabelecida com base em uma prática conveniente dos protocolos de purificação de RNA, onde os RNAs menores de 200 nucleotídeos são excluídos (Mercer et al. 2009). Dentre os ncRNAs curtos (< 200 nt) incluem-se os miRNAs, siRNAs, *piwi-interacting* RNAs (piRNAs), pequenos RNAs associados à regiões promotoras e de iniciação da transcrição, entre outros (Girard et al. 2006; Carthew and Sontheimer 2009; Fejes-Toth et al. 2009; Taft et al. 2009).

Embora a literatura seja dominada por trabalhos de ncRNAs curtos (< 200 nt), nos últimos anos tem havido um aumento no número de estudos que têm reportado a funcionalidade e diversidade de lncRNAs que mapeiam em regiões intrônicas e intergênicas (Birney et al. 2007; Nakaya et al. 2007; Dinger et al. 2009; Kapranov et al. 2010). A definição e a nomeação dos lncRNAs ainda estão evoluindo na literatura, e diferentes classes ou categorias de lncRNAs têm sido descritas (Wright and Bruford 2011).

Algumas classes já têm sido bem estudadas e caracterizadas. Uma delas é a dos ncRNAs longos intergênicos (lincRNA), definidos como transcritos em uma região distante pelo menos 1 kb de algum gene codificador de proteína, e que podem ou não sofrer *splicing*. Outra classe é a dos ncRNAs longos antissenso (lncRNA-AS), que possuem sobreposição

com introns de genes codificadores de proteina, e podem ou não sofrer *splicing*. Essas duas classes de lncRNAs apresentam características semelhantes a dos mRNAs codificadores de proteínas tais como: a presença da estrutura cap na ponta 5' do RNA, transcrição a partir de uma região promotora independente (identificada por meio da marca de histona H3K4me3), e presença da RNA Polimerase II no sítio de inicio de transcrição (TSS) (Guttman et al. 2009; Conley and King Jordan 2012). Além disso, alguns trabalhos têm mostrado que esses transcritos podem ser poli-adenilados ou não-poliadenilados, e que podem se acumular diferentemente no compartimento nuclear ou citoplasmático da célula, ou até mesmo estar igualmente distribuídos nos dois compartimentos (Kapranov et al. 2007; Kapranov et al. 2010).

Apesar do grande repertório dos lncRNAs no transcriptoma, estes apresentam de maneira geral uma baixa expressão quando comparada a dos gene codificadores de proteínas, e muitos possuem expressão tecido específica (Birney et al. 2007; Cabili et al. 2011). Além disso, evidências têm apontado para o envolvimento de lncRNAs com diversas doenças humanas (Taft et al. 2010; Esteller 2011; Wapinski and Chang 2011), especialmente com o câncer (Prensner and Chinnaiyan 2011; Reis and Verjovski-Almeida 2012; Spizzo et al. 2012).

Está se tornando claro, embora ainda pouco caracterizado em detalhe, que lncRNAs podem desempenhar papéis regulatórios e estruturais em importantes processos biológicos, podendo atuar por meio de diversos mecanismos moleculares (Clark and Mattick 2011; Wang et al. 2011). De maneira geral, há fortes evidências que o mecanismo de ação dos lncRNAs envolve a interação entre lncRNA e proteínas (Khalil and Rinn 2011), especialmente proteínas que compõem complexos remodeladores da cromatina com funções tanto ativadoras como repressoras da expressão gênica (Khalil et al. 2009; Flynn and Chang 2012). Alguns trabalhos têm indicado que além de interagir com estes complexos proteicos remodeladores

da cromatina, os lncRNAs se ligam ao DNA formando um híbrido RNA/DNA, guiando os complexos para sítios específicos do genoma para executar sua função (Chu et al. 2011; Mattick 2012). Além disso, os lncRNA podem promover ativação dos complexos proteicos alvos mediante modificações alostéricas (Mattick 2012).

Um exemplo, muito bem caracterizado quanto à sua interação com proteínas e funções é o lincRNA *HOTAIR* (Rinn et al. 2007), que é transcrito no locus *HOX*C, e silencia genes *HOXD* e outros alvos. Inicialmente foi descrito que o HOTAIR se liga ao *polycomb repressive complex 2* (PRC2) e direciona esse complexo para a região promotora do locus HOXD (Rinn et al. 2007). Esse complexo trimetila o resíduo de lisina 27 da histona 3 (H3K27me3), o que resulta em silenciamento gênico. Posteriormente foi observado também que esse lincRNA age como um *scaffold* para dois complexos remodeladores da cromatina (Tsai et al. 2010). Na porção 5´ do *HOTAIR* se liga a proteína SUZ12, integrante do PRC2. Já na sua porção 3´ se liga a proteína LSD1, integrante do CoREST/REST *repressor complex*, que retira a dimetilação do resíduo de lisina 4 da histona H3 (H3K4me2), o que também está relacionado com a repressão da expressão gênica e silenciamento dos genes *HOXD* (Tsai et al. 2010). Além disso, foram encontradas 832 regiões de ligação de *HOTAIR* no genoma anotadas como promotores e introns; essas regiões foram identificadas através de experimentos de *Chromatin Isolation by RNA Purification* (ChIRP) seguido de sequenciamento de larga escala, um método de mapeamento das regiões de ligação no genoma dos lncRNAs (Chu et al. 2011). Esse resultado indica que este lncRNA regula não apenas genes *HOX*, mas também possui outros alvos ao longo do genoma.

Um outro exemplo de interação entre lncRNA/DNA/proteínas é o do lincRNA *Mistral*, que é transcrito em uma região intergênica, entre os genes *HOX6* e *HOX7*, cuja transcrição é ativada durante a diferenciação celular (Bertani et al. 2011). O mecanismo descrito para o lincRNA *Mistral* é interessante no contexto do atual projeto porque ele recruta

o complexo remodelador da cromatina MLL1 para os promotores de sua vizinhança (*HOX6* e *HOX7*) e causa ativação da transcrição, que é dada pela trimetilação da histona H3K4.

Em 2004 nosso grupo utilizou microarranjos de cDNA enriquecidos com sondas para transcritos longos não-codificadores que mapeiam em regiões intrônicas dos genes codificadores de proteína humanos, estudamos a expressão gênica dos tumores de próstata de 27 pacientes, e descrevemos uma assinatura de expressão gênica correlacionada com o grau de diferenciação (ou grau de Gleason) dos tumores (Reis et al. 2004). Dentre os 56 transcritos com expressão significativamente alterada, e correlacionados com o grau de Gleason dos tumores, apontamos para um conjunto de 24 lncRNAs dentro de regiões intrônicas, que foram identificados como diferencialmente expressos (Reis et al. 2004). O lncRNA transcrito na direção antissenso no *locus* do gene *RASSF1* era o candidato que apresentou a maior correlação com o grau de Gleason dos tumores, e motivou o interesse do grupo em estudar este *locus* gênico.

2. OBJETIVOS

2.1 Objetivo geral

Este trabalho buscou caracterizar funcionalmente *ANRASSF1*, um RNA não codificador intrônico longo *unspliced*, expresso na orientação antissenso do gene *RASSF1*. Para isso, foi investigado o envolvimento desse lncRNA em eventos regulatórios transcricionais do gene *RASSF1* e suas isoformas.

2.2 Objetivos específicos

1. Caracterização de *ANRASSF1* quanto a sua expressão antissenso, seu tamanho completo, sua biogênese, presença da modificação cap 5', estabilidade, localização sub-celular e abundância.
2. Identificação de uma região promotora de *ANRASSF1* para validar este lncRNA como uma unidade transcricional independente.
3. Avaliar os efeitos da super-expressão e silenciamento de *ANRASSF1* na expressão das isoformas de *RASSF1*.
4. Identificar o possível mecanismo de regulação das isoformas de *RASSF1* pelo lncRNA *ANRASSF1*.
5. Identificar um possível efeito fenotípico da super-expressão ou silenciamento de *ANRASSF1* em células em cultura

3. MATERIAIS E MÉTODOS

3.1 Linhagem celular humana e condições de cultivo

Para a realização deste trabalho foi utilizada principalmente a linhagem celular humana de adenocarcinoma cervical, HeLa (CCL-2), que é amplamente utilizada na comunidade científica para diversos estudos. Outras linhagens celulares humanas utilizadas foram HEK293 (célula embrionária de rim); LNCaP (célula tumoral de próstata); MCF7 (célula tumoral de mama); 786-O (célula tumoral de rim); MDA-MB-231 (célula tumoral e metastática de mama); HepG3 (célula tumoral de fígado); RC-124 (célula normal de rim) e NHF (célula normal de fibroblasto). Este conjunto de linhagens celulares foi utilizado apenas no experimento de RT-PCR fita específica item 3.6.2. e a linhagem tumoral de próstata LNCaP também foi utilizada para sequenciamento de larga escala no item 3.8.

Todas as linhagem foram adquiridas da ATCC (*American Type Culture Collection*) e cultivadas à 37 ^{o}C em atmosfera de 5% de CO_2 em meio de cultura apropriado, conforme recomendações da ATCC. O meio recomendado e utilizado para a linhagem celular HeLa foi DMEM (*Dulbecco's Modified Eagle Medium*) (Cultilab), suplementado com SFB (Soro Fetal Bovino) (Cultilab) 10%, bicarbonato de sódio 1,2 g/L, glicose 4,2 g/L e com os antibióticos penicilina (100 U/mL) e estreptomicina (100 μg/mL). Para a manutenção das células em cultura, ao atingirem aproximadamente 80% da densidade de saturação as células foram lavadas com solução salina PBS (*Phosphate Buffered Saline*) (NaCl 140 mM, KCl 2,7 mM, Na_2HPO_4 8mM e KH_2PO_4 1,5mM, pH 7,2) e subcultivadas após o uso de solução de tripsina 0,1% (Cultilab). Os estoques celulares foram mantidos no meio de cultivo com 10% de DMSO (dimetilsulfóxido) (Sigma) e armazenados em freezer a -80 ^{o}C.

Para os experimentos de inibição de RNA Polimerase II, as células foram semeadas (~2 x 10^6 células HeLa em placas de 56 cm^2) e cultivadas por 24 horas, quando o meio de cultura foi trocado por meio fresco com α-amanitina (Sigma) 50 μg/mL ou com o veículo

(H_2O MilliQ estéril) e tratadas por 9 horas (Lee et al. 2004; Nakaya et al. 2007; Raha et al. 2010).

Para os experimentos de meia vida, as células foram semeadas (~2 x 10^6 células HeLa em placas 56 cm^2) e cultivadas por 24 horas, quando o meio de cultura foi trocado por meio fresco com actinomicina-D (Invitrogen) 10 μg/mL ou com o veículo (DMSO) e coletadas em diferentes tempos: 0, 1, 3, 6 e 8 horas (Ebralidze et al. 2008).

3.2 Extração de RNA, purificação e tratamento com DNase

Células HeLa foram cultivadas em frascos apropriados até a confluência desejada, lavadas com PBS e descoladas da superfície do frasco por tratamento com 0,1% de tripsina seguido da adição de meio de cultura contendo SFB, ou apenas removidas do frasco raspando as células com *cell scraper* ("*rodo*" de borracha) em PBS (no caso dos tratamentos com α-amanitina, actinomicina-D, item 3.1). Em seguida, as células foram sedimentadas por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. Os sedimentos de células foram utilizados imediatamente (no caso do fracionamento celular, item 3.11) ou mantidos a -80 ^oC até extração do RNA. O isolamento do RNA foi realizado utilizando TRizol (Invitrogen), de acordo com o protocolo do fabricante, e eluído em um volume de 100 μL de água tratada com DEPC (Invitrogen). Em seguida, o RNA é submetido a uma etapa de purificação e tratamento com DNaseI utilizando o kit RNAspin (GE Healthcare), de acordo com o protocolo do fabricante, com alteração do tempo de incubação da DNaseI de 45 min para 1 hora. O RNA purificado é armazenado a -80 ^oC.

3.3 Avaliação da qualidade do RNA

As amostras foram quantificadas por sua densidade óptica a 260nm, e sua pureza atestada pela razão 260/280nm. As medidas foram realizadas no espectrofotômetro NanoDrop (Thermo Scientific).

O RNA foi analisado quanto à sua qualidade utilizando-se o equipamento Bioanalyzer 2100 (Agilent Technologies), que realiza uma eletroforese capilar de alta tensão. A integridade das amostras foi avaliada com o programa 2100 Expert (Agilent). Este programa atribui um valor de integridade de RNA *(RNA Integrity Number*, RIN), que permite uma estimativa da integridade por meio da análise das intensidades das bandas de todo o traçado eletroforético da amostra, e não apenas da razão entre as bandas dos RNA ribossomais. Apenas amostras que tinham uma boa integridade (RIN > 8) foram utilizadas em todos os experimentos.

Além disso, somente RNAs que não apresentavam contaminação com DNA genômico foram utilizados para os ensaios. Para certificar a ausência de DNA genômico nas amostras de RNA, foi realizada uma PCR (50 μL, em 40 ciclos) do gene α-tubulina utilizando os oligonucleotídeos iniciadores para α-Tubulina (Tabela 1) e 1μg de RNA total de cada amostra, sem transcrição reversa, como molde para a reação de PCR (conforme será descrito no item 3.6.1).

3.4 Desenho de oligonucleotídeos iniciadores (*Primers*)

Os oligonucleotídeos iniciadores utilizados nos ensaios de transcrição reversa fita específica, clonagem, presença da modificação 5´cap, RACE e caminhada com oligonucleotídeos (*primer walking*) foram desenhados utilizando-se o programa Primer3 (Rozen and Skaletsky 2000) e estão descritos na Tabela 1.

Tabela 1. Oligonucleotídeo iniciadores utilizados para RT-fita específica, clonagem, ensaio da presença de 5'-cap, RACE e caminhada com oligonucleotídeos iniciadores (*Primer walking*)

Alvo	Nome dos oligonucleotídeos iniciadores	Sequência
RT-fita específica		
ANRASSF-SENSO	*ANRASSF1*_Senso_F	CCGCGCAGAATTAGCCTCT
ANRASSF-ANTISSENSO	*ANRASSF1*_Antissenso_R	CCAATGAGGAAAGGGGAAGT
PCR contaminação DNA genômico		
α-TUBULINA	α-TUBULINA_RT_F	TCAACACCTTCTTCAGTGAAACG
	α-TUBULINA_RT_R	AGTGCCAGTGCGAACTTCATC
clonagem (super-expressão e ensaio de promotor)		
Antisense-Transcript	ANRASSF_KpnI_F	TGGGTACCGCAGCGGGTGGAGTACTTG
	ANRASSF_HindIII_R	TGAAGCTTCCAATGAGGAAAGGGGAAGT
Antisense-Promoter	pANRASSF_F_BglII-AS	AGATCTGACCCAGGACGCGGCAAC
	pANRASSF_R_HindIII-AS	AAGCTTCCAGAGGCCACTCTTGTGC
Sense-Promoter	pANRASSF_F_HindIII-S	AAGCTTGACCCAGGACGCGGCAAC
	pANRASSF_R_BglII-S	AGATCTCCAGAGGCCACTCTTGTGC
ensaio da presença de estrutura 5'-cap		
snRNA U15A	snRNA U15A_F	GAAGAGATGATGACGAGTCTGACTTG
snRNA U15A	snRNA U15A_R	GAAATTACTTCAACCAGGGCTCTTT
ANRASSF1	*ANRASSF1*_RT_F	GGCAATTAGAACGCTCCTTG
ANRASSF1	*ANRASSF1*_RT_R	CTGTGCTAGGCGATAGAGATCC
α-TUBULIN	α-TUBULIN_RT_F	TCAACACCTTCTTCAGTGAAACG
α-TUBULIN	α-TUBULIN_RT_R	AGTGCCAGTGCGAACTTCATC
RACE e caminhada com oligonucleotídeos iniciadores (*Primer Walking*)		
	ANRASSF_ RACE(3)External_R	ACCTCACACTGCTACGCGGACT
	ANRASSF_RACE(3)_nested_R	GCCTGCTCAACAGTTGGATCTC
	ANRASSF_RACE(5)_External_F	GAGCGTTCTAATTGCCGATTTCC
	ANRASSF_RACE(5)_nested_F	CGGCGGCACAGAGAGGCTAATTCTGCG
	ANRASSF_Walk_F1	GTCCGCTTGCAGCGGGTG
	ANRASSF_Walk_F2	AGCCTGGGTCAGCCTGGG
	ANRASSF_Walk_F3	CACAAGAGTGGCCTCTGGC
	ANRASSF_Walk_F4	ACAGACCCCACCTACCACAG
	ANRASSF_Walk_F10	GGAAGGTGCGGGAAGTG
	ANRASSF_Walk_TSS_F1	GGAATGACCTCATCGCTCCG
	ANRASSF_Walk_TSS_F2	CAACCGTTAAGACTGAAACGTAGATCG
	ANRASSF_Walk_TSS_F3	ACGCACGCTTCGCCCC
	ANRASSF_Walk_R	GTATGGGAAGGGCGAGGAT

Já para os experimentos que utilizaram PCR quantitativo, os oligonucleotídeos iniciadores utilizados foram desenhados no programa Primer Express versão 3.0 (Applied Biosystems), de acordo com os parâmetros padrão e estão descritos na tabela 2.

Tabela 2. Oligonucleotídeos iniciadores utilizados para qPCR, ChIP e metilação do DNA.

Alvo	**Nome dos oligonucleotídeos iniciadores**	**Sequência**
qPCR		
ANRASSF1	*ANRASSF1*_RT_F	GGCAATTAGAACGCTCCTTG
	*ANRASSF1*_RT_R	CTGTGCTAGGCGATAGAGATCC
RASSF1C	RASSF1C_RT_F	GAAATGACCTGGAGCAGCAC
	RASSF1C_RT_R	TTGATCTGGGCATTGTACTCC
RASSF1A	RASSF1A_RT_F	ACCTCTGTGGCGACTTCATC
	RASSF1A_RT_R	GTTCGTGTCCCGCTCCAC
MALAT1	MALAT1_RT_F	TTTGATCTAGCACAGACCCTTCAC
	MALAT1_RT_R	GACCTCGACACCATCGTTACCT
LincSFPQ	LincSFPQ_RT_F	TCGTACTGTTAGGCCCTTGG
	LincSFPQ_RT_R	AACCTTGCATGAAGAGCACC
HOTAIR	HOTAIR_RT_F	GAGAGCCAGAGGAGGGAAGAG
	HOTAIR_RT_R	CACGTTTGTTCCGGGAACTG
α-TUBULINA	α-TUBULINA_RT_F	TCAACACCTTCTTCAGTGAAACG
	α-TUBULINA_RT_R	AGTGCCAGTGCGAACTTCATC
pre-tRNA^{Tyr}	pre-tRNA^{Tyr}_F	AAAAAACCGCACTTGTCTCCTTCG
	pre-tRNA^{Tyr}_R	CCTTCGATAGCTCAGCTGGTAGAG
c-myc	c-Myc_RT_F	TCAAGAGGTGCCACGTCTCC
	c-Myc_RT_R	TCTTGGCAGCAGGATAGTCCTT
HPRT1	HPRT1_F	TGACACTGGCAAAACAATGCA
	HPRT1_R	GGTCCTTTTCACCAGCAAGCT
ChIP		
Promotor RASSF1A	Promotor-RASSF1A_F	TAGAGGAAGAGGGTCCCCACAT
	Promotor-RASSF1A_R	TGGGTAGGCCAAGTGTGTTG
Promotor *HYAL2*	Promotor- HYAL2_F	GACGTGGCCGCCCATT
	Promotor- HYAL2_R	CTCGTGGCTGAGACTCCCATA
Promotor *TUSC2*	Promotor- TUSC2_F	GGCTGTCTCCACGGAAACC
	Promotor- TUSC2_R	CCCCACGACTTGAGCAACA

Metilação do DNA		
Promotor RASSF1A	Promotor-RASSF1A_F	TAGAGGAAGAGGGTCCCCACAT
	Promotor-RASSF1A_R	TGGGTAGGCCAAGTGTGTTG
Promoter *RASSF1C*	Promoter-RASSF1C_F	TTCGCGCGGTGAAGTACTG
	Promoter-RASSF1C_R	GCGCCTTCTTTCGAAATGAC
Promoter *ZMYND10*	Promoter- ZYMD10_F	AATCCGGGCGGGTCTCTAG
	Promoter- ZYMD10_R	GAAAGCGCTGGGAATCCAA
Promoter *NPRL2*	Promoter- NPRL2_F	GCCTCACAGTTGTCTGCGAAT
	Promoter- NPRL2_R	GCTCAACTGGCTGCCTGAA
Promoter *HOXA9*	Promoter- HOXA9_F	TGACCTTGAATGGCCCAAAG
	Promoter- HOXA9_R	CTGGTCACAGCAGGTAGGGTTAG

3.5 Transcrição reversa e síntese de cDNA

3.5.1 Transcrição reversa

Para a transcrição reversa e síntese de cDNA fita simples foi utilizado 1 µg de RNA total de cada condição experimental. Cada amostra foi transcrita com o kit SuperScript III First Strand Synthesis Mix (Invitrogen), utilizando oligonucleotideos iniciadores oligo-dT(20) ou hexameros randômicos, segundo as recomendações do fabricante, com incubação à 50 ^{o}C. Uma pré-incubação de 3 min a 85 ^{o}C foi introduzida no protocolo, antes da fase de anelamento, para desnaturar completamente o RNA. Após o término da fase de anelamento, a reação foi mantida a 50 ^{o}C, para a adição da enzima em tampão pré-aquecido à 50 ^{o}C, eliminando-se a etapa de resfriamento no gelo.

3.5.2 Transcrição reversa fita específica

Para validar a hipótese de que *ANRASSF1* era transcrito antissenso, foi realizada uma transcrição reversa fita-específica utilizando oligonucleotídeos iniciadores específicos complementares à sequência genômica da fita mais ou da fita menos no *locus* de *ANRASSF1*, em regiões que tivessem sobreposição apenas com os introns do gene codificar *RASSF1A*,

conforme indicado na Figura 4B. Foi utilizado para a transcrição reversa 3 µg de RNA total das linhagens de células testadas, Transcriptase Reversa (SuperScript III First Strand Synthesis Mix, Invitrogen) segundo as recomendações do fabricante, com incubação a 57 ^{o}C. Uma pré-incubação de 3 min a 85 ^{o}C foi introduzida no protocolo, durante a fase de anelamento, para desnaturar completamente o RNA. Após o termino desta fase o meio é mantido a 57 ^{o}C, para acrescentar a enzima, eliminando-se a etapa de resfriamento no gelo. Como controle adicional, uma transcrição reversa sem oligonucleotídeos iniciadores foi incluída para mostrar que o cDNA gerado não é um artefato causado por iniciador gerado a partir do auto-anelamento de um RNA ou por contaminação de DNA genômico.

3.6 Reação da Polimerase em Cadeia (*Polymerase Chain Reaction* - PCR)

3.6.1 PCR

Para amplificar os transcritos de interesse, foi realizada reação em cadeia da polimerase (PCR) utilizando-se 0,2 µM de oligonucleotídeos iniciadores específicos para cada transcrito e 1U de Taq Polimerase (GoTaq Green Master Mix; Promega), segundo as instruções do fabricante. Utilizou-se 250ng de DNA genômico de células HeLa como controle positivo da reação e H_2O DEPC (Invitrogen) como controle negativo. O produto da reação foi submetido à eletroforese em gel de agarose 1,7% contendo brometo de etídeo 0,5 µg/mL em tampão TAE (Tris-Acetato 40mM, EDTA 1mM). O produto da reação foi visualizado em um fotodocumentador MiniBis Pro (DNR Bio-Imaging Systems) sob luz UV e as imagens foram adquiridas através do programa GelCapture (DNR Bio-Imaging Systems).

3.6.2 PCR quantitativa em tempo real

As reações da PCR quantitativa em tempo real (qPCR) foram feitas em triplicata em um volume de 20µL contendo 10µL de Sybr Green Master Mix (Applied Biosystems), 5µL de cDNA da transcrição reversa (diluído 1:6) e 3,2µM de oligonucleotídeos iniciadores *Foward* e *Reverse* específicos para cada gene. A reação foi realizada no aparelho 7500 Real–Time PCR System (Applied Biosystems) utilizando os parâmetros padrão do aparelho.

O ciclo da PCR no qual a intensidade de fluorescência do *Sybr Green*, ligado à dupla fita de DNA amplificado, é detectada determina o Ct do gene (*Cycle threshold* – ciclo referência). Este Ct é uma medida relativa da quantidade inicial de moléculas do transcrito alvo presente no cDNA ou DNA genômico. Quanto menor o Ct, mais expresso é o transcrito. Para tornar as medidas de expressão comparáveis entre as amostras usa-se como referência o Ct de um gene constitutivo ou que não tenha sua expressão alterada nas diferentes condições de estudo. O 5S rRNA foi utilizado como normalizador no experimento de α-amanitina, α-Tubulina como normalizador nos demais experimentos em que utilizou-se cDNA. Para os experimentos de ChIP e RIP utilizou-se a medida da quantidade inicial (*input*) do alvo presente na amostra (Tabela 2) como normalizador, e por fim, nos ensaios de detecção de metilação do DNA utilizou-se a medida da condição sem adição de enzima como normalizador.

Calcula-se o delta-Ct (ΔCt) pela diferença entre a média das triplicatas (réplicas técnicas) do Ct do gene em análise e o Ct médio do normalizador. A partir dos valores de ΔCt calcula-se o valor de delta-delta-Ct (ΔΔCt) que consiste na subtração do ΔCt obtido nas diferentes condições testadas (tratamento e controle). A diferença de expressão relativa do gene em estudo, *fold change*, é calculada como $2^{-\Delta\Delta Ct}$. Para a análise de meia-vida, o *fold change* foi calculado a partir do delta-Ct obtido pela diferença entre a média do Ct detectado para o gene em análise em células tratadas com actinomicina D e o Ct médio do mesmo gene

detectado em células controle não tratadas. A significância estatística da diferença de expressão dos transcritos testados por qPCR entre condições distintas (tratamento e controle) foi medida utilizando o teste t de *Student* (pareado, bicaudal, homocedástico), considerando-se como limiar de significância p-valor <0,05. Neste teste, foram considerados os valores médios de ΔCt de cada condição, referentes à pelo menos 3 réplicas biológicas

3.7 RACE e caminhada com oligonucleotídeos iniciadores (*primer walking*)

Para determinar o tamanho completo do transcrito, ensaios de RACE (*Rapid Amplification of cDNA Ends*) foram realizados para obtenção da sequência das extremidades 5´e 3´. Foi utilizado cDNA comercial de colo de útero (*Human Colon Marathon-Ready*; Clontech) preparado a partir de RNA poliA, contendo adaptadores em cada uma das extremidades, e *Advantage 2 polymerase* (Clontech) de acordo com o manual do fabricante para a amplificação por RACE (5´ e 3´). Dois oligonucleotídeos específicos complementares a sequência do *ANRASSF1* foram desenhados ancorados em suas pontas, orientados para as extremidades da mensagem, de forma a possibilitar a utilização de cada oligo em reações de PCR sequênciais (*Nested*), que aumentam a especificidade dos produtos gerados; os oligonucleotídeos estão descritos na tabela 1.

Para obter a sequência final da extremidade 5´ do transcrito realizamos o ensaio de caminhada com oligonucleotídeos (*primer walking*). Realizamos predições *in silico* para identificar o possível local do início da transcrição (TSS) com o programa TSSW (http://www.softberry.com), utilizando como entrada a sequencia da região identificada como promotora do *ANRASSF1*, ver item 3.13. Foram desenhados oligonucleotídeos específicos "*forward*" para cada TSS predita e um oligonucleotídeo específico "*forward*" alternativo a 30 nt mais acima de cada região. Amplificação por PCR (item 3.6.1) foi realizada com cada um desses oligonucleotídeos específicos a partir de cDNA fita específica (item 3.5.2), utilizando

como segundo membro do par o oligonucleotídeo específico que se anele em uma porção conhecida da extremidade 5´; os oligonucleotídeos estão descritos na tabela 1. Os produtos da amplificação por PCR foram clonados conforme será descrito no item 3.12 e enviados para sequenciamento, para confirmar se a sequência amplificada era a mesma da predita.

3.8 Sequenciamento em larga escala e análise de bioinformática

Optamos por contratar o serviço de sequenciamento Illumina da *facility* da Scripps Clinic de San Diego, onde foi utilizado o equipamento Hiseq-2000 (Illumina) com o kit *Directional mRNA-Seq Sample Preparation* (Illumina) que gera fragmento de 100 pb e aproximadamente 100 milhões de sequências por *lane*. Utilizamos o programa TopHat (Trapnell et al. 2009) para mapear as sequências no genoma humano e no *locus* de *RASSF1A*.

3.9 Análise in silico da co-expressão de *ANRASSF1* e do gene *RASSF1*

A análise *in silico* foi realizada utilizando dados públicos de expressão de microarranjo da plataforma Affymetrix HG-U133 Plus2, que estão depositados e disponíveis no site do *National Center for Biotechnology Information* (NCBI) no banco de dados Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/). O método de processamento do RNA e da hibridação podem ser encontrados nas publicações originais. Os dados de expressão e a anotação das amostras foram carregados do NCBI GEO: GSE5823 (Cappellen et al. 2007), GSE11118 (Byun et al. 2009). Para cada estudo, foi realizada uma análise de correlação de Pearson entre os dados de expressão para o *probe set* (conjunto de sondas) 240278_at (*ANRASSF1*) e para o *probe set* 204346_s_at (*locus* inteiro do RASSF1, incluindo todas as isoformas).

3.10 Tratamento enzimático para análise de presença da modificação cap 5'

A enzima *Tobacco Acid Pyrophospahatase* (TAP; Epicentre Biotechnologies) cliva a ligação pirofosfato do nucleotídeo com a estrutura cap, guanosina metilada 5'-terminal, deixando um monofosfato livre na extremidade 5'. A enzima *Terminator 5'-Phosphate-Dependent Exonuclease* (5'-Exo; Epicentre Biotechnologies) é uma exonuclease processiva 5' → 3' que digere RNAs 5'- monofosfato. Inicialmente, 5 µg de RNA total de células HeLa foram tratados com 15 U da enzima Tobacco, na presença de 40 U do inibidor de RNAse RNAseOUT (Invitrogen) durante 2h a 37 ^{o}C. Em paralelo, foram feitos dois tratamentos sem enzima, um deles foi utilizado como controle negativo e o outro foi submetido à próxima etapa de clivagem enzimática pela 5'-Exo (TAP^{-}/5'-Exo^{+}). Após isso, as 3 amostras foram purificadas usando RNeasy Micro Kit (Qiagen), de acordo com o protocolo do fabricante. O RNA recuperado após o tratamento com TAP foi submetido ao tratamento com a enzima 5'-Exo (1 U para 1 µg de RNA), na presença de 40 U de RNAseOUT (Invitrogen) (TAP^{+}/5'-Exo^{+}), 2h a 37 ^{o}C. A amostra TAP^{-}/5'-Exo^{+} foi sujeita ao mesmo tratamento enzimático com 5'-Exo, enquanto o controle negativo foi novamente sujeito às mesmas condições de reação, exceto pelo fato de não receber adição de enzima (TAP^{-}/5'-Exo^{-}). Em seguida, as amostras foram purificadas usando RNeasy Micro Kit (Qiagen) de acordo com o protocolo do fabricante.

3.11 Fracionamento celular

O fracionamento celular foi realizado conforme o protocolo descrito em Topisrovic et al. (Topisirovic et al. 2003). As células (~3 x 10^{7}) foram coletadas e sedimentadas (conforme descrito no item 3.2), em seguida foram lavadas em PBS pH 7,4 gelado, ressuspendidas gentilmente em tampão de lise (Tris pH 8,4 10 mM, NaCl 140 mM, $MgCl_2$ 1,5 mM, 0,5% NP-40, DTT (ditiotreitol) 1mM e RNasin (100 U/ml) (Promega)) e incubadas em gelo durante

5 minutos. A eficiência da lise, presença de células intactas *versus* núcleos livres, foi verificada através da análise de uma alíquota de 1 µL do lisado por microscopia de luz visível, utilizando o microscópio Olympus CKX41. Em seguida, as suspensões celulares foram centrifugadas a 1.000 g durante 5 min a 4 ^{o}C e o sobrenadante foi salvo como fração citoplasmática. Os sedimentos (frações nucleares) foram ressuspendidos em tampão de lise acrescido de 3,3% de deoxicolato de sódio e 6,6% de Tween 40 e submetidos à nova centrifugação a 1.000 g para remoção completa de resquícios da fração citoplasmática. O sobrenadante foi descartado e o sedimento foi coletado como fração nuclear. As frações nucleares e citoplasmáticas foram ressuspendidas em 1 mL e 4 mL de TRizol (Invitrogen), respectivamente, e armazenadas a -80 ^{o}C.

3.12 Clonagem

Após amplificação dos transcritos de interesse por PCR, o produto da reação foi submetido à eletroforese em gel de agarose 1,7% contendo brometo de etídeo 0,5 µg/mL em tampão TAE (Tris-Acetato 40mM, EDTA 1mM). O produto da reação foi visualizado em um fotodocumentador MiniBis Pro (DNR Bio-Imaging Systems) sob luz UV. As bandas correspondentes aos produtos de PCR obtidos foram cortadas do gel e purificadas com o kit Quiack Gel Extraction (Qiagen), de acordo com o protocolo do fabricante. Os produtos da PCR purificados foram clonados em pGEM-T Easy Vector System (Promega), de acordo com o protocolo do fabricante. Em seguida, foi realizada diálise por 2 horas utilizando-se uma membrana VS de 0,025 µm (Millipore) para a remoção dos sais da reação. A reação dialisada foi utilizada para a transformação de células competentes de *Escherichia coli*, cepa DH10B (Invitrogen), por meio de eletroporação. Imediatamente após a eletroporação, 400 µL de meio SOC (bacto-triptona 20 mg/ml, extrato de levedura 5 mg/ml, NaCl 0,584 mg/ml, KCl 0,186 mg/ml, $MgCl_2$ 10 mM, MgSO4 10 mM, glicose 20 mM, pH 7,0) foram adicionados às

células, que foram incubadas a 37 ^{o}C por 1h. Posteriormente, as bactérias foram semeadas em meio LB-ágar (peptona de caseína 1%, extrato de levedura 0,5%, NaCl 1%, ágar bacteriológico 1,4%, pH 7,0) contendo ampicilina (10 mg/mL), X-Gal 20 mg/ml (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside) e IPTG (Isopropyl β-D-1-thiogalactopyranoside) 20 mg/ml e mantidas a 37 ^{o}C durante uma noite. As colônias selecionadas foram coletadas e inoculadas em 5 mL de meio LB líquido com ampicilina (10 mg/mL) a 37 ^{o}C, durante uma noite, com agitação constante. A extração do DNA plasmidial contendo o inserto foi realizada com o Plasmid Mini Kit (Qiagen), de acordo com o protocolo do fabricante. O DNA obtido foi quantificado por sua densidade óptica a 260 nm e sua pureza foi atestada pela razão 260/280 nm. As medidas foram realizadas no espectrofotômetro NanoDrop (Thermo Scientific). As construções foram utilizadas para a preparação dos vetores de transfecção do ensaio de promotor e super-expressão, como descrito a seguir nos itens 3.13 e 3.14 e para sequenciamento de produtos de amplificação por PCR.

O seqüenciamento foi realizado no aparelho ABI PRISM 3100 (Applied Biosystems) conforme recomendações do fabricante, sendo que as reações de seqüenciamento foram realizadas com Big dye terminator mix (Perkin Elmer) de acordo com as instruções do fabricante.

3.13 Ensaio de promotor

Uma possível região promotora bidirecional que abrangia o sitio de inicio de transcrição (TSS) de *ANRASSF1* (denominado promotor antissenso; está na fita mais do DNA genômico) e *RASSF1C* (denominado promotor senso; está na fita menos do DNA genômico) foi predita *in silico* no trabalho de (Davuluri et al. 2001), esquema na Figura A. Essas possíveis regiões promotoras foram amplificadas por PCR conforme descrito no item 3.6.1 utilizando oligonucleotídeos iniciados específicos (Tablea 1) e DNA genômico como molde,

clonadas no pGEM-T vector (Promega) conforme descrito item 3.12. Em seguida, foram excisadas desse vetor pelas endonucleases *Bgl*II e *Hind*III (New England Biolabs) conforme descrição do fabricante e foram inseridos no plasmídeo pGL3-Basic vector (Promega), previamente digerido com as mesmas endonucleases. As sequências dos promotores antissenso e senso foram inserida a montante da região do vetor pGL3 que codifica para a enzima *Firefly luciferase* do vagalume *Photinus pyralis*, que foi utilizada no ensaio de promotor. A ligação do produto da excisão no pGL3-Basic vector, transformação de bactérias competentes com o plasmídeos e extração dos vetores das bactérias foram realizadas conforme descrito no item 3.12.

Os ensaios de atividade promotora foram realizados na linhagem celular humana HeLa, cujas células foram semeadas em placas de 24 poços (4 x 10^5 células), e as transfecções foram realizadas com o reagente FuGENE 6 Transfection Reagent (Roche), segundo instruções do fabricante. As células foram transfectadas em triplicatas biológicas com um dos plasmídeos pGL3_promotor-antissenso, pGL3_promotor-senso, pGL3-SV40 (controle positivo) ou pGL3-vazio (controle negativo) e todas foram co-transfectadas com o plasmídeo pRL-SV40 (Promega). Este último é utilizado para normalização da eficiência de transfecção, segundo protocol do fabricante.

Para avaliação da atividade de promotor, foi utilizado o sistema *Dual-Luciferase Reporter Assay System* (Promega), segundo instruções do fabricante. As reações foram realizadas em placa branca opaca (Falcon) de 96 poços e as medidas de emissão de luz foram obtidas em luminômetro *MicroLumat Plus* (Microplate Luminometer LB – Berthold Technologies) com o programa *WinGlow* (Berthold Technologies) utilizando-se os seguintes parâmetros: intervalo de medida 1 de 10 s, intervalo de medida 2 de 1 s e temperatura de 25 °C. A razão do valor obtido para cada medida das luciferases *Firefly* e *Renilla* de cada poço

foi calculada (Unidade Relativa de luz), assim como a média aritmética da razão de três réplicas da transfecção com cada plasmídeo.

3.14 Super-expressão

O transcrito *ANRASSF1* foi amplificado por PCR conforme descrito item 3.6.1 utilizando oligonucleotídeos iniciadores específicos (Tabela 1) e como molde da amplificação foi utilizado cDNA proveniente de RT-fita específica do *ANRASSF1* conforme descrito no item 3.5.2. O produto da reação de PCR foi clonado no pGEM-T vector (Promega) conforme descrito no item 3.12. Em seguida, foi excisado o *ANRASSF1* desse vetor pelas endonucleases *Kpn*I e *Hind*III (New England Biolabs) conforme recomendação do fabricante e foi inserido no plasmídeo pCEP4 (Invitrogen), previamente digerido com as mesmas endonucleases. O vetor pCEP4 possui gene que confere resistência ao antibiótico Higromicina B, que foi utilizada na etapa de seleção das células transfectadas com os vetores. A ligação do produto da excisão no vetor pCEP4, transformação de bactérias competentes com o plasmídeos e extração dos vetor das bactérias foram realizadas conforme descrito no item 3.12.

Visando a super-expressão do *ANRASSF1* de maneira estável, os vetores pCEP4_*ANRASSF1* e pCEP vazio (controle) foram linearizados com a endonuclease *Nhe*I (New England Biolabs), conforme descrição fabricante. Células HeLa foram semeadas em placas de 6 poços ($5x10^5$ células) e mantidas em cultura até atingirem a confluência de 80%, e em seguida foram misturadas com 1,6 µg do vetor pCEP4 *ANRASSF1* ou do vetor pCEP, linearizado, diluído no reagente FuGENE 6 Transfection Reagent (Roche), segundo instruções do fabricante. Foram realizadas 3 transfecções independentes de cada construção, pCEP4 *ANRASSF1* ou pCEP vazio (como controle).

Após 48 horas da transfecção, as células foram tratadas com 200 ng/ml de Higromicina B (Gibco), para selecionar as células transfectadas com o vetor pCEP4 que

confere as células resistência a Higromicina B. Como controle da seleção uma placa com células HeLa submetidas às mesmas condições de transfecção sem adição de vetor, foi submetida ao tratamento de Higromicina B. A seleção ocorreu com meio de cultura com antibiótico de 4 em 4 dias. A seleção foi mantida até uma semana após as células utilizadas como controle da seleção da transfecção terem 100% de morte celular quando observadas ao microscópio *Axiovert 200* (Zeiss).

3.15 RNAi

Nos ensaios de supressão da expressão do lncRNA *ANRASSF1*, os oligonucleotídeos dupla-fita de RNA foram desenhados no BLOCK-iT™ RNAi Designer (Invitrogen) segundo protocolo do fabricante, utilizando somente os trechos de sequência que mapeiam nos introns do gene codificador de proteína *RASSF1A* do mesmo *locus* do lncRNA antissenso para evitar o silenciamento do transcrito maduro do gene codificador de proteína. Como controle negativo foram desenhados no mesmo programa siRNAs com a mesma composição de bases, porém em sequência embaralhada; as sequências dos siRNAs estão descritas na Tabela 3. As células HeLa foram transfectadas com um *pool* de 3 siRNA contra *ANRASSF1* ou um *pool* de 3 siRNA controle na concentração total de 15 nM (5 nM cada siRNA) utilizando lipofectamine RNAiMAX (Invitrogen) conforme descrição do fabricante. Após 48 h do tratamento o RNA total das células em cultura foi extraído conforme descrito no item 3.2 e avaliada a expressão dos alvos por RT-qPCR conforme descrito nos itens 3.5.1 e 3.6.2.

Tabela 3. Sequências usadas no ensaio de siRNA

Alvo	Nome do oligonucleotídeo	Sequência
ANRASSF1	stealth_ANRASSF_1	GGGAAAUCGGCAAUUAGAACGCUCC
ANRASSF1	stealth_ANRASSF_1	GGAGCGUUCUAAUUGCCGAUUUCCC
	stealth_controle_1	GGGCUACGGUAAGAUCAACGAAUCC
	stealth_controle_1	GGAUUCGUUGAUCUUACCGUAGCCC
ANRASSF1	stealth_ANRASSF_3	UGGAUCUCUAUCGCCUAGCACAGAA
ANRASSF1	stealth_ANRASSF_3	UUCUGUGCUAGGCGAUAGAGAUCCA
	stealth_controle_3	UGGUCUCGCUAAUCCCACGAUAGAA
	stealth_controle_3	UUCUAUCGUGGGAUUAGCGAGACCA
ANRASSF1	stealth_ANRASSF_2	CGACCUAUCUCAGUGGGUUACCUCA
ANRASSF1	stealth_ANRASSF_2	UGAGGUAACCCACUGAGAUAGGUCG
	stealth_controle_2	CGAAUCUCUGAGUGGUUCACCCUCA
	stealth_controle_2	UGAGGGUGAACCACUCAGAGAUUCG

3.16 *Western blot*

Células HeLa foram cultivadas em frascos apropriados até a confluência desejada, lavadas com PBS e descoladas da superfície do frasco por tratamento com 0,1% de tripsina seguido da adição de meio de cultura contendo SFB. Em seguida, as células foram sedimentadas por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. O sedimento de células foi lavado com PBS, novamente sedimentado por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. Em seguida, o sedimento de células foi ressuspenso em tampão de lise gelado (imidazol 20 mM (pH 7.2), EDTA1 mM, sacarose 250 mM com *complete protease inhibitor cocktail* (Roche)) e sonicado no aparelho Sonic Dismembrator Model 500 (Fischer Scientific), submetido a três pulsos de 10 segundos com amplitude de 30 % e intervalo de 30 segundos em gelo. A concentração das proteínas presentes no lisado foi determinada utilizando o *BCA protein assay* (Bio-Rad) conforme descrição do fabricante e a absorbância medida no espectrofotômetro SpectraMax Paradigm (Molecular Devices).

Quantidades iguais de proteínas (40 µg) foram separadas por SDS-PAGE com 10% de poliacrilamida, ao lado de um padrão de massa molecular de proteínas *Prestained Protein Marker* (Fermentas) e posteriormente transferidas para uma membrana de nitrocelulose Hybond-C Extra (Amersham Biosciences) utilizando o transferidor semi-seco TE 77 PWR (Amersham Biosciences).

As membranas foram bloqueadas em 10 ml de solução de bloqueio (400 mg BSA /10 mL PBS; Sigma) durante 1 hora. Em seguida, foram incubadas com os anticorpos primários: anti-RASSF1A (Abcam – ab23950; na diluição 1:7500) e anti-actina (Millipore – mab1501 na diluição 1:9000) em uma nova solução de bloqueio por 16 hrs, sendo em seguida lavadas com PBS Tween 0,1%, uma vez por 15 minutos e duas vezes por 5 minutos. Após as lavagens as membranas foram incubadas com o anticorpo secundário que é marcado com Alexa flúor 680 (Invitrogen) por 1 hora, e em seguida as lavagens descritas acima foram repetidas. O sinal de fluorescência foi analisado com *Odyssey Infrared Imaging System* (LI-COR Biosciences) e as intensidades foram quantificadas com *Odyssey Software Aplicativo v3* (LI-COR Biosciences). O valor de densitometria foi definido como a razão entre as médias das medidas das intensidades de RASSF1A normalizadas pela actina, em cada condição (super-expressando *ANRASSF1* ou controle) em um total de três réplicas biológicas.

A significância estatística da diferença entre as condições comparadas foi medida utilizando o teste t de *Student* (pareado, bicaudal, homocedástico), considerando-se como limiar de significância p-valor <0,05.

3.17 Ensaio de proliferação

Para determinar se as células HeLa super-expressando *ANRASSF1* ou depletadas de *ANRASSF1* por RNAi sofrem alteração do coeficiente de proliferação celular, por alteração dos níveis de *RASSF1A*, utilizamos o kit CellTiter 96 AQueous One Solution Cell

Proliferation Assay (Promega), um método colorimétrico de determinação do número de células viáveis, que foi realizado com duração de 48 hrs, conforme descrição do fabricante. Cada par de réplicas biológicas foi plaqueado na mesma placa de 96 poços e três réplicas foram testadas tanto nos ensaios de super-expressão como RNAi. O coeficiente de proliferação foi avaliado pela medida de absorbância de formazan nos tempos de 24 e 48 hrs, como recomendado no kit CellTiter 96 Aqueous One Solution Cell Proliferation Assay (Promega) no espectrômetro SpectraMax Paradigm (Molecular Devices). O coeficiente de proliferação foi definido como a razão entre as médias das medidas dos tempos de 48 hrs e 24 horas. No caso do ensaio de proliferação juntamente com RNAi (item 3.14), após 24 hrs da transfecção dos siRNA a proliferação foi avaliada conforme descrito acima. A significância estatística da diferença entre as condições comparadas foi medida utilizando o teste t de *Student* (pareado, bicaudal, homocedástico), considerando-se como limiar de significância p-valor <0,05.

3.18 Imunoprecipitação de complexos ribonucleoproteicos - RIP

Células HeLa foram cultivadas em frascos apropriados até a confluência desejada, lavadas com PBS e descoladas da superfície do frasco por tratamento com 0,1% de tripsina seguido da adição de meio de cultura contendo SFB. Em seguida, as células foram sedimentadas por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. Os sedimentos de células foram utilizados imediatamente.

A imunoprecipitação de complexos ribonucleoproteicos (RIP - *RNA-Binding Protein Immunoprecipitation*) foi realizada em $2x10^7$ células HeLa em triplicata, utilizando o Magna RIP RNA-Binding Protein Immunoprecipitation Kit (Millipore), segundo protocolo do fabricante.

Os seguintes anticorpos foram utilizados da Millipore anti-SUZ12 (03-179) e IgG de camundongo não imunizado (12-371). Os RNAs co-imunoprecipitados foram purificados por meio da utilização de fenol:clorofórmio:álcool isoamílico seguida de clorofórmio, e precipitados com etanol 100% a -80 °C por 16 horas. Os RNAs foram quantificados utilizando o reagente RiboGreen (Invitrogen) conforme protocolo do fabricante e sua qualidade avaliada conforme item 3.3. Os RNAs alvos co-imunoprecipitados foram avaliados quanto ao seu nível de expressão e de ligação nos complexos proteicos por RT-qPCR conforme descrito nos itens 3.5.1 e 3.6.2.

3.19 Imunoprecipitação da cromatina - ChIP

Para realizar a imunoprecipitação da cromatina (ChIP - *Chromatin Immunoprecipitation*) inicialmente foi adicionado ao meio de cultura das células HeLa formaldeído 1,1%, e as placas foram mantidas à temperatura ambiente por 10 min, a fim de promover ligações cruzadas para fixação das proteínas associadas à cromatina (*crosslink*). Em seguida, o meio foi incubado com glicina 125 mM por 5 min, para interromper a formação de ligações cruzadas, e, em seguida, removido. A partir desse passo seguimos o protocolo EZ-Magna ChIP A Chromatin Immunoprecipitation Kit (Millipore) que consistia em etapas de neutralização do formaldeído, lavagens, recuperação das células, fracionamento celular, lise nuclear, fragmentação da cromatina e a imunoprecipitação com os anticorpos específicos. Os seguintes anticorpos foram utilizados da abcam anti-H3K27me3 (ab6002), da Millipore anti-SUZ12 (03-179) e IgG de camundongo não imunizado (12-371). A avaliação da ligação das proteínas ao DNA foi realizada por qPCR descrito no item 3.6.2.

3.20 Ensaio de detecção de metilação do DNA

Células HeLa foram cultivadas em frascos apropriados até a confluência desejada, lavadas com PBS e descoladas da superfície do frasco por tratamento com 0,1% de tripsina seguido da adição de meio de cultura contendo SFB. Em seguida, as células foram sedimentadas por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. Os sedimentos de células foram utilizados imediatamente.

O DNA das células HeLa super-expressando *ANRASSF1* e controle foi extraído com fenol/clorofórmio e proteinase K conforme descrito por (Sambrook and Russell 2001). O DNA extraído (10 µg) foi fragmentado aleatoriamente por sonicação utilizando o aparelho Sonic Dismembrator Model 500 (Fischer Scientific), submetido a quatro pulsos de 10 segundos com amplitude de 11 % e intervalo de 30 segundos em gelo. A eficiência de sonicação foi avaliada por eletroforese em gel de agarose 2%. Em seguida, 1 µg do DNA sonicado foi tratado com 100 U da endonuclease dependente de metilação McrBC (New England Biolabs) a 37°C por 16 hrs, segundo recomendações do fabricante. Em paralelo como controle foi incubado 1 µg do DNA sonicado nas mesmas condições sem adição de enzima. A avaliação da quantidade de DNA não metilado foi realizada por qPCR (oligonucleotídeos iniciadores estão descritos na tabela 2) como descrito no item 3.6.2 em três réplicas biológicas.

3.21 Permeabilização celular e tratamento com RNase

Células HeLa foram cultivadas em frascos apropriados até a confluência desejada, lavadas com PBS e descoladas da superfície do frasco por tratamento com 0,1% de tripsina seguido da adição de meio de cultura contendo SFB. Em seguida, as células foram sedimentadas por centrifugação a 325 x g por 5 minutos e o sobrenadante foi removido. Os sedimentos de células foram utilizados imediatamente.

O tratamento com RNase foi realizado em $1,5x10^7$ células HeLa permeabilizadas em 1 mL de PBT (PBS; 0,05% Tween 20), posteriormente adicionado 1000 U/mL de RNase A (Sigma), 1000 U/mL de RNase H (Invitrogen) ou 1000 U/mL SUPERNase In (Ambion) inibidor de RNase e mantidas a 25^oC por 8 h sob agitação de 500 rpm no TermoMix (Eppendorff) conforme descrito em (Bertani et al. 2011). Após o tratamento de RNase, uma alíquota (70% vol) foi utilizada para realizar ChIP conforme descrito item 3.19 com os anticorpos da Milipore anti-SUZ12 (03-179), anti-RNA Pol II clone CTD4H8 (05-623B) e IgG de camundongo não imunizado (12-371). A restante alíquota (30% vol) foi utilizada para extração de RNA conforme descrito no item 3.2, sua qualidade averiguada conforme item 3.3, e avaliada a expressão dos alvos por RT-qPCR conforme descrito nos itens 3.5.1 e 3.6.2.

4. RESULTADOS

4.1 – Identificação do RNA não codificador antissenso ao gene *RASSF1* (*ANRASSF1- Antisense Noncoding RNA RASSF1)*

Em 2004 nosso grupo apontou para um conjunto de 24 lncRNAs dentro de regiões intrônicas, que foram identificados como diferencialmente expressos em 27 pacientes com diferentes graus de malignidade de tumor de próstata (Reis et al. 2004). Nesse conjunto, encontramos um lncRNA antissenso no segundo intron da isoforma *RASSF1A* e no primeiro intron da isoforma *RASSF1C*. Sua expressão foi confirmada por RT-PCR fita específica e *northern blot* e foi determinado que seu tamanho era de 1.1 kb. Esse transcrito antissenso apresentou uma maior expressão em tumor de alto grau de Gleason e menor expressão em tumores de baixo Gleason.

Posteriormente, em outro trabalho em 2007, nosso grupo identificou, por meio de buscas no banco de dados público de ESTs (*dbEST*), um conjunto de 67.731 possíveis longos ncRNAs *unspliced* que mapearam em regiões intrônicas de genes que codificam proteínas (Nakaya et al. 2007). Como o *locus* do gene *RASSF1* se mostrou interessante nos dados de microarranjo obtidos em 2004, fomos procurar se havia alguma predição de transcrição de lncRNAs nesse *locus*. Identificamos três lncRNAs no *locus* de *RASSF1* (Figura 2), sendo que dois desses lncRNAs (lncRNA_1 e lncRNA_2) tinham sobreposição com o transcrito antissenso identificado no trabalho de 2004 (Reis et al. 2004) (Figura 2). O outro lncRNA (lncRNA_3), representado por um cluster de ESTs de 580 pb, se localizava um pouco antes do início de transcrição da isoforma *RASSF1C* e tinha sobreposição com os introns 1 e 2 e exon 2 da isoforma *RASSF1A* (Figura 2).

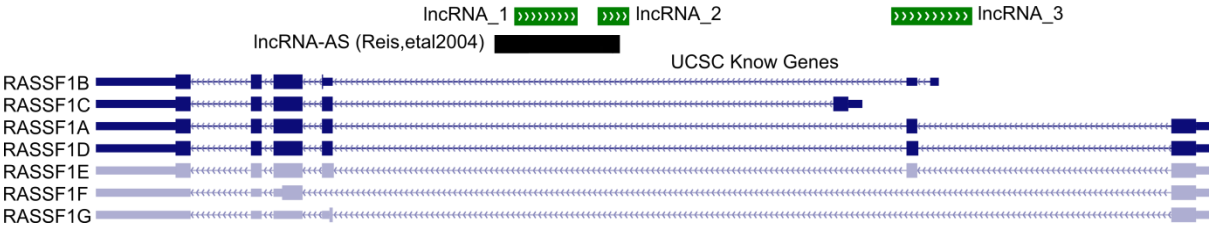

Figura 2 – Região genômica do gene *RASSF1*. (A) Em azul escuro são os *RefSeq genes* e azul claro *UCSC know genes*, em preto lncRNA antissenso identificado no trabalho de (Reis et al. 2004) e em verde predições de lncRNA identificados no trabalho de (Nakaya et al. 2007). Os retângulos mais grossos em azul representam exons codificadores, retângulos mais finos representam UTRs, linhas horizontais finas com setas representam introns e as setas mostram a orientação dos transcritos no genoma. O gene *RASSF1* está na fita – do genoma enquanto os lncRNAs estão na fita + do genoma.

A plataforma de microarranjo desenvolvida por nosso grupo em 2007 (Nakaya et al. 2007) continha sonda apenas para o lncRNA_1 do *locus RASSF1* e esta não apresentou diferença de expressão entre os três tecidos testados (fígado, próstata e rim). Com isso, como alternativa para avançar no estudo desse *locus*, olhamos se os microarrranjos comerciais da Affymetrix apresentavam sondas específicas para a região de interesse. Na plataforma *GeneChip* Human Genome *U133 Plus 2.0* Array (Affymetrix) identificamos sondas específicas para o lncRNA_3, que estão na orientação antissenso do gene codificador de proteína, além de sondas para o todo o *locus RASSF1*.

Dessa forma, realizamos uma análise *in silico* da co-expressão dos dois transcritos, com dados públicos desse microarranjo (GSE 5823 (Cappellen et al. 2007) e GSE 11118 (Byun et al. 2009)) e encontramos uma correlação negativa estatisticamente significante entre a expressão de *RASSF1* e a expressão de lncRNA_3 em 3 linhagens celulares (MCF-7, HeLa e MDA-MB-231) (Figura 3A). Adicionalmente, em células Jurkat submetidas a estresse mitótico, *RASSF1* mostrou um aumento de 1,5 vezes após uma hora de tratamento com ésteres de forbol e ionomicina e foi observada uma correlação negativa de sua expressão com a expressão do lncRNA_3 (Figura 3B). Nesse sentido, foi descrito que RASSF1, em conjunto com Daxx, constitui um *checkpoint* mitótico em situações de estresse, em que a células param

de se dividir e eventualmente morrem (Giovinazzi et al. 2012). Devido a correlação negativa observada entre o lncRNA_3 e *RASSF1*, levantamos a hipótese que este lncRNA poderia estar envolvido na regulação da expressão gênica desse *locus* e optamos por focar apenas no estudo desse transcrito.

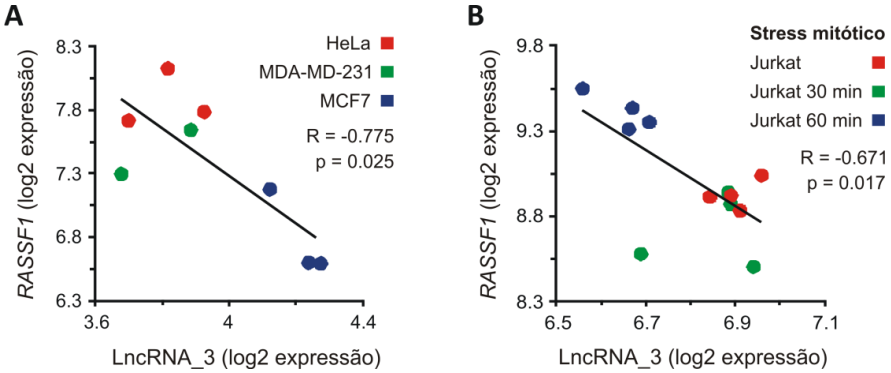

Figura 3 - Correlação inversa entre os genes *RASSF1* e lncRNA_3. (A) Expressão de *RASSF1* e lncRNA_3 obtidas por medida de microarranjos da Affymetrix em células HeLa, MDA-MB-231 e MCF-7, a partir de dados públicos GSE5823 (Cappellen et al. 2007). **(B)** Expressão de *RASSF1* e lncRNA_3 obtidas por medida de microarranjos da Affymetrix em células Jurkat submetidas a estresse mitótico induzido por tratamento com ésteres de forbol e ionomicina durante 30 e 60 min, a partir de dados públicos GSE11118 (Byun et al. 2009).

Uma vez que identificamos a expressão do lncRNA_3 com orientação antissenso em relação ao gene codificador de proteína no microarranjo da Affymetrix, e que as ESTs utilizadas para montar o lncRNA_3 foram originadas de bibliotecas de cDNA sem orientação e expressas em diversos tecidos humanos, buscamos confirmar este resultado por RT-PCR fita específica em diversas linhagens celulares.

Demonstramos que o lncRNA_3, de fato, era antissenso ao gene codificador de proteína e é expresso em 9 linhagens celulares (Figura 4A). Denominamos o lncRNA_3 como *Antisense Noncoding RNA RASSF1* (*ANRASSF1* – RNA não codificador antissenso ao gene *RASSF1*). Vale ressaltar que *ANRASSF1* também foi detectado nas linhagens celulares HeLA, MCF7 e MDA-MB-231, as mesmas presentes na meta-análise do microarranjo da Affymetrix (Figura 3A), onde se observou uma correlação negativa entre a expressão do gene *RASSF1* e de *ANRASSF1*.

Em seguida, realizamos a técnica de RACE (*Rapid Amplification of cDNA Ends*) para estender ambas as extremidades 5’ e 3’ do transcrito *ANRASSF1*. O produto do RACE-PCR para a extremidade 3' estendeu-se por 38 nt a partir da predição de ESTs existente, com 5 nt mapeando no genoma e com uma sequência adicional de 33 adeninas, correspondendo a uma possível cauda poli (A) (Figura 4B). Este achado é amparado pela identificação de um sinal conservado de poliadenilação (ATTAAA) a 17 nt a montante da cauda poli (A) nesta região, utilizando a ferramenta de predição de sinal de poliadenilação polyadq (Colgan e Manley, 1997; Tabaska e Zhang, 1999). Foram combinadas as abordagens de RACE-PCR para a extremidade 5' e caminhada com oligonucleotídeos iniciadores (*primer-walking* PCR), resultando na extensão de 205 nt na extemidade 5’ (Figura 4B, caixa vermelha), o que permitiu a descrição do transcrito completo de *ANRASSF1* com tamanho total de 790 nt.

Além disso, nosso grupo realizou sequenciamento em larga escala (com sequenciador de nova geração) de RNA poli(A)+ de células LNCaP, utilizando um método fita-específico, capaz de revelar se a transcrição se origina na fita mais ou na fita menos do genoma. Esses dados de sequenciamento foram utilizados para investigar a fundo o *locus* de *RASSF1*, e foi possível a confirmação do comprimento e da orientação antissenso de *ANRASSF1* (Figura 4C), como já foi apontado por RACE, RT-PCR fita específica (Figura 4B) e microarranjo Affymetrix.

Para excluir a possibilidade de que esse lncRNA tenha potencial para codificar proteínas, fizemos uma análise utilizando o programa CPC (*Coding Potential Calculator*) (Kong et al. 2007). Obtivemos como resultado que *ANRASSF1* apresenta potencial negativo de codificação. Além disso, o programa *ORF Finder* (Sayers et al. 2010) conseguiu predizer apenas dois peptídeos curtos, com 52 e 79 aminoácidos; nenhum dos peptídeos curtos preditos exibiu similaridade significativa com qualquer peptídeo conhecido ou domínio de proteína, portanto *ANRASSF1* foi considerado um lncRNA.

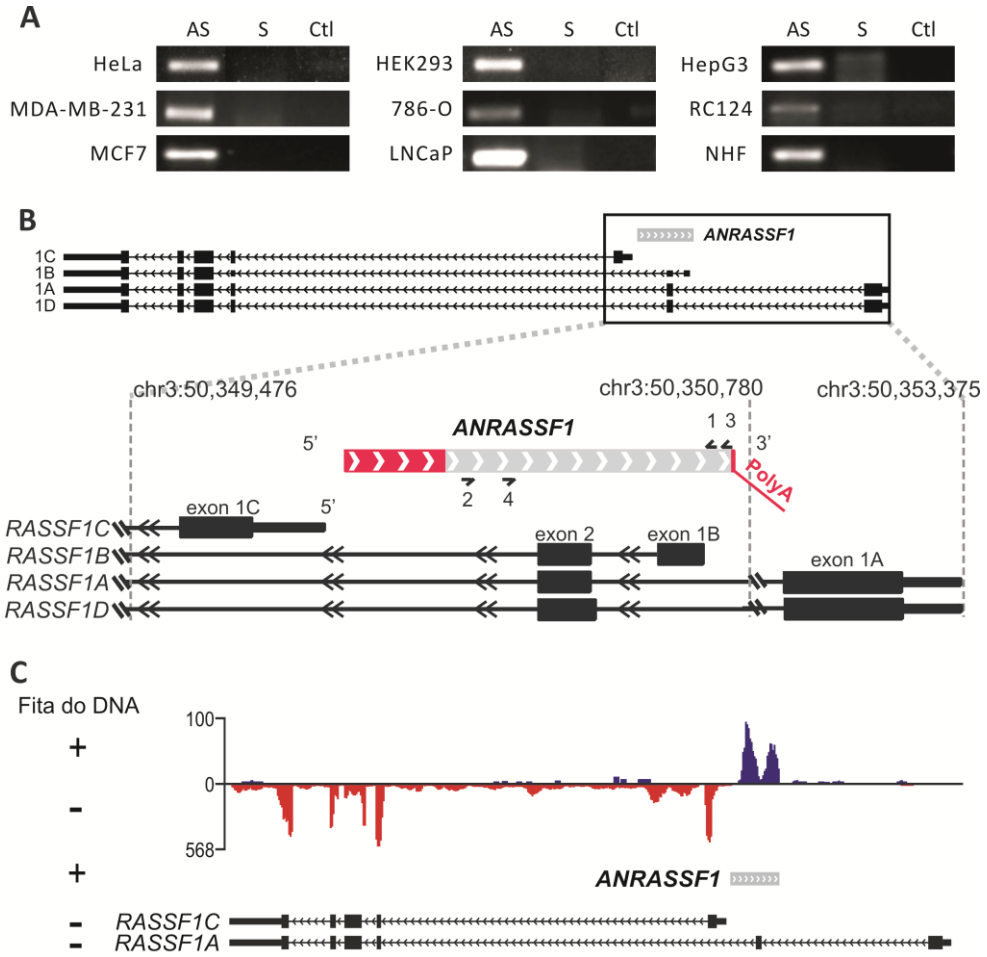

Figura 4 - RNA não codificador antissenso *ANRASSF1* é expresso no *locus* genômico de *RASSF1*. (A) *ANRASSF1* foi detectado em várias linhagens celulares humanas por PCR apenas na orientação antissenso (AS) e não na orientação senso (S), relativa ao gene codificador de proteína. Oligonucleotídeos iniciadores específicos foram usados para a reação de transcrição reversa (RT) (setas curtas de números 2 e 3 sob e sobre o *ANRASSF1*, respectivamente, no painel B). Reação de RT na ausência de oligonucleotídeo iniciador foi utilizada como controle negativo (Ctl). Setas curtas de números 1 e 4 (painel B) indicam oligonucleotídeos iniciadores usados para a PCR. **(B)** Representação esquemática do *locus RASSF1*. Isoformas do gene codificador de proteína *RASSF1* (RefSeq-anotadas) estão mostradas em preto; montagem de *ANRASSF1* a partir das evidências de ESTs públicas em cinza claro; porções estendidas por 5'- RACE e caminhada com oligonucleotídeos iniciadores (*primer-walking)* ou 3'- RACE estão em vermelho, com um segmento poli-A na extremidade 3'. Setas indicam a orientação das sequências. **(C)** RNA-seq fita-específico de RNA polyA+ de células LNCaP. Sequências mapeadas no *locus* genômico de *RASSF1* estão mostradas no diagrama, e as cores correspondem à fita de cada transcrito.

Em seguida, medimos a expressão do lncRNA *ANRASSF1* e do gene *RASSF1A* em linhagens celulares tumorais e normal de mama (Figura 5A) e de próstata (Figura 5B). Verificamos que um menor nível de expressão de *ANRASSF1* em linhagens de células

normais, em comparação com as linhagens tumorais em ambos os tecidos, e um padrão oposto para *RASSF1A*, que foi detectado com uma maior expressão em células normais, em comparação com as linhagens celulares tumorais. Dessa forma, a correlação inversa entre a expressão de *ANRASSF1* e a expressão de *RASSF1*, que encontramos nos dados públicos de microarranjos, foi confirmada aqui em duas linhas de células tumorais e uma linhagem de células normais em cada um dos dois diferentes tecidos, apontando para um possível papel na tumorigênese.

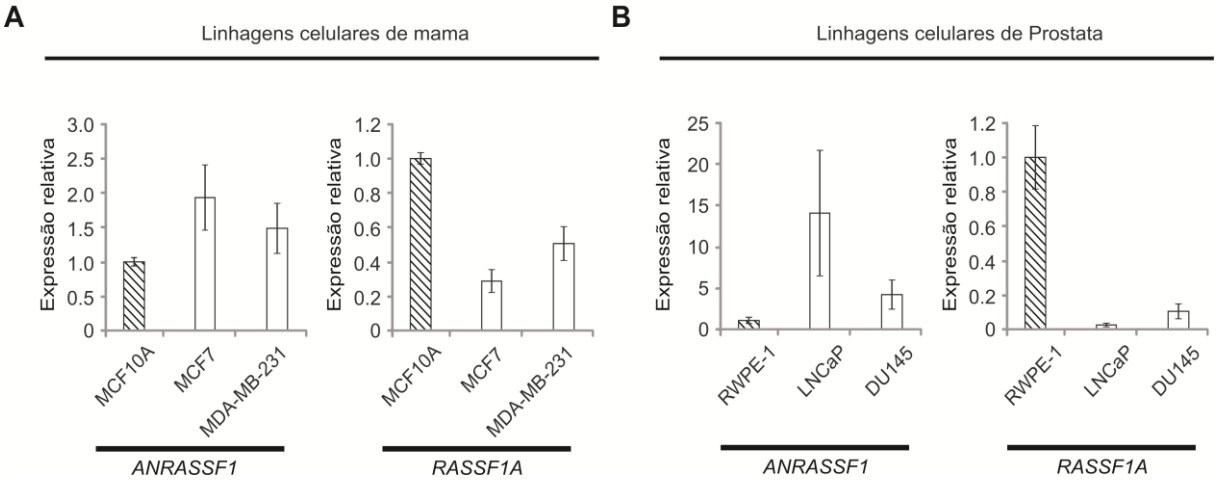

Figura 5 - Correlação inversa entre a expressão de *ANRASSF1* e *RASSF1A* em linhagens celulares normais e tumorais. Expressão de *ANRASSF1* e do gene *RASSF1A* em linhagens de (A) mama e (B) próstata. Linhagens celulares tumorais (barras brancas) e normais (barras hachuradas). Dados referentes à média de três ou duas réplicas biológicas de cada linhagem. Os valores de expressão nas linhagens tumorais são mostrados em relação à expressão detectada nas células normais; os dados foram normalizados pela expressão de *HPRT1*.

4.2 - Caracterização de *ANRASSF1*

Partimos para a caracterização de *ANRASSF1* quanto à sua biogênese, presença de cap 5', estabilidade, localização celular, presença de região promotora e abundância. Inicialmente, com o objetivo de responder se *ANRASSF1* é transcrito pela RNA Polimerase II (RNAP II), tratamos células HeLa com α- amanitina. A α-amanitina é um peptídeo cíclico de oito

aminoácidos que na concentração usada em nossos experimentos inibe a enzima RNAP II por bloquear sua translocação sobre o DNA durante a elongação da transcrição (Bushnell et al. 2002; Gong et al. 2004). Realizamos a extração de RNA de três réplicas biológicas de células HeLa tratadas e não-tratadas (controle) com α-amanitina (10 μg/ml) durante 9 h, seguindo essencialmente protocolos de trabalhos já publicados (Lee et al. 2004; Nakaya et al. 2007; Raha et al. 2010). Como controle negativo e normalizador do experimento (RT-qPCR), utilizamos o transcrito pre-$tRNA^{Tyr}$, que é sintetizado pela RNAP III, que não é sensível a baixas concentrações de α- amanitina, tais como as utilizadas nestes experimentos (Jacob et al. 1970; Kedinger et al. 1970; Lindell et al. 1970). Por RT-qPCR, identificamos que o *ANRASSF1* é transcrito pela RNAP II, visto que houve uma diminuição da sua abundância após o tratamento com α-amanitina (**Erro! Fonte de referência não encontrada.**6A). O mesmo resultado foi observado para o controle positivo, alfa tubulina (**Erro! Fonte de referência não encontrada.**6A), gene sabidamente transcrito pela RNAP II.

Para investigar se *ANRASSF1* possui a modificação 7-metil guanosina na extremidade 5' (cap 5'), característica de mRNAs transcritos pela RNAP II, foi utilizado o tratamento enzimático para remoção do cap 5' com a enzima *Tobacco Acid Pyrophosphatase* (TAP) e posterior digestão dos RNAs por uma enzima 5'-exonuclease (5'-Exo), que digere RNAs que possuem 5' - monofosfato. Os RNAs com a extremidade 5' protegida pela presença de cap 5' são resistentes à digestão pela 5'-Exo, enquanto RNAs sem cap 5' ou parcialmente degradados são digeridos pela enzima. Foram feitas duas condições de tratamento: i) amostra tratada com TAP e 5'-Exo (TAP^+/5'-Exo^+) e ii) amostra tratada apenas com 5'-Exo (TAP^-/5'-Exo^+). A amostra TAP^-/5'-Exo^+ foi submetida à incubação com o mesmo tampão e temperatura da amostra a ser tratada com a enzima TAP, previamente ao tratamento com a enzima 5´-Exo, o que assegura que as amostras em estudo (TAP^+/5'-Exo^+ e TAP^-/5'-Exo^+) ficassem sujeitas as

mesmas condições e portanto, susceptíveis ao mesmo nível de degradação inespecífica do RNA por RNases contaminantes.

A detecção da presença ou ausência de cap 5' via tratamento enzimático (TAP/5'-Exo) foi verificada através de PCR. A **Erro! Fonte de referência não encontrada.**6B mostra que *ANRASSF1* possui cap na sua extremidade 5´, por não ser degradado com tratamento apenas com 5'-Exo. Ao removermos o cap 5' com a enzima TAP e depois tratarmos essa fração com a 5'-Exo, observamos uma diminuição do produto amplificado em relação ao tratamento apenas com 5'-Exo. O mesmo foi observado no controle positivo com alfa tubulina, demonstrando que as condições de tratamento com as enzimas foram eficientes na remoção do cap 5´. O snoRNA (*small nucleolar* RNA), snRNA U15A, que não possui cap 5' (Tycowski et al. 1993), foi usado como controle negativo para o tratamento com as enzimas (**Erro! Fonte de referência não encontrada.**B).

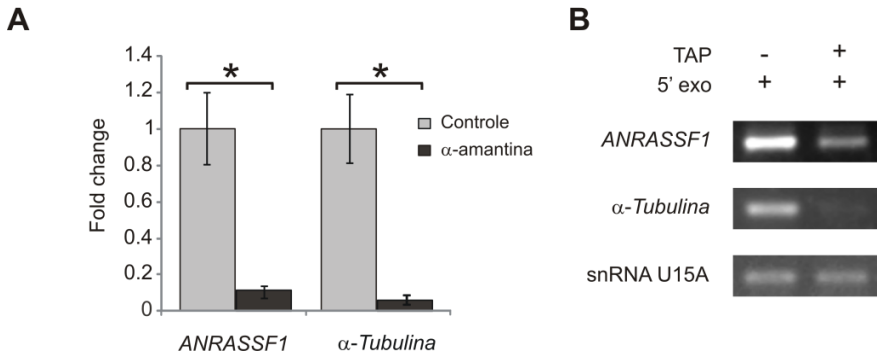

Figura 6 - *ANRASSF1* é transcrito pela RNAPII e possui a estrutura cap 5'. (A) Células HeLa foram tratadas com o inibidor de RNA Polimerase 2 (RNAPII) α-amanitina (barra preta) ou com o veículo (barra cinza) por 24 h. A abundância do transcrito *ANRASSF1* foi medida por RT-qPCR. Os resultados foram normalizados pelo nível de pre-tRNA^{tyr}, que é transcrito pela RNAPIII. O mRNA de *α-tubulina*, que é transcrito pela RNAPII, foi utilizado como controle positivo. Foram realizadas três réplicas biológicas; barras de erro referem-se ao desvio padrão. Asteriscos indicam diminuição significativa de *ANRASSF1* após o tratamento com o inibidor de RNAPII (teste-t, $p < 0,0001$). (B) RNA total de células HeLa foi digerido exclusivamente com 5'-exonuclease (5'-Exo) ou em combinação com tobacco acid pyrophosphatase (TAP), como indicado. Subsequentemente, as amostras foram submetidas à transcrição reversa seguida de PCR, com oligonucleotídeos iniciadores para *ANRASSF1 (*27 ciclos de PCR), *α-tubulina* (controle positivo, 14 ciclos de PCR) ou snRNA *U15A*, que não possui 5'cap (controle negativo, 22 ciclos de PCR), como indicado.

Para estudar a estabilidade de *ANRASSF1*, medimos a mudança de sua expressão após a inibição da transcrição com actinomicina D, durante cinco diferentes tempos de tratamento (0, 1, 3, 6 e 8 h). A actinomicina D é um antibiótico polipeptídico cíclico, que forma um complexo estável com o DNA fita simples no complexo de iniciação da transcrição, bloqueando a elongação da transcrição pelas RNA polimerases (Sobell 1985). Para este experimento, foi extraído RNA de duas réplicas biológicas independentes de células HeLa tratadas com actinomicina D ou com seu veículo (DMSO; controle) e realizado RT-qPCR em cada condição e tempo. A partir destas medidas foi calculada a meia-vida de *ANRASSF1* por meio do modelo de *one-phase exponential decay* (Graphpad 5.02), segundo descrito em (Clark et al. 2012). A meia vida de *ANRASSF1* foi de 53 min (Figura 7A) e a meia-vida do mRNA *c-Myc*, controle do experimento, foi de 20 min (Figura 7A), compatível com medidas de estabilidade já descritas na literatura para este mRNA, em torno 15 min (Dani et al. 1984; He et al. 2000).

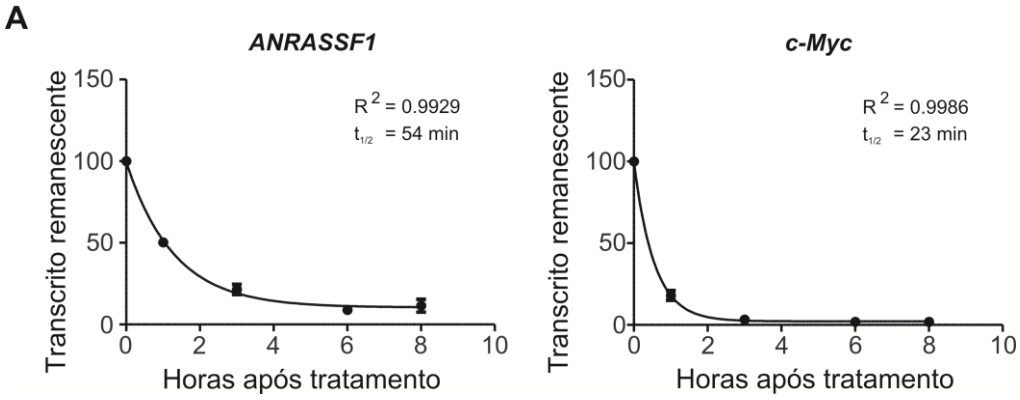

Figura 6 - Meia-vida de *ANRASSF1* (A) Estabilidade do transcrito *ANRASSF1* em células HeLa. As células foram tratadas com o inibidor transcricional actinomicina D ou com o veículo por 0, 1, 3, 6 e 8 h. Os níveis de *ANRASSF1* e *c-Myc* foram medidos por RT-qPCR e normalizados por uma amostra não tratada. Foram realizadas duas réplicas biológicas; barras de erro referem-se ao desvio padrão. A meia-vida ($t_{1/2}$) obtida para cada uma das curvas ajustadas usando *one-phase exponential decay* está mostrada.

Para identificar a localização sub-celular de *ANRASSF1*, foram realizados experimentos de fracionamento celular para obtenção de frações nucleares e citoplasmáticas de células HeLa, por meio de lise em condições que rompem a membrana plasmática, mas mantém o envelope nuclear intacto. Através de eletroforese capilar, foi observado um sensível enriquecimento do precursor rRNA 32S na fração nuclear em relação à fração citoplasmática, atestando a boa qualidade do fracionamento (Figura 8A). Este resultado mostrou-se semelhante ao obtido com o método original de fracionamento descrito na literatura (Gondran et al. 1999) e indica que a contaminação nuclear nas frações citoplasmáticas obtidas é ausente ou muito limitada. Além disso, como controle da qualidade do fracionamento, foi realizado RT-qPCR de amostras de RNA obtidas de cada compartimento, utilizando oligonucleotídeos iniciadores para o ncRNA *MALAT1* (Hutchinson et al. 2007), que é sabidamente retido no núcleo. A quantificação feita por RT-qPCR mostra uma detecção do *ANRASSF1* 100 vezes maior no núcleo em relação ao citoplasma (Figura **8**8B).

Para avaliar a abundância relativa de *ANRASSF1*, comparamos seu nível de expressão com os níveis de expressão de alguns lncRNAs descritos na literatura como muito abundantes, e com os níveis de expressão das duas principais isoformas do gene codificador de proteína do mesmo *locus, RASSF1A* e *RASSF1C*, usando RT-qPCR. Foi possível observar que *ANRASSF1* apresenta uma baixa expressão quando comparado com os lincRNAs *MALAT1* (Ji et al. 2003), *HOTAIR* (Rinn et al. 2007) e lincRNA *SFPQ* (Khalil et al. 2009) e com os genes do *locus RASSF1* (Figura **8**8C). Este último achado, em parte, já havia sido observado nos dados de intensidade das sondas específicas para *ANRASSF1* e *RASSF1* do microarranjo da Affymetrix (Figura 3).

Por último, identificamos um possível promotor bidirecional na região que abrange os sítios de início de transcrição de *ANRASSF1* e *RASSF1C*, que foram denominados promotor antissenso e promotor senso, respectivamente (Figura 9A). Estes promotores foram preditos *in*

silico no trabalho de (Davuluri et al. 2001), disponível como *track* no *UCSC Genome Browser*. Para avaliação da atividade promotora, foi utilizado o sistema *Dual-Luciferase Reporter Assay System*, que realiza a medida da atividade da enzima *Firefly luciferase*, seguida da inativação desta enzima e posterior medição da atividade da *Renilla luciferase*.

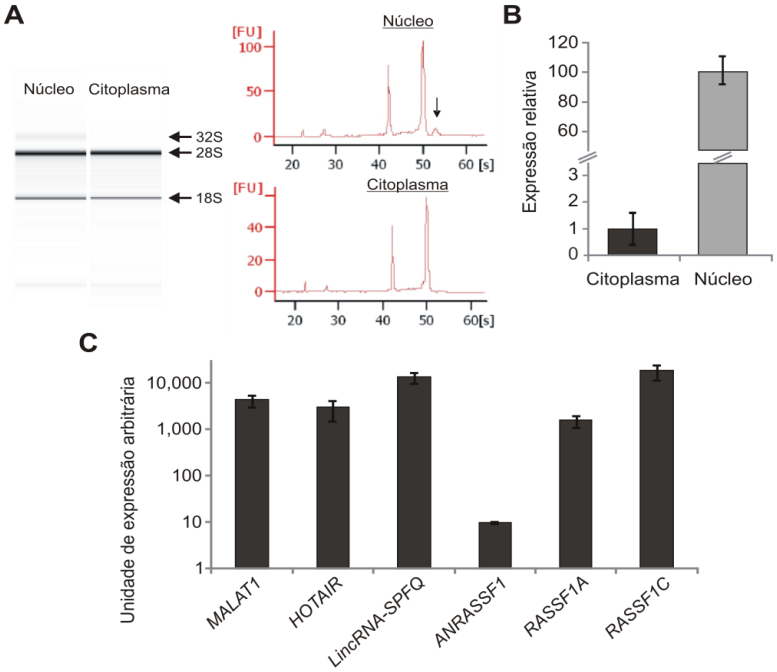

Figura 8 - *ANRASSF1* é enriquecido na fração nucelar e possui baixa abundância comparado com outros lncRNAs (A) Os painéis representam o perfil analítico das preparações de RNA total de citoplasma e núcleo após tratamento com DNAse, obtidos a partir de eletroforese capilar no Agilent 2100 Bioanalyzer (Agilent Technologies). O painel à esquerda mostra a imagem de eletroforese virtual reconstruída após análise no Bioanalyzer. O painel à direita apresenta o eletroferograma, com os dois picos maiores representando os rRNAs 18S e 28S (da esquerda para a direita, respectivamente). A seta indica a presença do precursor 32S rRNA, observado apenas na fração nuclear. **(B)** Abundância relativa do transcrito *ANRASSF1* nos compartimentos nuclear/citoplasmático. Frações celulares enriquecidas em RNAs nucleares ou citoplasmáticos foram preparadas em duplicata; barras de erro referem-se ao desvio padrão. Quantidades iniciais comparáveis de cada fração foram usadas nas RT-qPCR para medir *ANRASSF1*. **(C)** Níveis de expressão de lncRNAs abundantes e das principais isoformas de *RASSF1* relativo a *ANRASSF1*. Os níveis de expressão foram medidos em células HeLa por RT-qPCR e normalizados por *ANRASSF1*. Está mostrada a média de três experimentos independentes; barras de erro referem-se ao desvio padrão.

Esta segunda medida é utilizada para a normalização dos valores da atividade *Firefly luciferase*, a fim de se reduzir as variações decorrentes de diferenças nas eficiências de transfecção. Os promotores candidatos foram clonados no plasmídeo pGL3, a montante do gene da *luciferase*, e a quantificação da atividade do promotor foi feita medindo-se a intensidade luminosa emitida pela *luciferase* após a reação com seu substrato. Também foram utilizados nos ensaios de atividade de promotor, como controle positivo e negativo, os plasmídeos pGL3-SV40 e pGL3-vazio, respectivamente.

A atividade do possível promotor antissenso do *ANRASSF1* foi detectada (Figura 9B) em HeLa, bem como a atividade do promotor de *RASSF1C* (Figura B). Esse resultado confirma que o promotor é bidirecional, sendo responsável pela transcrição de *ANRASSF1* e de *RASSF1C*. Principalmente, o resultado mostra a detecção de um promotor específico para *ANRASSF1*, o que indica que esse lncRNA representa uma unidade transcricional independente.

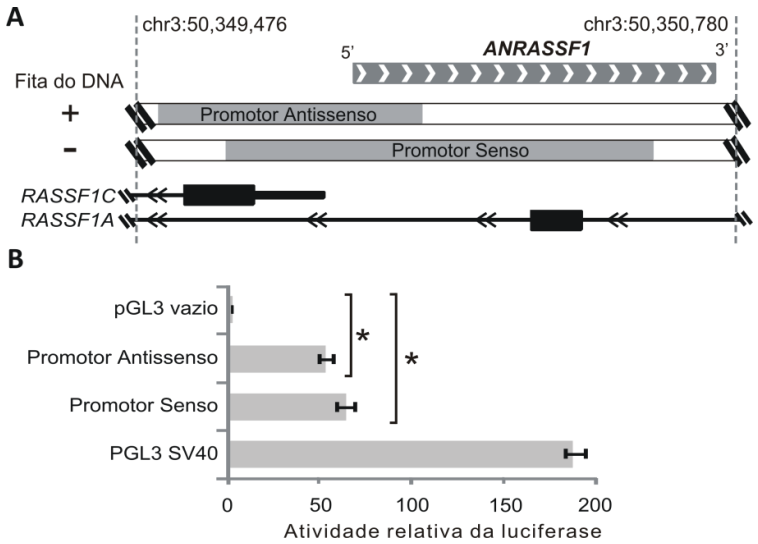

Figura 9 – Ensaio de atividade do promotor de *ANRASSF1*. (A) Localização genômica e tamanho das construções do promotor clonado *upstream* ao gene da Luciferase. **(B)** Atividade das diversas construções medidas em unidades relativas de luz. Os plasmídeos pGL3-Control e pGL3-Basic foram usados como controles positivo e negativo, respectivamente.

4.3 – Abundância de *ANRASSF1* modula a expressão de *RASSF1A* e afeta a taxa de proliferação de células em cultura

Visto que encontramos uma correlação negativa em células HeLa entre a expressão de *ANRASSF1* e a expressão do gene *RASSF1*, testamos se a abundância do lncRNA poderia modular a expressão do gene codificador de proteína no mesmo *locus* (*RASSF1A* e *RASSF1C*), atuando em *cis*. Inicialmente, o transcrito antissenso foi clonado no plasmídio de expressão em mamíferos pCEP4 (pCEP4-*ANRASSF1*), e foi transfectado de maneira estável (integrado no genoma) em 3 réplicas biológicas de células HeLa. Realizamos também uma transfecção com o vetor vazio linearizado (pCEP4-vazio) para ser utilizada como controle negativo do experimento. A seleção da população de células resistentes foi feita com Higromicina B, um antibiótico que inibe a síntese proteica através da interrupção da translocação da sub-unidade 80S do ribossomo e também por causar erros no processo de tradução. Após a seleção, a medida dos níveis de expressão de *ANRASSF1*, *RASSF1A* e *RASSF1C* foram realizadas por RT-qPCR nas células transfectadas com pCEP4-*ANRASSF1* ou pCEP4-vazio.

As células HeLa transfectadas com pCEP4-*ANRASSF1* apresentaram níveis de expressão do *ANRASSF1*, em média, 40 vezes maior quando comparado ao controle (Figura 10A), células HeLa transfectadas com pCEP4-vazio, confirmando a super-expressão do lncRNA. Nas células super-expressando o *ANRASSF1* foi observado que os níveis endógenos de *RASSF1A* diminuem para 20% quando comparados com células controle (Figura 10B). Interessantemente, a abundância da isoforma *RASSF1C* não foi afetada pela super-expressão do lncRNA (Figura 10C) .

Em paralelo, os níveis da proteína RASSF1A foram determinados por *western blot* nas duas linhagens de células. Observamos que houve uma redução na quantidade de proteína RASSF1A nas células transfectadas com pCEP4-*ANRASSF1*, quando comparadas com as células controle (Figura 10D). Análise de densitometria óptica de três réplicas de *western blot*

revelou que as células transfectadas com pCEP4-*ANRASSF1* apresentaram uma redução de 57% nos níveis da proteína RASSF1A, quando comparadas com os níveis endógenos das células controle (Figura 10E).

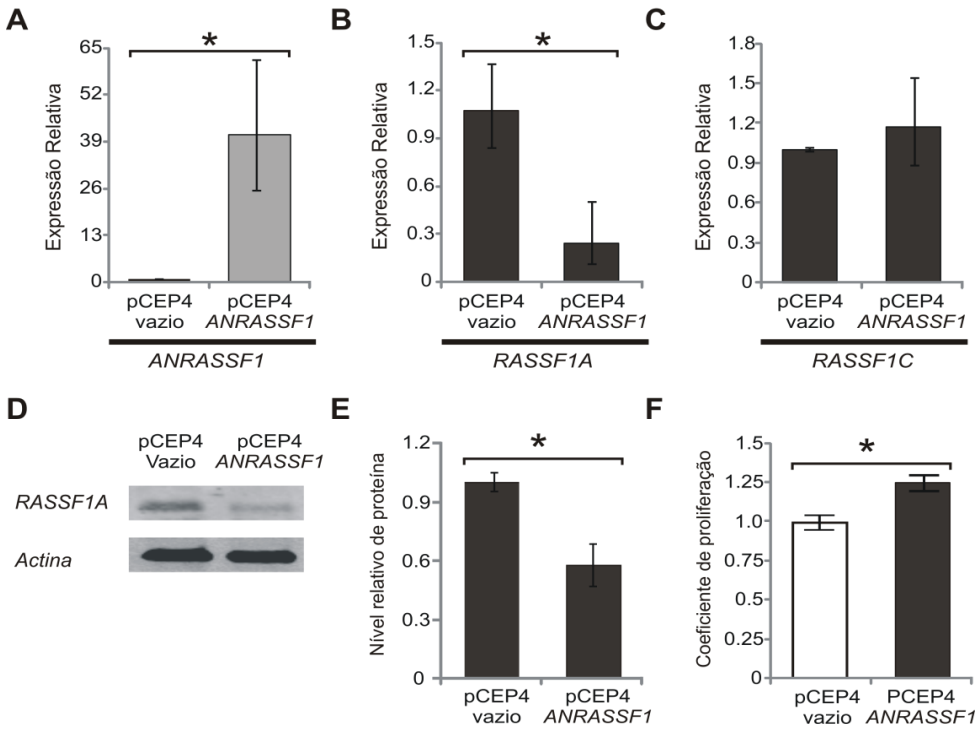

Figura 10 - Super-expressão de *ANRASSF1* promove a diminuição da isoforma *RASSF1A* e aumenta a proliferação celular. (A) Expressão de *ANRASSF1* em células HeLa que foram transfectadas com vetor contendo *ANRASSF1* (pCEP4-*ANRASSF1*) relativa a expressão em células HeLa transfectadas com o vetor vazio (pCEP4-vazio). Expressão relativa foi detectada por RT-qPCR e normalizada por α-tubulina; dados referentes à média de três experimentos de transfecção independentes; barras de erro referem-se ao desvio-padrão. O asterisco é indicativo de diferença estatisticamente significante (teste-t, $p < 0,01$, n=3). **(B)** Expressão relativa da isoforma *RASSF1A* em células transfectadas com *ANRASSF1* e células controle, como descrito em (A). O asterisco é indicativo de diferença estatisticamente significante (teste-t, $p < 0,01$, n=3). **(C)** Expressão relativa da isoforma *RASSF1C* em células transfectadas com *ANRASSF1* e células controle, como descrito em (A). **(D)** *Western blot* do lisado proteico de células HeLa controle (pCEP4-vazio) ou de células transfectadas com *ANRASSF1* (pCEP4-*ANRASSF1*), utilizando anticorpo contra a proteína RASSF1A. Anticorpo contra actina foi usado como controle. **(E)** Análises de densitometria do *western blot* mostrado em (D), onde os níveis proteicos relativos de RASSF1A foram normalizados pela actina; os dados são referentes à média de três réplicas biológicas; barras de erro referem-se ao desvio-padrão. O asterisco é indicativo de diferença estatisticamente significante (teste-t, $p < 0,02$, n=3). **(F)** Coeficiente de proliferação de células transfectadas com *ANRASSF1* relativo a células controle. Coeficiente de proliferação foi definido como a média das medidas dos pontos de 48 horas dividida pela média de 24 horas. Os ensaios foram feitos em triplicata; barras de erro referem-se ao desvio-padrão. O asterisco indica a existência de uma taxa de proliferação estatisticamente significante maior em pCEP4-*ANRASSF1* comparado ao controle (teste-t, $p < 0,03$, n=3).

Para checar se a diminuição da quantidade da proteína RASSF1A, causada pela super-expressão de *ANRASSF1*, resultaria em alterações detectáveis ao nível celular (alterações fenotípicas), foram realizados ensaios de proliferação. Foi observado que as células super-expressando *ANRASSF1* proliferaram significativamente mais rápido do que as células controle (Figura 10F), com um aumento médio de 1,22 vezes na taxa de proliferação ($p < 0,03$). Esse aumento observado na taxa de proliferação, em consequência da diminuição na abundância RASSF1A, está de acordo com a função conhecida desta proteína de atuar como supressor tumoral.

Posteriormente, para melhor documentar o efeito de *ANRASSF1* sobre a abundância da expressão de gene *RASSF1A*, foram desenhados oligonucleotídeos de RNA específicos para o silenciamento de *ANRASSF1* por siRNAs. Células HeLa foram transfectadas com um conjunto de três siRNAs específicos para *ANRASSF1* (siRNA-*ANRASSF1*) ou com um conjunto de três siRNA controle, com sequência embaralhada. O conjunto siRNA-*ANRASSF1* diminuiu significativamente a expressão de *ANRASSF1* para 39% do seu nível endógeno, em comparação com o controle (Figura 11A). Em paralelo, foi observado um aumento de 2,25 vezes na expressão do gene *RASSF1A* (Figura 11B). Por outro lado, a abundância da isoforma *RASSF1C* não foi afetada pela diminuição de *ANRASSF1* (Figura 11C), o que é análogo ao que foi descrito mais acima, quando *RASSF1C* não sofreu efeito aparente com a super-expressão de *ANRASSF1*.

Em seguida, foram realizados ensaios de proliferação celular para testar se o silenciamento de *ANRASSF1*, e o consequente aumento de seu nível de expressão, acarretaria em uma diminuição na taxa de proliferação da célula. Foi observado que as células nas quais *ANRASSF1* foi silenciado proliferavam significativamente mais lentamente do que as células controle (Figura 11D), com uma diminuição média de 15% na taxa de proliferação em comparação com as células controle ($p < 0,02$).

Em conjunto, os dados de super-expressão e silenciamento de *ANRASSF1* apontam para um mecanismo de modulação da expressão de *RASSF1A* mediado por *ANRASSF1*.

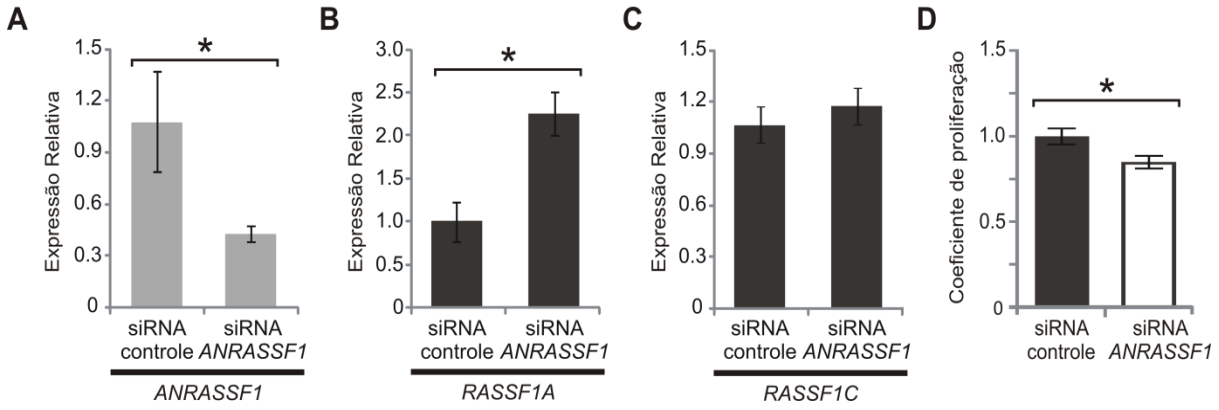

Figura 7 - Efeito da depleção de *ANRASSF1* na expressão das isoformas de *RASSF1* e na proliferação celular (A) Níveis de expressão de *ANRASSF1* em células tratadas com siRNA-*ANRASSF1* relativo a expressão de *ANRASSF1* em células controle tratadas com siRNA-controle. Expressão relativa detectada por RT-qPCR e normalizada por α-tubulina referente à média de três experimentos independentes de transfecção; barras de erro referem-se ao desvio-padrão. O asterisco indica diferença estatisticamente significante (teste-t, $p < 0,01$, n=3). **(B)** Níveis de expressão relativa de *RASSF1A* em células tratadas com siRNA-*ANRASSF1* e em células tratadas com siRNA-controle, como descrito em (A). O asterisco é indicativo de diferença estatisticamente significante (teste-t, $p < 0,01$, n=3). **(C)** Níveis de expressão relativa de *RASSF1C* em células tratadas com siRNA-*ANRASSF1* e em células tratadas com siRNA-controle, como descrito em (A). **(D)** Coeficiente de proliferação de células tratadas com siRNA-*ANRASSF1* relativo a células controle (siRNA-controle). Coeficiente de proliferação foi definido como a média das medidas dos pontos de 48 horas dividida pela média de 24 horas. Os ensaios foram feitos em triplicata, e é mostrada a média; barras de erro referem-se ao desvio-padrão. O asterisco indica a existência de uma taxa de proliferação estatisticamente significante e menor em células tratadas com siRNA-*ANRASSF1* comparado às células controle (teste-t, $p < 0,02$, n=3).

4.4 – A trimetilação da histona H3K27 no promotor de *RASSF1A* é dependente do nível de *ANRASSF1* ligado à PRC2

Procuramos ter uma visão mecanística sobre o envolvimento do lncRNA *ANRASSF1* na regulação da expressão *RASSF1A*. Já foi descrito que *RASSF1A* é regulado epigeneticamente por metilação do DNA de sua região promotora (Dammann et al. 2000). Como vimos que *ANRASSF1* se localiza no núcleo de HeLa e que sua super-expressão

provoca a diminuição da expressão de *RASSF1A*, existe uma possibilidade que este lncRNA possa estar envolvido no recrutamento de algum dos fatores regulatórios (MYC/PRC2) associados ao silenciamento de *RASSF1A* por meio da metilação do DNA do seu promotor ou por trimetilação da histona H3K27 (Palakurthy et al. 2009). Adicionalmente trabalhos recentes têm mostrado que uma fração dos lncRNAs retidos no núcleo estão envolvidos com o recrutamento de complexos remodeladores da cromatina, incluído *Polycomb* (Khalil et al. 2009; Tsai et al. 2010; Guttman et al. 2011).

Com isso, inicialmente, avaliamos a possibilidade de *ANRASSF1* endógeno encontrar-se fisicamente associado com SUZ12, proteína integrante do PRC2 (Nayak et al. 2011), e investigamos se esta associação poderia estar relacionada ao recrutamento de PRC2 para a região promotora de *RASSF1A*. Realizamos o ensaio denominado *RNA-Binding Protein Immunoprecipitation* (RIP), imunoprecipitando PRC2 em células transfectadas com pCEP4-vazio, condição expressando níveis endógenos do lncRNA, e medindo a eventual presença de *ANRASSF1* no precipitado. Para isso utilizamos o anticorpo anti-SUZ12 e o anticorpo IgG de camundongo não imunizado, como controle negativo de imunoprecipitação inespecífica. Detectamos um enriquecimento de 20 vezes de *ANRASSF1* endógeno na fração de RNA imunoprecipitada com anticorpo específico para SUZ12, quando comparado com a fração controle, que foi imunoprecipitada com IgG não-específico (Figura 8). Como controle positivo foi utilizado o lincRNA *SFPQ*, que é conhecido por se ligar a SUZ12 (Khalil et al, 2009) e este exibiu um enriquecimento de 32 vezes comparado ao controle (Figura 82). Estes resultados confirmam que *ANRASSF1* estava fisicamente associado com PRC2.

Como é sabido que SUZ12 se liga à região promotora de *RASSF1A* (Palakurthy et al. 2009), investigamos se a super-expressão de *ANRASSF1* poderia alterar a quantidade de SUZ12 ligada a esta região. Realizamos imunoprecipitação da cromatina (ChIP – *Chromatin Immunoprecipitation*), em células super-expressando *ANRASSF1* e em células controle

(*ANRASSF1* endógeno, pCEP4-vazio), utilizando o anticorpo anti-SUZ12 e o anticorpo IgG de camundongo não imunizado, como controle negativo. Medimos em cada condição, por qPCR, a presença de DNA da região promotora de *RASSF1A* na fração imunoprecipitada. Foi observado um aumento de 36 vezes do DNA do promotor de *RASSF1A* na fração anti-SUZ12 de células super-expressando *ANRASSF1* em relação à células controle (**Erro! Fonte de referência não encontrada.**A), evidenciando o aumento da ligação de SUZ12 no promotor com aumento de *ANRASSF1*. Além disso, nenhuma mudança na ocupação por SUZ12 foi observada com a super-expressão de *ANRASSF1* tanto na região promotora da isoforma *RASSF1C* como no promotor de cada um de dois genes vizinhos, de cada lado do *locus RASSF1* (Figura 13A). Como um controle adicional, nenhuma alteração foi detectada no promotor de *HOXA9*, um gene que é sabidamente regulado por PRC2 (Cao e Zhang, 2004). Estes resultados indicam que existe uma correlação entre níveis mais elevados de *ANRASSF1* e uma alta ocupação de PRC2 no promotor de *RASSF1A*, e também apontam para uma modulação epigenética local e altamente específica.

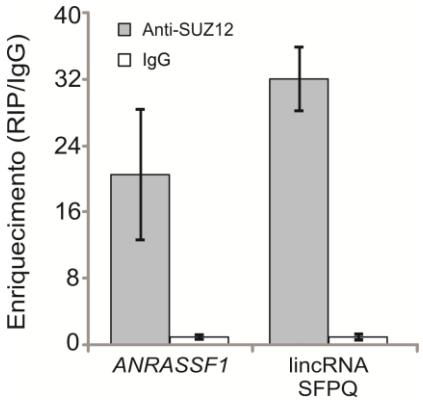

Figura 8 - *ANRASSF1* interage com PRC2. Níveis endógenos de *ANRASSF1* ligado a PRC2 foram medidos por RIP em células HeLa controle usando anticorpo anti-SUZ12 seguido de RT-qPCR com oligonucleotídeos iniciadores específicos para *ANRASSF1*. Como controle negativo, foi feito RIP com anticorpo anti-IgG não específico. LincRNA *SFPQ* foi utilizado como controle positivo para RIP com anti-SUZ12 (Khalil et al., 2009). Enriquecimento foi calculado pela comparação das médias das amostras com anti-SUZ12 contra amostras com IgG, RIP não-específico; os valores foram normalizados pela quantidade de *ANRASSF1* presente no *input*. Os ensaios foram realizados em triplicata e é mostrada a média ± desvio-padrão.

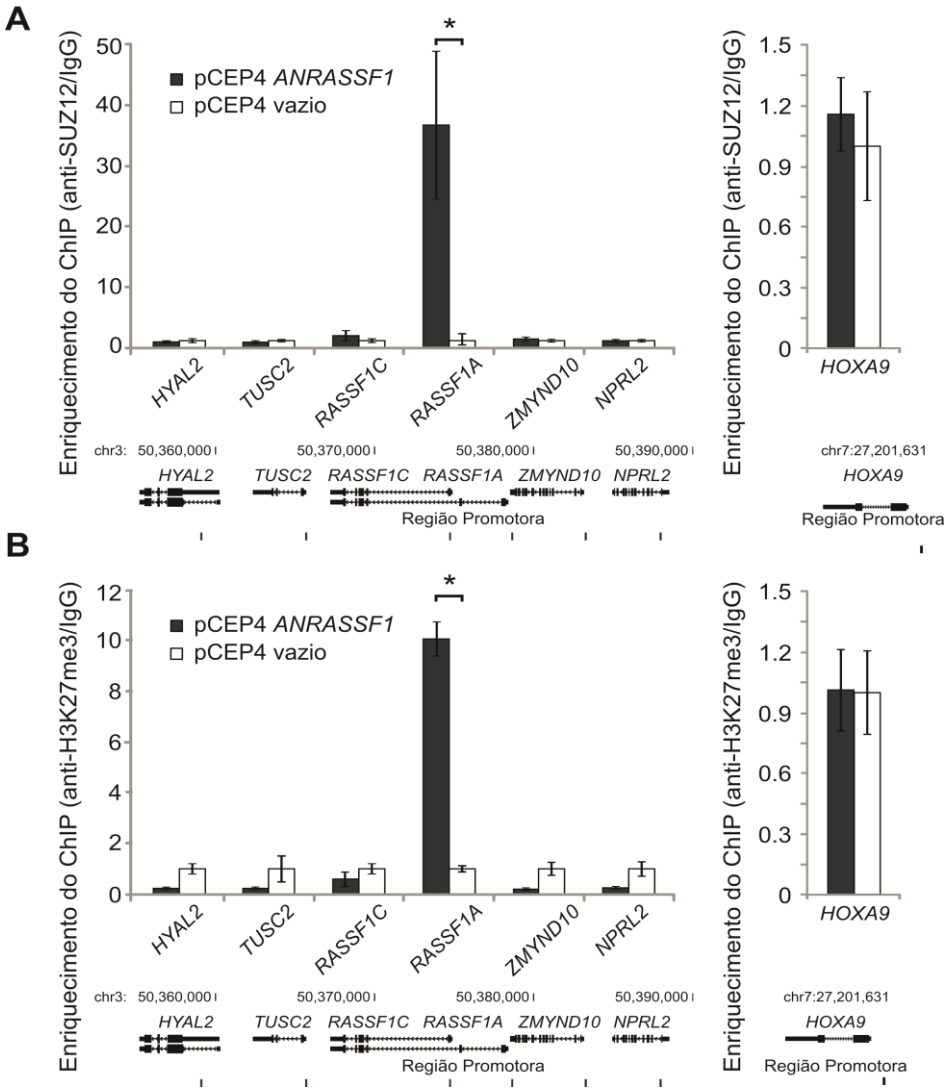

Figura 93 - *ANRASSF1* afeta especificamente a ocupação de PRC2 e a metilação da histona H3K27 no promotor de *RASSF1A*. (**A**) Ensaio de ChIP realizado com o anticorpo anti-SUZ12 em células HeLa transfectadas com *ANRASSF1* (pCEP4-*ANRASSF1*, barras pretas) ou em células controle (pCEP4-vazio, barras brancas). Foram investigadas as regiões promotoras das isoformas *RASSF1A* e *RASSF1C* e de outros dois genes de cada lado do *locus* de *RASSF1* (os promotores estão marcados pelas linhas verticais no esquema na parte inferior da figura). Como controle *HOXA9* foi utilizado como um gene que é sabidamente regulado por PRC2 (Cao and Zhang 2004). Enriquecimento foi calculado pela comparação dos valores de qPCR em amostras com anti-SUZ12 contra amostras com IgG não-específico; os valores foram normalizados pela quantidade de DNA amplificado por qPCR da região promotora de *RASSF1A* no *input*. Os ensaios foram realizados em triplicata e é mostrada a média ± desvio-padrão. Asterisco indica diferença estatisticamente significante (teste-t, $p < 0,04$, n=3). (**B**) Ensaio de ChIP com o anticorpo anti-H3K27me3, em um experimento semelhante ao descrito em (A). O asterisco indica diferença estatisticamente significante (teste-t, $p < 0,01$, n=3).

Como PRC2 é um complexo enzimático responsável pela di- e trimetilação da histona H3K27, testamos se o aumento da ocupação do promotor de *RASSF1A* por PRC2, causado pelo aumento dos níveis de *ANRASSF1*, aumentaria a quantidade de H3K27me3 nesta região. Observamos que ocorre um aumento de 10 vezes de H3K27me3 na região do promotor de *RASSF1A*, em células super-expressando o lncRNA (**Erro! Fonte de referência não encontrada.**3B), quando comparado com células controle, confirmando que a trimetilação da H3K27 no promotor de *RASSF1A* é dependente do nível de ligação de *ANRASSF1* à PRC2. Nenhuma mudança nos níveis de H3K27me3 foi observada nas regiões promotoras dos outros genes na vizinhança de *RASSF1A* e do promotor de *HOXA9* (Figura 13B). Este resultado é consistente com a ausência de efeito de *ANRASSF1* sobre a ocupação de SUZ12 nestes *loci* (Figura 13A).

Uma vez que a metilação do DNA do promotor de *RASSF1A* envolve o recrutamento de PRC2 e DNMT3B para este promotor (Palakurthy et al. 2009), testamos se o aumento do recrutamento de PRC2 causado pela super-expressão de *ANRASSF1* também afetaria a metilação do DNA neste *locus*, além da metilação na histona H3K27, já descrita acima. Por meio de ensaios utilizando endonuclease dependente de metilação, identificamos uma pequena mudança no padrão de metilação do DNA do promotor de *RASSF1A* em células super-expressando *ANRASSF1* quando comparados com células controle (Figura 104). No entanto, a heterogeneidade entre as triplicatas biológicas de transfecção impossibilitou a detecção de um aumento estatisticamente significativo no nível de metilação ($p = 0,10$). De fato, por meio deste ensaio, mesmo sem significância estatística, verificamos que o resultado é consistente com o descrito por (Palakurthy et al. 2009), onde um aumento da metilação do DNA no promotor de *RASSF1A* envolve a participação de PRC2. O pequeno aumento observado na metilação do DNA da região promotora em células HeLa super-expressando *ANRASSF1* pode refletir uma eventual disponibilidade limitada de outros fatores tais como

HOXB3 e DNMT3B que estão envolvidos na metilação do DNA do promotor de *RASSF1A* (Palakurthy et al. 2009). Em concordância com este resultado, já foi demonstrado que a região promotora de *RASSF1A* é hipometilada em células HeLa (Dammann et al. 2000).

Com isso, esses dados indicam que a marca de repressão da histona H3K27me3 é o alvo predominante de *ANRASSF1*, em contraposição à marca de metilação do DNA e que está regulação possa ser local e específica.

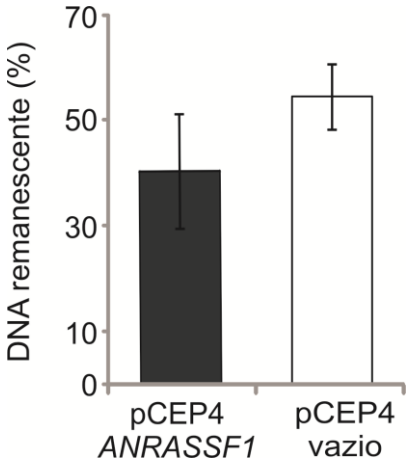

Figura 10 - A metilação do DNA da região promotora de *RASSF1A*. A metilação do DNA da região promotora de *RASSF1A* foi detectada mediante ensaio com uma endonuclease dependente de metilação, McrBC, em células super-expressando *ANRASSF1* ou em células controle, seguido de qPCR. A porcentagem de DNA remanescente foi calculada pela comparação da quantidade de DNA amplificada por qPCR em uma amostra submetida a tratamento com a endonuclease McrBC contra uma amostra não submetida ao tratamento. Os ensaios foram feitos em triplicata e é mostrada a média ± desvio-padrão.

4.5 – *ANRASSF1* se liga ao DNA e é necessário para a ocupação de SUZ12 no promotor de *RASSF1A*

Recentes estudos têm demonstrado que alguns lincRNAs como o $ncRNA_{CCND1}$ (Wang et al. 2008) e o Mistral (Bertani et al. 2011) se ligam ao DNA e recrutam proteínas que atuam nas regiões promotoras dos genes codificadores de proteínas da sua vizinhança para regular sua expressão. Como vimos que *ANRASSF1* está ligado a PRC2, testamos se *ANRASSF1*

poderia atuar de maneira semelhante a estes lincRNAs, por meio do recrutamento do complexo PRC2 para o promotor de *RASSF1A*, agindo em *cis*. Para esta finalidade, realizamos o ensaio denominado RNase ChIP em células HeLa controle (pCEP4_vazio), que inicialmente foram permeabilizadas e, em seguida, tratadas com RNase A ou H, para degradação de *ANRASSF1* endógeno livre ou eventualmente associado ao promotor de *RASSF1A*. Como controle do experimento, as células permeabilizadas foram tratadas com inibidor de RNase. Posteriormente, realizamos ChIP com o anticorpo anti-SUZ12, com anticorpo IgG de camundongo não imunizado, que foi usado como controle, para cada condição experimental, isto é, para células tratadas com RNase A, com RNase H ou com inibidor de RNase.

Verificamos por ChIP-qPCR que a máxima ocupação da região promotora de *RASSF1A* por SUZ12 foi detectada no tratamento com inibidor de RNase (Figura 115A, barra preta). Na condição de tratamento com RNase H, enzima que digere híbridos de RNA/DNA, SUZ12 foi quase totalmente liberada da região promotora (Figura 115A, barra de vermelho), visto que houve uma considerável diminuição de DNA da região promotora detectada pelo qPCR. Em paralelo, o RNA destas células foi extraído e realizada a medida dos níveis endógenos de *ANRASSF1*, por RT-qPCR, para testar se a diminuição de SUZ12 ligada ao promotor de *RASSF1A* observada era devida à diminuição de *ANRASSF1* por meio da digestão por RNase H.

Comparando-se os níveis de *ANRASSF1* nas células tratadas com inibidor de RNase e nas células tratadas com RNase H (Figura 115C, barra preta e vermelha, respectivamente), observamos que nesta última houve uma redução de 87% dos níveis endógenos de *ANRASSF1*. A digestão por RNase H indica que *ANRASSF1* faz parte de um híbrido RNA/DNA, e comprova que a ligação de SUZ12 no promotor de *RASSF1A* é dependente de *ANRASSF1*.

Na condição em que células HeLa foram permeabilizadas e tratadas com RNase A, enzima que digere apenas o RNA fita simples, foi observada uma menor redução tanto da ocupação de SUZ12 na região promotora (Figura 115A, barra azul) como dos níveis de *ANRASSF1* endógeno (Figura 11C, barra azul), quando comparado com o tratamento com RNase H. Isso nos sugere que o RNA fita simples encontra-se em rápido equilíbrio no nucleoplasma com o híbrido RNA/DNA, de modo que, por alguns instantes, este fica suscetível ao ataque da RNase A. Alternativamente, a interação RNA/proteínas PRC2 também pode estar em rápido equilíbrio com PRC2 livre, liberando assim uma porção da fita de RNA que fica livre para ser digerida. Este resultado reforça a conclusão anterior de que *ANRASSF1* faz parte de um híbrido RNA/DNA e que a ligação de SUZ12 no promotor de *RASSF1A* é dependente de *ANRASSF1*.

Como controle negativo realizamos ChIP-qPCR com anticorpo anti-RNA polimerase II, um fator cuja ocupação no promotor de *RASSF1A* não deve ser afetado pelo tratamento das RNase H ou RNase A. Verificamos que o tratamento com RNase H não alterou a ocupação da RNA Polimerase II (Figura 15B). Além de um controle adicional ao tratamento, medimos os níveis de expressão do gene *alfa-tubulina*, um mRNA que não forma híbridos RNA/DNA, já que este mRNA é exportado para o citoplasma para ser traduzido. Como esperado, observamos que o mRNA da *alfa-tubulina* não foi digerido pela RNase H (Figura 11D, barra vermelha) e só foi digerido pela RNase A (Figura 11D, barra azul).

Para confirmar este resultado, repetimos esses experimentos de RNase ChIP em células super-expressando *ANRASSF1*. Igualmente, observamos uma redução na ocupação de SUZ12 na região promotora de *RASSF1A* quando as células super-expressando *ANRASSF1* foram tratadas com RNase H e A (Figura 126A), um padrão semelhante ao encontrado na condição endógena (Figura 11A). Além disso, encontramos um padrão muito semelhante na porcentagem de *ANRASSF1* remanescente após o tratamento com as enzimas (Figura 126B),

ao comparar-se com a condição endógena (Figura 11C). É intressante notar que restou uma maior ocupação SUZ12 na região promotora de *RASSF1A* (Figura 126A), comparado com a ocupação nas células não super-expressando ANRASSF1 (Figura 115A). Isto se deve provavelmente ao fato de que há um excesso de *ANRASSF1* nas células devido à super-expressão, e apesar de uma grande fração de *ANRASSF1* ser degradada pela RNase H, resta ainda uma quantidade suficiente de *ANRASSF1* para se ligar ao DNA e ainda recrutar uma certa quantidade de PRC2 ao promotor de *RASSF1A*.

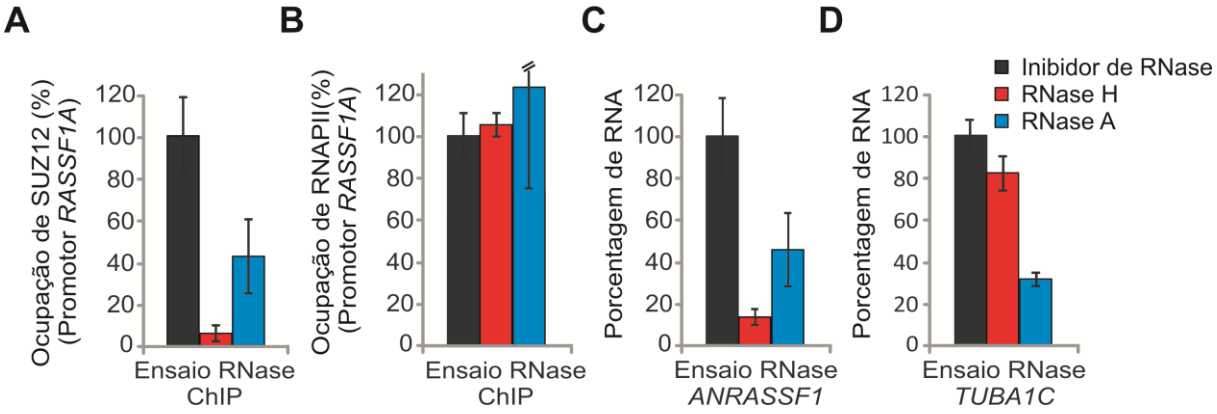

Figura 11 - *ANRASSF1* endógeno promove o recrutamento de SUZ12 para o promotor de *RASSF1A*. (A) Ensaio de RNase-ChIP, onde o DNA da região promotora de *RASSF1A* foi medido por qPCR na fração imunoprecipitada com anticorpo anti-SUZ12, proveniente de células HeLa controle permeabilizadas e tratadas com inibidor de RNase (barra preta), com RNase H (barra vermelha) ou com RNase A (barra azul). Medidas de qPCR obtidas utilizando anticorpo anti-SUZ12 foram comparadas com aquelas obtidas em paralelo com amostras tratadas com IgG não específico; a normalização foi realizada pela medida do *input*. A porcentagem de ocupação de SUZ12 na região promotora de *RASSF1A* mostrada para as duas RNAses foi calculada como relativa aos valores correspondentes para o inibidor de RNase, a partir de duas réplicas experimentais medidas em triplicata; barras de erro referem-se ao desvio-padrão.(**B**) Controle negativo do ensaio RNase ChIP para a região promotora de *RASSF1A*, utilizando anticorpo anti-RNA Pol II, nas mesmas condições descritas em (A). **(C)** Ensaio de RNase para a detecção do RNA de *ANRASSF1* por RT-qPCR em células HeLa permeabilizadas e tratadas com o inibidor de RNase (barra de erro), com RNase H (barra vermelha) ou com RNase A (barra azul). Porcentagem de RNA para cada um dos dois tratamentos com RNase foi expressa relativa aos valores correspondentes ao inibidor de RNase. Três ensaios foram feitos e é mostrada a média ± desvio-padrão. **(D)** Como controle, mRNA de alfa-tubulina foi medido em paralelo por RT-qPCR, nas mesmas condições descritas em (C).

No controle do tratamento, com o gene *alfa-tubulina*, não encontramos diferença nos resultados entre células controle (Figura 115C) e células super-expressando *ANRASSF1* (Figura 12C).

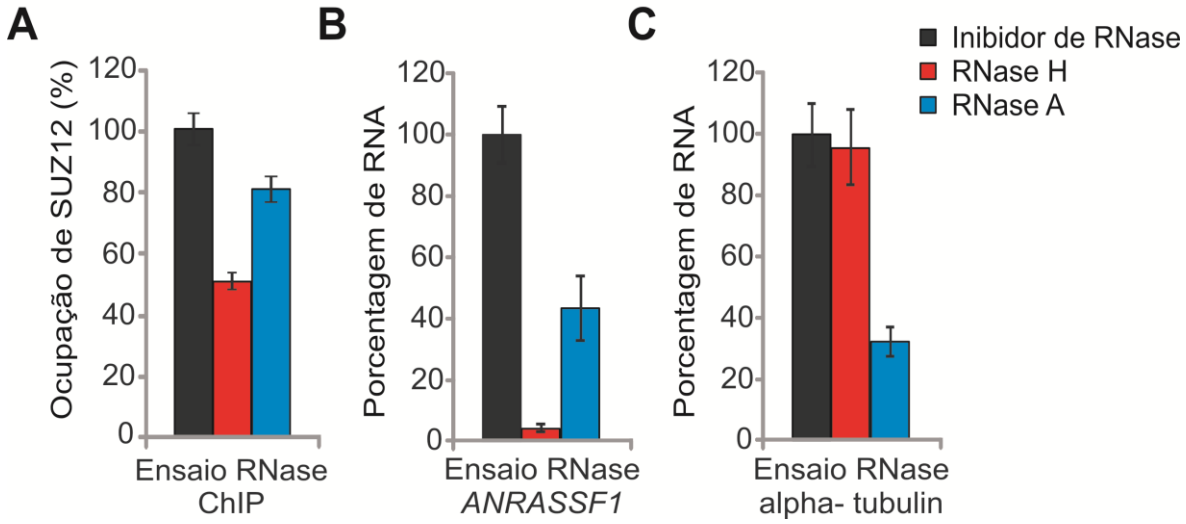

Figura 12 Super-expressão de *ANRASSF1* promove o recrutamento de SUZ12 para o promotor de *RASSF1A*. (A) Ensaio de RNase-ChIP, onde o DNA da região promotora de *RASSF1A* foi medido por qPCR na fração imunoprecipitada com anticorpo anti-SUZ12, proveniente de células HeLa super-expressando *ANRASSF1* permeabilizadas e tratadas com inibidor de RNase (barra preta), com RNase H (barra vermelha) ou com RNase A (barra azul). Medidas de qPCR obtidas utilizando anticorpo anti-SUZ12 foram comparadas com aquelas obtidas em paralelo com amostras tratadas com IgG não específico; a normalização foi realizada pela medida do *input*. A porcentagem de ocupação de SUZ12 na região promotora de *RASSF1A* mostrada para as duas RNAses foi calculada como relativa aos valores correspondentes para o inibidor de RNase, a partir de duas réplicas experimentais medidas em triplicata; barras de erro referem-se ao desvio-padrão. **(B)** Ensaio de RNase para a detecção do RNA de *ANRASSF1* por RT-qPCR em células HeLa super-expressando *ANRASSF1* permeabilizadas e tratadas com o inibidor de RNase (barra de erro), com RNase H (barra vermelha) ou com RNase A (barra azul). Porcentagem de RNA para cada um dos dois tratamentos com RNase foi expressa relativa aos valores correspondentes ao inibidor de RNase. Três ensaios foram feitos e é mostrada a média ± desvio-padrão. **(C)** Como controle, mRNA de alfa-tubulina foi medido em paralelo por RT-qPCR, nas mesmas condições descritas em (B).

Com esses resultados, demonstramos que o lncRNA antissenso *unspliced* intrônico, *ANRASSF1*, se liga ao PRC2 e é necessário para o recrutamento e ocupação deste complexo no promotor de *RASSF1A*. Neste sentido, postulamos um modelo mecanístico (Figura 137),

em que a interação de *ANRASSF1*, tanto com o DNA da região em que ocorre sua transcrição, bem como com o PRC2, pode induzir uma mudança de conformação criando uma volta no DNA, fazendo o recrutamento de PRC2 para o promotor de *RASSF1A*. Este recrutamento de PRC2 induz o acúmulo da marca de repressão H3K27me3, que culmina com a diminuição da expressão da isoforma *RASSF1A* (Figura 137).

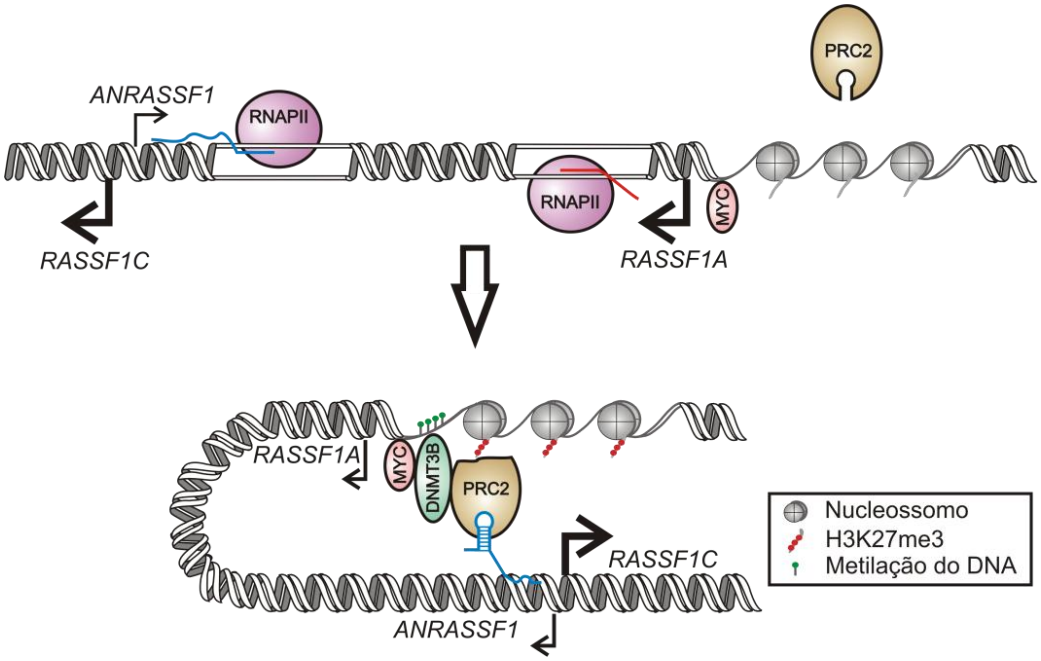

Figura 137 - Modelo funcional proposto para a ação de *ANRASSF1* no *locus* de *RASSF1*. Postulamos que a interação do lncRNA *ANRASSF1* (linha azul) com o DNA genômico, no seu sítio de transcrição, forma um híbrido RNA/DNA que recruta o complexo modificador de cromatina, PRC2, para a região promotora do gene *RASSF1A*. A ligação de *ANRASSF1* ao PRC2 causa uma modificação conformacional que ativa este complexo, alterando o padrão de metilação da histona H3K27 (círculos vermelhos). Isto causa uma diminuição específica da atividade transcricional de *RASSF1A*, com nenhum efeito sobre a transcrição de *RASSF1C*. O fator de transcrição MYC se liga à região promotora de *RASSF1A* e regula a atividade de liberação da pausa da RNA Pol II (Rahl et al. 2010). A DNMT3B é recrutada para o promotor de *RASSF1A*, através de interações com PRC2 e MYC (Palakurthy et al. 2009), resultando na hipermetilação do DNA (círculos verdes).

5. DISCUSSÃO

Após o sequenciamento do genoma humano, no início da década passada, e com os recentes avanços nas técnicas de larga escala em genômica, tem sido descrito que ocorre a transcrição de milhares de lncRNAs ao longo de mais de 90% de todo o genoma humano (Birney et al. 2007; Kapranov et al. 2007), e que estes lncRNAs podem agir por diversos mecanismos. No entanto, apenas alguns deles foram caracterizados funcionalmente em detalhe (Wilusz et al. 2009; Wang et al. 2011; Rinn and Chang 2012). A definição e a nomeação dos lncRNAs ainda encontra-se em evolução na literatura, e diferentes classes ou categorias de lncRNAs têm sido descritas (Wright and Bruford 2011). Porém, atualmente, a maioria dos grupos tem focado seus esforços no estudo de uma classe de lncRNA *spliced*, que tem sido chamada de RNA não codificador longo intergênico (lincRNA). Um dos motivos para que muitos grupos venham focando seus estudos em lincRNAs é que este sofrem *splicing* e estão fora do locus de um gene codificador de proteína, o que possibilita maior segurança quanto a existência de um transcrito independente. Já os lncRNAs intrônicos *unspliced* têm sido pouco caracterizados e até muitas vezes descartados, pois frequentemente suspeita-se de que eles sejam um possível artefato técnico ou um ruído transcricional, ou ainda, fragmentos de DNA genômico. Na realidade, na revisão recente de (Rinn and Chang 2012), o "mapa anatômico de lncRNAs" mostra somente lncRNAs *spliced* intrônicos e intergênicos, e exclui todos os lncRNAs *unspliced*.

Em contraste com as informações predominantes na literatura, identificamos *ANRASSF1*, um lncRNA intrônico *unspliced*, que é expresso com orientação antissenso em relação ao gene codificador de proteína e possui uma função regulatória *in cis* no promotor de *RASSF1A*. Ao contrário disso, têm sido demonstrado que a maioria dos lincRNAs atua principalmente *in trans* (Guttman et al. 2011). Os poucos lncRNAs intrônicos *unspliced* descritos até o momento que atuam *in cis*, participam principalmente de eventos epigenéticos

de genes *imprinted*, como discutido mais abaixo. Com isso, o mecanismo descrito para *ANRASSF1* acrescenta uma novidade para a classe, visto que envolve uma regulação epigenética *in cis* em um gene não-*imprinted*.

Por outro lado, *ANRASSF1* também apresenta características comuns a lincRNAs bem caracterizados, tais como a presença de modificação cap na ponta 5´do RNA, transcrição pela RNAPII, enriquecimento nuclear e ligação com proteínas do complexo PRC2 (Guttman et al. 2009; Khalil et al. 2009). Adicionalmente, verificamos que *ANRASSF1* possui uma meia-vida curta, e que sua estabilidade é próxima da estabilidade de genes regulatórios tais como c-*myc*. Um trabalho recente identificou que lncRNAs localizados no núcleo possuem baixa estabilidade (Clark et al. 2012). Este resultado é consistente com a idéia de que estes lncRNAs estejam agindo como moléculas regulatórias, da mesma forma que mRNAs codificadores de proteínas pouco estáveis, possivelmente para proporcionar um elevado grau de dinamismo aos processos celulares que regulam (Clark et al. 2012). Como vimos que *ANRASSF1* participa da maquinaria de regulação de *RASSF1A*, recrutando PRC2 para este *locus*, foi interessante observar uma meia-vida curta para este lncRNA, o que pode proporcionar um dinamismo no controle da expressão deste gene envolvido na regulação do ciclo celular.

Para a maioria dos lncRNAs estudados até o momento foi demonstrada evidência de uma região promotora, apenas por meio da identificação de marcas de histona H3K4me3, que estão relacionadas com promotores ativos, próximas ao início de transcrição (Guttman et al. 2009; Conley and King Jordan 2012). Entretanto, como a maioria desses estudos ocorreu em regiões intergênicas ou intrônicas antissenso, sem evidência de início de transcrição próximo nas duas fitas ou de promotores bidirecionais, as marcas de histonas foram suficientes para considerar esta região como um promotor.

Como os dados de predição *in silico* do trabalho de Davuluri e colaboradores (Davuluri et al. 2001) apontavam para a existência de uma região promotora bidirecional

ANRASSF1/*RASSF1C*, não foi possível utilizar as informações de histona H3K4me3 e foi necessária uma validação *in vitro* dos promotores para confirmar a região promotora bidirecional predita. A partir do estudo *in vitro*, demostramos que *ANRASSF1* é uma unidade transcricional independente.

Vale ressaltar que, em células HeLa, *ANRASSF1* parece ter um nível de expressão consideravelmente inferior ao de lincRNAs bem estudados, e mesmo assim, alterações na sua abundância, por super-expressão ectópica ou por silenciamento mediado por siRNA, afetaram especificamente a expressão de *RASSF1A* por meio do recrutamento de PRC2. É importante notar que a alteração da abundância de *ANRASSF1* não afetou a expressão de *RASSF1C*, outra isoforma expressa no *locus RASSF1* sob um promotor diferente, nem afetou o recrutamento de PRC2 para o promotor de *RASSF1C* e para o promotor de genes vizinhos ao *locus* de *RASSF1*, incluindo outros dois genes supressores tumorais bem caracterizados, *TUSC2* (Uno et al. 2004) e *NPRL2* (Kurata et al. 2008). Portanto, os dados apontam para uma regulação epigenética altamente local-específica, direcionada pelo lncRNA.

Já tem sido identificado o envolvimento de lncRNAs no silenciamento gênico por meio de enzimas envolvidas no remodelamento da cromatina, especialmente com o complexo PRC2, causando modificação pós-traducional de histonas em genes alvo (Rinn et al. 2007; Khalil et al. 2009; Guttman et al. 2011). Recentemente, foi evidenciado, por meio de *crosslinked* RIP-seq em linhagem humana tumoral, um grande número de sequências intrônicas *unspliced* capazes de se ligarem a EZH2, um dos principais componentes do PRC2 (Guil et al. 2012). Neste trabalho, os autores caracterizaram em maior detalhe um RNA intrônico *unspliced*, com orientação senso em relação ao gene *SMYD3* que se liga a EZH2. Foi visto que sua super-expressão do lncRNA senso reduz os níveis de *SMYD3*, por meio do recrutamento de EZH2 para o *locus* de *SMYD3* e consequente aumento da marca repressiva H3K27me3 neste local. Portanto, esse lncRNA intrônico *unspliced* senso regula *in cis* o gene

SMYD3 por meio de eventos epigenéticos e causa uma diminuição da proliferação celular (Guil et al. 2012). Guil e colaboradores não conseguiram determinar a origem desse lncRNA intrônico *unspliced* senso, e sugerem que EZH2 se ligue no pré-mRNA ou após o processamento do intron, indicando ser mais provável que ocorra após o *splicing* (Guil et al. 2012).

Outro exemplo de lncRNA intrônico *unspliced* que recruta EZH2 é o *Kcnq1ot1*, que é transcrito no *locus* de *KCNQ1* na fita oposta do gene codificador de proteína, o que o caracteriza como um antissenso (Pandey et al. 2008). O lncRNA *Kcnq1ot1* regula a expressão de onze genes *imprinted* no *cluster* do gene *KCNQ1* (Pandey et al. 2008), um mecanismo nitidamente diferente do encontrado para *ANRASSF1*, que consiste em uma regulação local e específica *in cis*.

O mecanismo de ação proposto nesse trabalho para *ANRASSF1*, que envolve um híbrido RNA/DNA que recruta o complexo PRC2 e consequentemente dispara uma alteração conformacional da cromatina, foi baseado principalmente nos mecanismos descritos para os $ncRNAs_{CCND1}$ (Wang et al. 2008) e para o lincRNA *Mistral* (Bertani et al. 2011). Os $ncRNAs_{CCND1}$ consistem de uma série de ncRNAs, que são transcritos na fita senso e antissenso a partir da região promotora do gene *CCND1* e se ligam nesta região, formando híbridos de RNA/DNA. Estes híbridos recrutam a proteína TLS para esta região promotora, causando uma repressão gênica (Wang et al. 2008) por meio de uma regulação local e específica *in cis*. Os $ncRNAs_{CCND1}$ são expressos em baixos níveis, apresentam tamanhos variados, são poliadenilados e não possuem cap 5´ (Wang et al. 2008). Sendo assim, estes transcritos representam um exemplo de outra classe de ncRNAs que regulam genes específicos e que exercem papéis funcionais na transcrição gênica pelo recrutamento de co-reguladores via ligação do RNA para as unidades transcricionais específicas.

O lincRNA *Mistral* também envolve a formação de um híbrido lncRNA/DNA e o recrutamento de um complexo proteico, e adicionalmente seu mecanismo envolve uma modificação da conformação da cromatina. Este lincRNA age de maneira análoga ao mecanismo aqui proposto. O mecanismo de regulação de Mistral envolve a enzima remodeladora da cromatina MLL1, que reconhece e liga o híbrido Mistral/DNA (Bertani et al. 2011) e consequentemente desencadeia alterações dinâmicas na conformação cromossômica. Porém, diferente do modelo que propomos para *ANRASSF1*, que reprime a expressão de *RASSF1A*, o lincRNA Mistral promove a ativação da transcrição dos genes Hoxa6 e Hoxa7 (Bertani et al. 2011).

Há boas evidências de que lncRNAs estejam envolvidos em uma série de doenças humanas (Taft et al. 2010; Esteller 2011; Wapinski and Chang 2011), especialmente em câncer (Esteller 2011; Prensner and Chinnaiyan 2011; Reis and Verjovski-Almeida 2012; Spizzo et al. 2012). Neste trabalho, comprovamos que *ANRASSF1* regula negativamente a expressão do gene supressor tumoral *RASSF1A*, por meio do enriquecimento da marca de histona associada a repressão, H3K27me3, no promotor de *RASSF1A*. Têm sido demonstrado que a hipermetilação da ilha CpG da região promotora de *RASSF1A* e sua resultante inativação são eventos epigenéticos frequentes em diversos tipos de câncer humano (Dammann et al. 2000). Além disso, já foi visto que essa metilação do DNA do promotor de *RASSF1A* é desencadeada pela trimetilação dos resíduos de lisina nove e vinte e sete da histona 3 (H3K9 e H3K27) (Strunnikova et al. 2005; Kawamoto et al. 2007; Starlard-Davenport et al. 2010). Encontramos um discreto aumento na metilação do DNA da ilha CpG do promotor de *RASSF1A* e um aumento na ocupação de PRC2 e do níveis de H3K27me3 neste promotor após a super-expressão de *ANRASSF1*. Esses resultados são consistentes com o mecanismo de metilação do DNA desta ilha CpG, já descrito, o qual envolve o recrutamento de PRC2 (Palakurthy et al. 2009). Nossos dados acrescentam uma nova camada no

mecanismo de regulação do supressor tumoral *RASSF1A*, por meio de *ANRASSF1*, que afeta predominantemente o passo inicial da trimetilação da histona; se sabe que é a trimetilação da histona quem dispara a metilação do DNA. Apesar de termos identificado uma maior expressão de *ANRASSF1* em linhagens celulares tumorais, ainda são necessários estudos adicionais sobre o possível envolvimento de *ANRASSF1* no processo de tumorigênese.

ANRASSF1 é um dos milhares de lncRNAs *unspliced*, que são transcritos no genoma humano a partir de regiões intrônicas de 74% dos genes codificadores de proteínas (Nakaya et al. 2007; Engelhardt and Stadler 2012). A formação de um híbrido RNA/DNA no *locus* de transcrição de cada um dos lncRNAs intrônicos *unspliced* poderia resultar em um efeito altamente localizado e específico de regulação, por meio do recrutamento de complexos remodeladores da cromatina. Estamos propondo que uma fração destes milhares de lncRNAs intrônicos *unspliced* não caracterizados no genoma humano possa contribuir como integrantes de uma ampla rede de modulação epigenética, em que cada um atue especificamente no *locus* em que este lncRNA seja transcrito.

6. CONCLUSÃO

Este trabalho identificou um lncRNA intrônico *unspliced* antissenso ao gene *RASSF1*, denominado *ANRASSF1* e caracterizou seu papel funcional na regulação transcricional do gene supressor tumoral *RASSF1A*. Essa regulação ocorre por meio de mudanças na conformação da cromatina, disparadas pela formação de um híbrido RNA/DNA em combinação com o recrutamento do complexo repressivo PRC2 para a região promotora de *RASSF1A*. Com isso, o complexo PRC2 promove a trimetilação da histona H3K27, o que causa uma redução da transcrição de *RASSF1A*. Essa alteração da histona possibilita o desencadeamento do processo de metilação do DNA do promotor.

Além disso, os dados obtidos nesse trabalho trazem à tona a importância funcional dos lncRNAs intrônicos *unspliced*. Embora existam evidências de que esses lncRNAs sejam expressos no mesmo *locus* em 74% dos genes codificadores de proteína, a maior parte dos pesquisadores que trabalham com ncRNAs tem, de certa forma, negligenciado o seu estudo. É possível que, além do *ANRASSF1*, existam muitos outros lncRNAs intrônicos *unspliced* que atuem como moduladores específicos de eventos epigenéticos, no *locus* em que são transcritos.

7. REFERÊNCIAS

(ENCODE, 2011). "A user's guide to the encyclopedia of DNA elements (ENCODE)." PLoS Biol **9**(4): e1001046.

Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G. Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. M. Cherry, S. Cawley, C. Dahlke, L. B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. D. Mays, I. Dew, S. M. Dietz, K. Dodson, L. E. Doup, M. Downes, S. Dugan-Rocha, B. C. Dunkov, P. Dunn, K. J. Durbin, C. C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A. E. Gabrielian, N. S. Garg, W. M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. H. Gorrell, Z. Gu, P. Guan, M. Harris, N. L. Harris, D. Harvey, T. J. Heiman, J. R. Hernandez, J. Houck, D. Hostin, K. A. Houston, T. J. Howland, M. H. Wei, C. Ibegwam, M. Jalali, F. Kalush, G. H. Karpen, Z. Ke, J. A. Kennison, K. A. Ketchum, B. E. Kimmel, C. D. Kodira, C. Kraft, S. Kravitz, D. Kulp, Z. Lai, P. Lasko, Y. Lei, A. A. Levitsky, J. Li, Z. Li, Y. Liang, X. Lin, X. Liu, B. Mattei, T. C. McIntosh, M. P. McLeod, D. McPherson, G. Merkulov, N. V. Milshina, C. Mobarry, J. Morris, A. Moshrefi, S. M. Mount, M. Moy, B. Murphy, L. Murphy, D. M. Muzny, D. L. Nelson, D. R. Nelson, K. A. Nelson, K. Nixon, D. R. Nusskern, J. M. Pacleb, M. Palazzolo, G. S. Pittman, S. Pan, J. Pollard, V. Puri, M. G. Reese, K. Reinert, K. Remington, R. D. Saunders, F. Scheeler, H. Shen, B. C. Shue, I. Siden-Kiamos, M. Simpson, M. P. Skupski, T. Smith, E. Spier, A. C. Spradling, M. Stapleton, R. Strong, E. Sun, R. Svirskas, C. Tector, R. Turner, E. Venter, A. H. Wang, X. Wang, Z. Y. Wang, D. A. Wassarman, G. M. Weinstock, J. Weissenbach, S. M. Williams, WoodageT, K. C. Worley, D. Wu, S. Yang, Q. A. Yao, J. Ye, R. F. Yeh, J. S. Zaveri, M. Zhan, G. Zhang, Q. Zhao, L. Zheng, X. H. Zheng, F. N. Zhong, W. Zhong, X. Zhou, S. Zhu, X. Zhu, H. O. Smith, R. A. Gibbs, E. W. Myers, G. M. Rubin and J. C. Venter (2000). "The genome sequence of Drosophila melanogaster." Science **287**(5461): 2185-2195.

Agathanggelou, A., W. N. Cooper and F. Latif (2005). "Role of the Ras-association domain family 1 tumor suppressor gene in human cancers." Cancer Res **65**(9): 3497-3508.

Almeida, G. T., M. S. Amaral, F. C. Beckedorff, J. P. Kitajima, R. Demarco and S. Verjovski-Almeida (2011). "Exploring the Schistosoma mansoni adult male transcriptome using RNA-seq." Exp Parasitol.

Amaral, P. P. and J. S. Mattick (2008). "Noncoding RNA in development." Mamm Genome **19**(7-8): 454-492.

Avruch, J., R. Xavier, N. Bardeesy, X. F. Zhang, M. Praskova, D. Zhou and F. Xia (2009). "Rassf family of tumor suppressor polypeptides." J Biol Chem **284**(17): 11001-11005.

Baksh, S., S. Tommasi, S. Fenton, V. C. Yu, L. M. Martins, G. P. Pfeifer, F. Latif, J. Downward and B. G. Neel (2005). "The tumor suppressor RASSF1A and MAP-1 link death receptor signaling to Bax conformational change and cell death." Mol Cell **18**(6): 637-650.

Bertani, S., S. Sauer, E. Bolotin and F. Sauer (2011). "The noncoding RNA Mistral activates Hoxa6 and Hoxa7 expression and stem cell differentiation by recruiting MLL1 to chromatin." Mol Cell **43**(6): 1040-1046.

Bertone, P., V. Stolc, T. E. Royce, J. S. Rozowsky, A. E. Urban, X. Zhu, J. L. Rinn, W. Tongprasit, M. Samanta, S. Weissman, M. Gerstein and M. Snyder (2004). "Global identification of human transcribed sequences with genome tiling arrays." Science **306**(5705): 2242-2246.

Birney, E., J. A. Stamatoyannopoulos, A. Dutta, R. Guigo, T. R. Gingeras, E. H. Margulies, Z. Weng, M. Snyder, E. T. Dermitzakis, R. E. Thurman, M. S. Kuehn, C. M. Taylor, S. Neph, C. M. Koch, S. Asthana, A. Malhotra, I. Adzhubei, J. A. Greenbaum, R. M. Andrews, P. Flicek, P. J. Boyle, H. Cao, N. P. Carter, G. K. Clelland, S. Davis, N. Day, P. Dhami, S. C. Dillon, M. O. Dorschner, H. Fiegler, P. G. Giresi, J. Goldy, M. Hawrylycz, A. Haydock, R. Humbert, K. D. James, B. E. Johnson, E. M. Johnson, T. T. Frum, E. R. Rosenzweig, N. Karnani, K. Lee, G. C. Lefebvre, P. A. Navas, F. Neri, S. C. Parker, P. J. Sabo, R. Sandstrom, A. Shafer, D. Vetrie, M. Weaver, S. Wilcox, M. Yu, F. S. Collins, J. Dekker, J. D. Lieb, T. D. Tullius, G. E. Crawford, S. Sunyaev, W. S. Noble, I. Dunham, F. Denoeud, A. Reymond, P. Kapranov, J. Rozowsky, D. Zheng, R. Castelo, A. Frankish, J. Harrow, S. Ghosh, A. Sandelin, I. L. Hofacker, R. Baertsch, D. Keefe, S. Dike, J. Cheng, H. A. Hirsch, E. A. Sekinger, J. Lagarde, J. F. Abril, A. Shahab, C. Flamm, C. Fried, J. Hackermuller, J. Hertel, M. Lindemeyer, K. Missal, A. Tanzer, S. Washietl, J. Korbel, O. Emanuelsson, J. S. Pedersen, N. Holroyd, R. Taylor, D. Swarbreck, N. Matthews, M. C. Dickson, D. J. Thomas, M. T. Weirauch, J. Gilbert, J. Drenkow, I. Bell, X. Zhao, K. G. Srinivasan, W. K. Sung, H. S. Ooi, K. P. Chiu, S. Foissac, T. Alioto, M. Brent, L. Pachter, M. L. Tress, A. Valencia, S. W. Choo, C. Y. Choo, C. Ucla, C. Manzano, C. Wyss, E. Cheung, T. G. Clark, J. B. Brown, M. Ganesh, S. Patel, H. Tammana, J. Chrast, C. N. Henrichsen, C. Kai, J. Kawai, U. Nagalakshmi, J. Wu, Z. Lian, J. Lian, P. Newburger, X. Zhang, P. Bickel, J. S. Mattick, P. Carninci, Y. Hayashizaki, S. Weissman, T. Hubbard, R. M. Myers, J. Rogers, P. F. Stadler, T. M. Lowe, C. L. Wei, Y. Ruan, K. Struhl, M. Gerstein, S. E. Antonarakis, Y. Fu, E. D. Green, U. Karaoz, A. Siepel, J. Taylor, L. A. Liefer, K. A. Wetterstrand, P. J. Good, E. A. Feingold, M. S. Guyer, G. M. Cooper, G. Asimenos, C. N. Dewey, M. Hou, S. Nikolaev, J. I. Montoya-Burgos, A. Loytynoja, S. Whelan, F. Pardi, T. Massingham, H. Huang, N. R. Zhang, I. Holmes, J. C. Mullikin, A. Ureta-Vidal, B. Paten, M. Seringhaus, D. Church, K. Rosenbloom, W. J. Kent, E. A. Stone, S. Batzoglou, N. Goldman, R. C. Hardison, D. Haussler, W. Miller, A. Sidow, N. D. Trinklein, Z. D. Zhang, L. Barrera, R. Stuart, D. C. King, A. Ameur, S. Enroth, M. C. Bieda, J. Kim, A. A. Bhinge, N. Jiang, J. Liu, F. Yao, V. B. Vega, C. W. Lee, P. Ng, A. Yang, Z. Moqtaderi, Z. Zhu, X. Xu, S. Squazzo, M. J. Oberley, D. Inman, M. A. Singer, T. A. Richmond, K. J. Munn, A. Rada-Iglesias, O. Wallerman, J. Komorowski, J. C. Fowler, P. Couttet, A. W. Bruce, O. M. Dovey, P. D. Ellis, C. F. Langford, D. A. Nix, G. Euskirchen, S. Hartman, A. E. Urban, P. Kraus, S. Van Calcar, N. Heintzman, T. H. Kim, K. Wang, C. Qu, G. Hon, R. Luna, C. K. Glass, M. G. Rosenfeld, S. F. Aldred, S. J. Cooper, A. Halees, J. M. Lin, H. P. Shulha, M. Xu, J. N. Haidar, Y. Yu, V. R. Iyer, R. D. Green, C. Wadelius, P. J. Farnham, B. Ren, R. A. Harte, A. S. Hinrichs, H. Trumbower, H. Clawson, J. Hillman-Jackson, A. S. Zweig, K. Smith, A. Thakkapallayil, G. Barber, R. M. Kuhn, D. Karolchik, L. Armengol, C. P. Bird, P. I. de Bakker, A. D. Kern, N. Lopez-Bigas, J. D. Martin, B. E. Stranger, A. Woodroffe, E. Davydov, A. Dimas, E. Eyras, I. B. Hallgrimsdottir, J. Huppert, M. C. Zody, G. R. Abecasis, X. Estivill, G. G. Bouffard, X. Guan, N. F. Hansen, J. R. Idol, V. V. Maduro, B. Maskeri, J. C. McDowell, M. Park, P. J. Thomas, A. C. Young, R.

W. Blakesley, D. M. Muzny, E. Sodergren, D. A. Wheeler, K. C. Worley, H. Jiang, G. M. Weinstock, R. A. Gibbs, T. Graves, R. Fulton, E. R. Mardis, R. K. Wilson, M. Clamp, J. Cuff, S. Gnerre, D. B. Jaffe, J. L. Chang, K. Lindblad-Toh, E. S. Lander, M. Koriabine, M. Nefedov, K. Osoegawa, Y. Yoshinaga, B. Zhu and P. J. de Jong (2007). "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project." Nature **447**(7146): 799-816.

Burbee, D. G., E. Forgacs, S. Zochbauer-Muller, L. Shivakumar, K. Fong, B. Gao, D. Randle, M. Kondo, A. Virmani, S. Bader, Y. Sekido, F. Latif, S. Milchgrub, S. Toyooka, A. F. Gazdar, M. I. Lerman, E. Zabarovsky, M. White and J. D. Minna (2001). "Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression." J Natl Cancer Inst **93**(9): 691-699.

Bushnell, D. A., P. Cramer and R. D. Kornberg (2002). "Structural basis of transcription: alpha-amanitin-RNA polymerase II cocrystal at 2.8 A resolution." Proc Natl Acad Sci U S A **99**(3): 1218-1222.

Byun, J. S., M. M. Wong, W. Cui, G. Idelman, Q. Li, A. De Siervi, S. Bilke, C. M. Haggerty, A. Player, Y. H. Wang, M. J. Thirman, J. J. Kaberlein, C. Petrovas, R. A. Koup, D. Longo, K. Ozato and K. Gardner (2009). "Dynamic bookmarking of primary response genes by p300 and RNA polymerase II complexes." Proc Natl Acad Sci U S A **106**(46): 19286-19291.

Cabili, M. N., C. Trapnell, L. Goff, M. Koziol, B. Tazon-Vega, A. Regev and J. L. Rinn (2011). "Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses." Genes Dev **25**(18): 1915-1927.

Cao, R. and Y. Zhang (2004). "SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex." Mol Cell **15**(1): 57-67.

Cappellen, D., T. Schlange, M. Bauer, F. Maurer and N. E. Hynes (2007). "Novel c-MYC target genes mediate differential effects on cell proliferation and migration." EMBO Rep **8**(1): 70-76.

Carninci, P., T. Kasukawa, S. Katayama, J. Gough, M. C. Frith, N. Maeda, R. Oyama, T. Ravasi, B. Lenhard, C. Wells, R. Kodzius, K. Shimokawa, V. B. Bajic, S. E. Brenner, S. Batalov, A. R. Forrest, M. Zavolan, M. J. Davis, L. G. Wilming, V. Aidinis, J. E. Allen, A. Ambesi-Impiombato, R. Apweiler, R. N. Aturaliya, T. L. Bailey, M. Bansal, L. Baxter, K. W. Beisel, T. Bersano, H. Bono, A. M. Chalk, K. P. Chiu, V. Choudhary, A. Christoffels, D. R. Clutterbuck, M. L. Crowe, E. Dalla, B. P. Dalrymple, B. de Bono, G. Della Gatta, D. di Bernardo, T. Down, P. Engstrom, M. Fagiolini, G. Faulkner, C. F. Fletcher, T. Fukushima, M. Furuno, S. Futaki, M. Gariboldi, P. Georgii-Hemming, T. R. Gingeras, T. Gojobori, R. E. Green, S. Gustincich, M. Harbers, Y. Hayashi, T. K. Hensch, N. Hirokawa, D. Hill, L. Huminiecki, M. Iacono, K. Ikeo, A. Iwama, T. Ishikawa, M. Jakt, A. Kanapin, M. Katoh, Y. Kawasawa, J. Kelso, H. Kitamura, H. Kitano, G. Kollias, S. P. Krishnan, A. Kruger, S. K. Kummerfeld, I. V. Kurochkin, L. F. Lareau, D. Lazarevic, L. Lipovich, J. Liu, S. Liuni, S. McWilliam, M. Madan Babu, M. Madera, L. Marchionni, H. Matsuda, S. Matsuzawa, H. Miki, F. Mignone, S. Miyake, K. Morris, S. Mottagui-Tabar, N. Mulder, N. Nakano, H. Nakauchi, P. Ng, R. Nilsson, S. Nishiguchi, S. Nishikawa, F. Nori, O. Ohara, Y. Okazaki, V. Orlando, K. C. Pang, W. J. Pavan, G. Pavesi, G. Pesole, N. Petrovsky, S. Piazza, J. Reed, J. F. Reid, B. Z. Ring, M. Ringwald, B. Rost, Y. Ruan, S. L. Salzberg, A. Sandelin, C. Schneider, C. Schonbach, K. Sekiguchi, C. A. Semple, S. Seno, L. Sessa, Y. Sheng, Y. Shibata, H. Shimada, K. Shimada, D. Silva, B. Sinclair, S. Sperling, E. Stupka, K. Sugiura, R. Sultana, Y. Takenaka, K. Taki, K. Tammoja, S. L. Tan, S. Tang, M. S. Taylor, J. Tegner, S. A.

Teichmann, H. R. Ueda, E. van Nimwegen, R. Verardo, C. L. Wei, K. Yagi, H. Yamanishi, E. Zabarovsky, S. Zhu, A. Zimmer, W. Hide, C. Bult, S. M. Grimmond, R. D. Teasdale, E. T. Liu, V. Brusic, J. Quackenbush, C. Wahlestedt, J. S. Mattick, D. A. Hume, C. Kai, D. Sasaki, Y. Tomaru, S. Fukuda, M. Kanamori-Katayama, M. Suzuki, J. Aoki, T. Arakawa, J. Iida, K. Imamura, M. Itoh, T. Kato, H. Kawaji, N. Kawagashira, T. Kawashima, M. Kojima, S. Kondo, H. Konno, K. Nakano, N. Ninomiya, T. Nishio, M. Okada, C. Plessy, K. Shibata, T. Shiraki, S. Suzuki, M. Tagami, K. Waki, A. Watahiki, Y. Okamura-Oho, H. Suzuki, J. Kawai and Y. Hayashizaki (2005). "The transcriptional landscape of the mammalian genome." Science **309**(5740): 1559-1563.

Carthew, R. W. and E. J. Sontheimer (2009). "Origins and Mechanisms of miRNAs and siRNAs." Cell **136**(4): 642-655.

Chow, L. S., K. W. Lo, J. Kwong, K. F. To, K. S. Tsang, C. W. Lam, R. Dammann and D. P. Huang (2004). "RASSF1A is a target tumor suppressor from 3p21.3 in nasopharyngeal carcinoma." Int J Cancer **109**(6): 839-847.

Chu, C., K. Qu, F. L. Zhong, S. E. Artandi and H. Y. Chang (2011). "Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions." Mol Cell **44**(4): 667-678.

Clark, M. B., R. L. Johnston, M. Inostroza-Ponta, A. H. Fox, E. Fortini, P. Moscato, M. E. Dinger and J. S. Mattick (2012). "Genome-wide analysis of long noncoding RNA stability." Genome Res **22**(5): 885-898.

Clark, M. B. and J. S. Mattick (2011). "Long noncoding RNAs in cell biology." Semin Cell Dev Biol **22**(4): 366-376.

Collins, F. S., E. S. Lander, J. Rogers, R. H. Waterston and I. H. G. S. Conso (2004). "Finishing the euchromatic sequence of the human genome." Nature **431**(7011): 931-945.

Conley, A. B. and I. King Jordan (2012). "Epigenetic regulation of human cis-natural antisense transcripts." Nucleic Acids Res **40**(4): 1438-1445.

Consortium, I. H. G. S. (1998). "Genome sequence of the nematode C. elegans: a platform for investigating biology." Science **282**(5396): 2012-2018.

Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-563.

Crick, F. H. (1958). "On protein synthesis." Symp Soc Exp Biol **12**: 138-163.

Dallol, A., A. Agathanggelou, S. L. Fenton, J. Ahmed-Choudhury, L. Hesson, M. D. Vos, G. J. Clark, J. Downward, E. R. Maher and F. Latif (2004). "RASSF1A interacts with microtubule-associated proteins and modulates microtubule dynamics." Cancer Res **64**(12): 4112-4116.

Dammann, R., C. Li, J. H. Yoon, P. L. Chin, S. Bates and G. P. Pfeifer (2000). "Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3." Nat Genet **25**(3): 315-319.

Dammann, R., U. Schagdarsurengin, C. Seidel, M. Strunnikova, M. Rastetter, K. Baier and G. P. Pfeifer (2005). "The tumor suppressor RASSF1A in human carcinogenesis: an update." Histol Histopathol **20**(2): 645-663.

Dani, C., J. M. Blanchard, M. Piechaczyk, S. El Sabouty, L. Marty and P. Jeanteur (1984). "Extreme instability of myc mRNA in normal and transformed human cells." Proc Natl Acad Sci U S A **81**(22): 7046-7050.

David, L., W. Huber, M. Granovskaia, J. Toedling, C. J. Palm, L. Bofkin, T. Jones, R. W. Davis and L. M. Steinmetz (2006). "A high-resolution map of transcription in the yeast genome." Proc Natl Acad Sci U S A **103**(14): 5320-5325.

Davuluri, R. V., I. Grosse and M. Q. Zhang (2001). "Computational identification of promoters and first exons in the human genome." Nat Genet **29**(4): 412-417.

Dinger, M. E., P. P. Amaral, T. R. Mercer and J. S. Mattick (2009). "Pervasive transcription of the eukaryotic genome: functional indices and conceptual implications." Brief Funct Genomic Proteomic **8**(6): 407-423.

Donninger, H., M. D. Vos and G. J. Clark (2007). "The RASSF1A tumor suppressor." J Cell Sci **120**(Pt 18): 3163-3172.

Dreijerink, K., E. Braga, I. Kuzmin, L. Geil, F. M. Duh, D. Angeloni, B. Zbar, M. I. Lerman, E. J. Stanbridge, J. D. Minna, A. Protopopov, J. Li, V. Kashuba, G. Klein and E. R. Zabarovsky (2001). "The candidate tumor suppressor gene, RASSF1A, from human chromosome 3p21.3 is involved in kidney tumorigenesis." Proc Natl Acad Sci U S A **98**(13): 7504-7509.

Ebralidze, A. K., F. C. Guibal, U. Steidl, P. Zhang, S. Lee, B. Bartholdy, M. A. Jorda, V. Petkova, F. Rosenbauer, G. Huang, T. Dayaram, J. Klupp, K. B. O'Brien, B. Will, M. Hoogenkamp, K. L. Borden, C. Bonifer and D. G. Tenen (2008). "PU.1 expression is modulated by the balance of functional sense and antisense RNAs regulated by a shared cis-regulatory element." Genes Dev **22**(15): 2085-2092.

Engelhardt, J. and P. F. Stadler (2012). "Hidden treasures in unspliced EST data." Theory Biosci **131**(1): 49-57.

Esteller, M. (2011). "Non-coding RNAs in human disease." Nat Rev Genet **12**(12): 861-874.

Fahrner, J. A., S. Eguchi, J. G. Herman and S. B. Baylin (2002). "Dependence of histone modifications and gene expression on DNA hypermethylation in cancer." Cancer Res **62**(24): 7213-7218.

Fajas, L., C. Paul, A. Vie, S. Estrach, R. Medema, J. M. Blanchard, C. Sardet and M. L. Vignais (2001). "Cyclin A is a mediator of p120E4F-dependent cell cycle arrest in G1." Mol Cell Biol **21**(8): 2956-2966.

Farajollahi, S. and S. Maas (2010). "Molecular diversity through RNA editing: a balancing act." Trends Genet **26**(5): 221-230.

Fejes-Toth, K., V. Sotirova, R. Sachidanandam, G. Assaf and J. Hannon G (2009). "Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs." Nature **457**(7232): 1028-1032.

Fenton, S. L., A. Dallol, A. Agathanggelou, L. Hesson, J. Ahmed-Choudhury, S. Baksh, C. Sardet, R. Dammann, J. D. Minna, J. Downward, E. R. Maher and F. Latif (2004). "Identification of the E1A-regulated transcription factor p120 E4F as an interacting partner of the RASSF1A candidate tumor suppressor gene." Cancer Res **64**(1): 102-107.

Flynn, R. A. and H. Y. Chang (2012). "Active chromatin and noncoding RNAs: an intimate relationship." Curr Opin Genet Dev **22**(2): 172-178.

Foley, C. J., H. Freedman, S. L. Choo, C. Onyskiw, N. Y. Fu, V. C. Yu, J. Tuszynski, J. C. Pratt and S. Baksh (2008). "Dynamics of RASSF1A/MOAP-1 association with death receptors." Mol Cell Biol **28**(14): 4520-4535.

Frith, M. C., M. Pheasant and J. S. Mattick (2005). "The amazing complexity of the human transcriptome." Eur J Hum Genet **13**(8): 894-897.

Gerstein, M. B., Z. J. Lu, E. L. Van Nostrand, C. Cheng, B. I. Arshinoff, T. Liu, K. Y. Yip, R. Robilotto, A. Rechtsteiner, K. Ikegami, P. Alves, A. Chateigner, M. Perry, M. Morris, R. K. Auerbach, X. Feng, J. Leng, A. Vielle, W. Niu, K. Rhrissorrakrai, A. Agarwal, R. P. Alexander, G. Barber, C. M. Brdlik, J. Brennan, J. J. Brouillet, A. Carr, M. S. Cheung, H. Clawson, S. Contrino, L. O. Dannenberg, A. F. Dernburg, A. Desai, L. Dick, A. C. Dose, J. Du, T. Egelhofer, S. Ercan, G. Euskirchen, B. Ewing, E. A. Feingold, R. Gassmann, P. J. Good, P. Green, F. Gullier, M. Gutwein, M. S. Guyer, L. Habegger, T. Han, J. G. Henikoff, S. R. Henz, A. Hinrichs, H. Holster, T. Hyman, A. L. Iniguez, J. Janette, M. Jensen, M. Kato, W. J. Kent, E. Kephart, V. Khivansara, E.

Khurana, J. K. Kim, P. Kolasinska-Zwierz, E. C. Lai, I. Latorre, A. Leahey, S. Lewis, P. Lloyd, L. Lochovsky, R. F. Lowdon, Y. Lubling, R. Lyne, M. MacCoss, S. D. Mackowiak, M. Mangone, S. McKay, D. Mecenas, G. Merrihew, D. M. Miller, 3rd, A. Muroyama, J. I. Murray, S. L. Ooi, H. Pham, T. Phippen, E. A. Preston, N. Rajewsky, G. Ratsch, H. Rosenbaum, J. Rozowsky, K. Rutherford, P. Ruzanov, M. Sarov, R. Sasidharan, A. Sboner, P. Scheid, E. Segal, H. Shin, C. Shou, F. J. Slack, C. Slightam, R. Smith, W. C. Spencer, E. O. Stinson, S. Taing, T. Takasaki, D. Vafeados, K. Voronina, G. Wang, N. L. Washington, C. M. Whittle, B. Wu, K. K. Yan, G. Zeller, Z. Zha, M. Zhong, X. Zhou, J. Ahringer, S. Strome, K. C. Gunsalus, G. Micklem, X. S. Liu, V. Reinke, S. K. Kim, L. W. Hillier, S. Henikoff, F. Piano, M. Snyder, L. Stein, J. D. Lieb and R. H. Waterston (2010). "Integrative analysis of the Caenorhabditis elegans genome by the modENCODE project." Science **330**(6012): 1775-1787.

Gingeras, T. R. (2007). "Origin of phenotypes: genes and transcripts." Genome Res **17**(6): 682-690.

Giovinazzi, S., C. R. Lindsay, V. M. Morozov, E. Escobar-Cabrera, M. K. Summers, H. S. Han, L. P. McIntosh and A. M. Ishov (2012). "Regulation of mitosis and taxane response by Daxx and Rassf1." Oncogene **31**(1): 13-26.

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

Gondran, P., F. Amiot, D. Weil and F. Dautry (1999). "Accumulation of mature mRNA in the nuclear fraction of mammalian cells." FEBS Lett **458**(3): 324-328.

Gong, X. Q., Y. A. Nedialkov and Z. F. Burton (2004). "Alpha-amanitin blocks translocation by human RNA polymerase II." J Biol Chem **279**(26): 27422-27427.

Goodliffe, J. M., E. Wieschaus and M. D. Cole (2005). "Polycomb mediates Myc autorepression and its transcriptional control of many loci in Drosophila." Genes Dev **19**(24): 2941-2946.

Guil, S., M. Soler, A. Portela, J. Carrere, E. Fonalleras, A. Gomez, A. Villanueva and M. Esteller (2012). "Intronic RNAs mediate EZH2 regulation of epigenetic targets." Nat Struct Mol Biol.

Guttman, M., I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn and E. S. Lander (2009). "Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals." Nature **458**(7235): 223-227.

Guttman, M., J. Donaghey, B. W. Carey, M. Garber, J. K. Grenier, G. Munson, G. Young, A. B. Lucas, R. Ach, L. Bruhn, X. Yang, I. Amit, A. Meissner, A. Regev, J. L. Rinn, D. E. Root and E. S. Lander (2011). "lincRNAs act in the circuitry controlling pluripotency and differentiation." Nature **477**(7364): 295-300.

Guttman, M., M. Garber, J. Z. Levin, J. Donaghey, J. Robinson, X. Adiconis, L. Fan, M. J. Koziol, A. Gnirke, C. Nusbaum, J. L. Rinn, E. S. Lander and A. Regev (2010). "Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs." Nat Biotechnol **28**(5): 503-510.

He, L., J. Liu, I. Collins, S. Sanford, B. O'Connell, C. J. Benham and D. Levens (2000). "Loss of FBP function arrests cellular proliferation and extinguishes c-myc expression." EMBO J **19**(5): 1034-1044.

Hesson, L., I. Bieche, D. Krex, E. Criniere, K. Hoang-Xuan, E. R. Maher and F. Latif (2004). "Frequent epigenetic inactivation of RASSF1A and BLU genes located within the critical 3p21.3 region in gliomas." Oncogene **23**(13): 2408-2419.

Hesson, L. B., W. N. Cooper and F. Latif (2007). "The role of RASSF1A methylation in cancer." Dis Markers **23**(1-2): 73-87.

Hung, T., Y. Wang, M. F. Lin, A. K. Koegel, Y. Kotake, G. D. Grant, H. M. Horlings, N. Shah, C. Umbricht, P. Wang, B. Kong, A. Langerod, A. L. Borresen-Dale, S. K. Kim, M. van de Vijver, S. Sukumar, M. L. Whitfield, M. Kellis, Y. Xiong, D. J. Wong and H. Y. Chang (2011). "Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters." Nat Genet **43**(7): 621-629.

Hutchinson, J. N., A. W. Ensminger, C. M. Clemson, C. R. Lynch, J. B. Lawrence and A. Chess (2007). "A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains." BMC Genomics **8**: 39.

Jacob, S. T., E. M. Sajdel and H. N. Munro (1970). "Specific action of alpha-amanitin on mammalian RNA polymerase protein." Nature **225**(5227): 60-62.

Ji, L., M. Nishizaki, B. Gao, D. Burbee, M. Kondo, C. Kamibayashi, K. Xu, N. Yen, E. N. Atkinson, B. Fang, M. I. Lerman, J. A. Roth and J. D. Minna (2002). "Expression of several genes in the human chromosome 3p21.3 homozygous deletion region by an adenovirus vector results in tumor suppressor activities in vitro and in vivo." Cancer Res **62**(9): 2715-2720.

Ji, P., S. Diederichs, W. Wang, S. Boing, R. Metzger, P. M. Schneider, N. Tidow, B. Brandt, H. Buerger, E. Bulk, M. Thomas, W. E. Berdel, H. Serve and C. Muller-Tidow (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): 8031-8041.

Kampa, D., J. Cheng, P. Kapranov, M. Yamanaka, S. Brubaker, S. Cawley, J. Drenkow, A. Piccolboni, S. Bekiranov, G. Helt, H. Tammana and T. R. Gingeras (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, S. P. Fodor and T. R. Gingeras (2002). "Large-scale transcriptional activity in chromosomes 21 and 22." Science **296**(5569): 916-919.

Kapranov, P., J. Cheng, S. Dike, D. A. Nix, R. Duttagupta, A. T. Willingham, P. F. Stadler, J. Hertel, J. Hackermuller, I. L. Hofacker, I. Bell, E. Cheung, J. Drenkow, E. Dumais, S. Patel, G. Helt, M. Ganesh, S. Ghosh, A. Piccolboni, V. Sementchenko, H. Tammana and T. R. Gingeras (2007). "RNA maps reveal new RNA classes and a possible function for pervasive transcription." Science **316**(5830): 1484-1488.

Kapranov, P., G. St Laurent, T. Raz, F. Ozsolak, C. P. Reynolds, P. H. Sorensen, G. Reaman, P. Milos, R. J. Arceci, J. F. Thompson and T. J. Triche (2010). "The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA." BMC Biol **8**: 149.

Kapranov, P., A. T. Willingham and T. R. Gingeras (2007). "Genome-wide transcription and the implications for genomic organization." Nat Rev Genet **8**(6): 413-423.

Katayama, S., Y. Tomaru, T. Kasukawa, K. Waki, M. Nakanishi, M. Nakamura, H. Nishida, C. C. Yap, M. Suzuki, J. Kawai, H. Suzuki, P. Carninci, Y. Hayashizaki, C. Wells, M. Frith, T. Ravasi, K. C. Pang, J. Hallinan, J. Mattick, D. A. Hume, L. Lipovich, S. Batalov, P. G. Engstrom, Y. Mizuno, M. A. Faghihi, A. Sandelin, A. M. Chalk, S. Mottagui-Tabar, Z. Liang, B. Lenhard and C. Wahlestedt (2005). "Antisense transcription in the mammalian transcriptome." Science **309**(5740): 1564-1566.

Kawamoto, K., S. T. Okino, R. F. Place, S. Urakami, H. Hirata, N. Kikuno, T. Kawakami, Y. Tanaka, D. Pookot, Z. Chen, S. Majid, H. Enokida, M. Nakagawa and R. Dahiya (2007). "Epigenetic modifications of RASSF1A gene through chromatin remodeling in prostate cancer." Clin Cancer Res **13**(9): 2541-2548.

Kedinger, C., M. Gniazdowski, J. L. Mandel, Jr., F. Gissinger and 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): 165-171.

Khalil, A. M., M. Guttman, M. Huarte, M. Garber, A. Raj, D. Rivea Morales, K. Thomas, A. Presser, B. E. Bernstein, A. van Oudenaarden, A. Regev, E. S. Lander and J. L. Rinn (2009). "Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression." Proc Natl Acad Sci U S A **106**(28): 11667-11672.

Khalil, A. M. and J. L. Rinn (2011). "RNA-protein interactions in human health and disease." Semin Cell Dev Biol **22**(4): 359-365.

Kitagawa, D., H. Kajiho, T. Negishi, S. Ura, T. Watanabe, T. Wada, H. Ichijo, T. Katada and H. Nishina (2006). "Release of RASSF1C from the nucleus by Daxx degradation links DNA damage and SAPK/JNK activation." EMBO J **25**(14): 3286-3297.

Kong, L., Y. Zhang, Z. Q. Ye, X. Q. Liu, S. Q. Zhao, L. Wei and G. Gao (2007). "CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine." Nucleic Acids Res **35**(Web Server issue): W345-349.

Kuzmin, I., J. W. Gillespie, A. Protopopov, L. Geil, K. Dreijerink, Y. Yang, C. D. Vocke, F. M. Duh, E. Zabarovsky, J. D. Minna, J. S. Rhim, M. R. Emmert-Buck, W. M. Linehan and M. I. Lerman (2002). "The RASSF1A tumor suppressor gene is inactivated in prostate tumors and suppresses growth of prostate carcinoma cells." Cancer Res **62**(12): 3498-3502.

Lee, Y., M. Kim, J. Han, K. H. Yeom, S. Lee, S. H. Baek and V. N. Kim (2004). "MicroRNA genes are transcribed by RNA polymerase II." EMBO J **23**(20): 4051-4060.

Lerman, M. I. and J. D. Minna (2000). "The 630-kb lung cancer homozygous deletion region on human chromosome 3p21.3: identification and evaluation of the resident candidate tumor suppressor genes. The International Lung Cancer Chromosome 3p21.3 Tumor Suppressor Gene Consortium." Cancer Res **60**(21): 6116-6133.

Li, J., F. Wang, A. Protopopov, A. Malyukova, V. Kashuba, J. D. Minna, M. I. Lerman, G. Klein and E. Zabarovsky (2004). "Inactivation of RASSF1C during in vivo tumor growth identifies it as a tumor suppressor gene." Oncogene **23**(35): 5941-5949.

Lindell, T. J., F. Weinberg, P. W. Morris, R. G. Roeder and W. J. Rutter (1970). "Specific inhibition of nuclear RNA polymerase II by alpha-amanitin." Science **170**(3956): 447-449.

Matallanas, D., D. Romano, K. Yee, K. Meissl, L. Kucerova, D. Piazzolla, M. Baccarini, J. K. Vass, W. Kolch and E. O'Neill (2007). "RASSF1A elicits apoptosis through an MST2 pathway directing proapoptotic transcription by the p73 tumor suppressor protein." Mol Cell **27**(6): 962-975.

Mathe, E. (2004). "RASSF1A, the new guardian of mitosis." Nat Genet **36**(2): 117-118.

Mattick, J. S. (2007). "A new paradigm for developmental biology." J Exp Biol **210**(Pt 9): 1526-1547.

Mattick, J. S. (2011). "The central role of RNA in human development and cognition." FEBS Lett **585**(11): 1600-1616.

Mattick, J. S. (2012). "RNA driving the epigenetic bus." EMBO J **31**(3): 515-516.

Mercer, T. R., M. E. Dinger and J. S. Mattick (2009). "Long non-coding RNAs: insights into functions." Nat Rev Genet **10**(3): 155-159.

Mercer, T. R., D. J. Gerhardt, M. E. Dinger, J. Crawford, C. Trapnell, J. A. Jeddeloh, J. S. Mattick and J. L. Rinn (2012). "Targeted RNA sequencing reveals the deep complexity of the human transcriptome." Nat Biotechnol **30**(1): 99-104.

Nagasaki, H., M. Arita, T. Nishizawa, M. Suwa and O. Gotoh (2005). "Species-specific variation of alternative splicing and transcriptional initiation in six eukaryotes." Gene **364**: 53-62.

Nakaya, H. I., P. P. Amaral, R. Louro, A. Lopes, A. A. Fachel, Y. B. Moreira, T. A. El-Jundi, A. M. da Silva, E. M. Reis and S. Verjovski-Almeida (2007). "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 **8**(3): R43.

Nayak, V., C. Xu and J. Min (2011). "Composition, recruitment and regulation of the PRC2 complex." Nucleus **2**(4).

Newton, A. C. (1995). "Protein kinase C. Seeing two domains." Curr Biol **5**(9): 973-976.

Numata, K., A. Kanai, R. Saito, S. Kondo, J. Adachi, L. G. Wilming, D. A. Hume, Y. Hayashizaki and M. Tomita (2003). "Identification of putative noncoding RNAs among the RIKEN mouse full-length cDNA collection." Genome Res **13**(6B): 1301-1306.

Oceandy, D., A. Pickard, S. Prehar, M. Zi, T. M. Mohamed, P. J. Stanley, F. Baudoin-Stanley, R. Nadif, S. Tommasi, G. P. Pfeifer, A. L. Armesilla, E. J. Cartwright and L. Neyses (2009). "Tumor suppressor Ras-association domain family 1 isoform A is a novel regulator of cardiac hypertrophy." Circulation **120**(7): 607-616.

Ota, T., Y. Suzuki, T. Nishikawa, T. Otsuki, T. Sugiyama, R. Irie, A. Wakamatsu, K. Hayashi, H. Sato, K. Nagai, K. Kimura, H. Makita, M. Sekine, M. Obayashi, T. Nishi, T. Shibahara, T. Tanaka, S. Ishii, J. Yamamoto, K. Saito, Y. Kawai, Y. Isono, Y. Nakamura, K. Nagahari, K. Murakami, T. Yasuda, T. Iwayanagi, M. Wagatsuma, A. Shiratori, H. Sudo, T. Hosoiri, Y. Kaku, H. Kodaira, H. Kondo, M. Sugawara, M. Takahashi, K. Kanda, T. Yokoi, T. Furuya, E. Kikkawa, Y. Omura, K. Abe, K. Kamihara, N. Katsuta, K. Sato, M. Tanikawa, M. Yamazaki, K. Ninomiya, T. Ishibashi, H. Yamashita, K. Murakawa, K. Fujimori, H. Tanai, M. Kimata, M. Watanabe, S. Hiraoka, Y. Chiba, S. Ishida, Y. Ono, S. Takiguchi, S. Watanabe, M. Yosida, T. Hotuta, J. Kusano, K. Kanehori, A. Takahashi-Fujii, H. Hara, T. O. Tanase, Y. Nomura, S. Togiya, F. Komai, R. Hara, K. Takeuchi, M. Arita, N. Imose, K. Musashino, H. Yuuki, A. Oshima, N. Sasaki, S. Aotsuka, Y. Yoshikawa, H. Matsunawa, T. Ichihara, N. Shiohata, S. Sano, S. Moriya, H. Momiyama, N. Satoh, S. Takami, Y. Terashima, O. Suzuki, S. Nakagawa, A. Senoh, H. Mizoguchi, Y. Goto, F. Shimizu, H. Wakebe, H. Hishigaki, T. Watanabe, A. Sugiyama, M. Takemoto, B. Kawakami, K. Watanabe, A. Kumagai, S. Itakura, Y. Fukuzumi, Y. Fujimori, M. Komiyama, H. Tashiro, A. Tanigami, T. Fujiwara, T. Ono, K. Yamada, Y. Fujii, K. Ozaki, M. Hirao, Y. Ohmori, A. Kawabata, T. Hikiji, N. Kobatake, H. Inagaki, Y. Ikema, S. Okamoto, R. Okitani, T. Kawakami, S. Noguchi, T. Itoh, K. Shigeta, T. Senba, K. Matsumura, Y. Nakajima, T. Mizuno, M. Morinaga, M. Sasaki, T. Togashi, M. Oyama, H. Hata, T. Komatsu, J. Mizushima-Sugano, T. Satoh, Y. Shirai, Y. Takahashi, K. Nakagawa, K. Okumura, T. Nagase, N. Nomura, H. Kikuchi, Y. Masuho, R. Yamashita, K. Nakai, T. Yada, O. Ohara, T. Isogai and S. Sugano (2004). "Complete sequencing and characterization of 21,243 full-length human cDNAs." Nat Genet **36**(1): 40-45.

Palakurthy, R. K., N. Wajapeyee, M. K. Santra, C. Gazin, L. Lin, S. Gobeil and M. R. Green (2009). "Epigenetic silencing of the RASSF1A tumor suppressor gene through HOXB3-mediated induction of DNMT3B expression." Mol Cell **36**(2): 219-230.

Pandey, R. R., T. Mondal, F. Mohammad, S. Enroth, L. Redrup, J. Komorowski, T. Nagano, D. Mancini-Dinardo and C. Kanduri (2008). "Kcnq1ot1 antisense noncoding RNA mediates lineage-specific transcriptional silencing through chromatin-level regulation." Mol Cell **32**(2): 232-246.

Ponjavic, J., C. P. Ponting and G. Lunter (2007). "Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs." Genome Res **17**(5): 556-565.

Ponting, C. P. and D. R. Benjamin (1996). "A novel family of Ras-binding domains." Trends Biochem Sci **21**(11): 422-425.

Ponting, C. P., P. L. Oliver and W. Reik (2009). "Evolution and functions of long noncoding RNAs." Cell **136**(4): 629-641.

Prensner, J. R. and A. M. Chinnaiyan (2011). "The emergence of lncRNAs in cancer biology." Cancer Discov **1**(5): 391-407.

Raha, D., Z. Wang, Z. Moqtaderi, L. Wu, G. Zhong, M. Gerstein, K. Struhl and M. Snyder (2010). "Close association of RNA polymerase II and many transcription factors with Pol III genes." Proc Natl Acad Sci U S A **107**(8): 3639-3644.

Rahl, P. B., C. Y. Lin, A. C. Seila, R. A. Flynn, S. McCuine, C. B. Burge, P. A. Sharp and R. A. Young (2010). "c-Myc regulates transcriptional pause release." Cell **141**(3): 432-445.

Reeves, M. E., S. W. Baldwin, M. L. Baldwin, S. T. Chen, J. M. Moretz, R. J. Aragon, X. Li, D. D. Strong, S. Mohan and Y. G. Amaar (2010). "Ras-association domain family 1C protein promotes breast cancer cell migration and attenuates apoptosis." BMC Cancer **10**: 562.

Reis, E. M., H. I. Nakaya, R. Louro, F. C. Canavez, A. V. Flatschart, G. T. Almeida, C. M. Egidio, A. C. Paquola, A. A. Machado, F. Festa, D. Yamamoto, R. Alvarenga, C. C. da Silva, G. C. Brito, S. D. Simon, C. A. Moreira-Filho, K. R. Leite, L. H. Camara-Lopes, F. S. Campos, E. Gimba, G. M. Vignal, H. El-Dorry, M. C. Sogayar, M. A. Barcinski, A. M. da Silva and S. Verjovski-Almeida (2004). "Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer." Oncogene **23**(39): 6684-6692.

Reis, E. M. and S. Verjovski-Almeida (2012). "Perspectives of Long Non-Coding RNAs in Cancer Diagnostics." Front Genet **3**: 32.

Rinn, J. L. and H. Y. Chang (2012). "Genome Regulation by Long Noncoding RNAs." Annu Rev Biochem **81**: 145-166.

Rinn, J. L., G. Euskirchen, P. Bertone, R. Martone, N. M. Luscombe, S. Hartman, P. M. Harrison, F. K. Nelson, P. Miller, M. Gerstein, S. Weissman and M. Snyder (2003). "The transcriptional activity of human Chromosome 22." Genes Dev **17**(4): 529-540.

Rinn, J. L., M. Kertesz, J. K. Wang, S. L. Squazzo, X. Xu, S. A. Brugmann, L. H. Goodnough, J. A. Helms, P. J. Farnham, E. Segal and H. Y. Chang (2007). "Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs." Cell **129**(7): 1311-1323.

Rong, R., W. Jin, J. Zhang, M. S. Sheikh and Y. Huang (2004). "Tumor suppressor RASSF1A is a microtubule-binding protein that stabilizes microtubules and induces G2/M arrest." Oncogene **23**(50): 8216-8230.

Roy, S., J. Ernst, P. V. Kharchenko, P. Kheradpour, N. Negre, M. L. Eaton, J. M. Landolin, C. A. Bristow, L. Ma, M. F. Lin, S. Washietl, B. I. Arshinoff, F. Ay, P. E. Meyer, N. Robine, N. L. Washington, L. Di Stefano, E. Berezikov, C. D. Brown, R. Candeias, J. W. Carlson, A. Carr, I. Jungreis, D. Marbach, R. Sealfon, M. Y. Tolstorukov, S. Will, A. A. Alekseyenko, C. Artieri, B. W. Booth, A. N. Brooks, Q. Dai, C. A. Davis, M. O. Duff, X. Feng, A. A. Gorchakov, T. Gu, J. G. Henikoff, P. Kapranov, R. Li, H. K. MacAlpine, J. Malone, A. Minoda, J. Nordman, K. Okamura, M. Perry, S. K. Powell, N. C. Riddle, A. Sakai, A. Samsonova, J. E. Sandler, Y. B. Schwartz, N. Sher, R. Spokony, D. Sturgill, M. van Baren, K. H. Wan, L. Yang, C. Yu, E. Feingold, P. Good, M. Guyer, R. Lowdon, K. Ahmad, J. Andrews, B. Berger, S. E. Brenner, M. R. Brent, L. Cherbas, S. C. Elgin, T. R. Gingeras, R. Grossman, R. A. Hoskins, T. C. Kaufman, W. Kent, M. I. Kuroda, T. Orr-Weaver, N. Perrimon, V. Pirrotta, J. W. Posakony, B. Ren, S. Russell, P. Cherbas, B. R. Graveley, S. Lewis, G. Micklem, B. Oliver, P. J. Park, S. E. Celniker, S. Henikoff, G. H. Karpen, E. C. Lai, D. M. MacAlpine, L. D. Stein, K. P. White and M. Kellis (2010). "Identification of

functional elements and regulatory circuits by Drosophila modENCODE." Science **330**(6012): 1787-1797.

Rozen, S. and H. Skaletsky (2000). "Primer3 on the WWW for general users and for biologist programmers." Methods Mol Biol **132**: 365-386.

Sambrook, J. and D. W. Russell (2001). Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Sayers, E. W., T. Barrett, D. A. Benson, E. Bolton, S. H. Bryant, K. Canese, V. Chetvernin, D. M. Church, M. Dicuccio, S. Federhen, M. Feolo, L. Y. Geer, W. Helmberg, Y. Kapustin, D. Landsman, D. J. Lipman, Z. Lu, T. L. Madden, T. Madej, D. R. Maglott, A. Marchler-Bauer, V. Miller, I. Mizrachi, J. Ostell, A. Panchenko, K. D. Pruitt, G. D. Schuler, E. Sequeira, S. T. Sherry, M. Shumway, K. Sirotkin, D. Slotta, A. Souvorov, G. Starchenko, T. A. Tatusova, L. Wagner, Y. Wang, W. John Wilbur, E. Yaschenko and J. Ye (2010). "Database resources of the National Center for Biotechnology Information." Nucleic Acids Res **38**(Database issue): D5-16.

Scheel, H. and K. Hofmann (2003). "A novel interaction motif, SARAH, connects three classes of tumor suppressor." Curr Biol **13**(23): R899-900.

Shivakumar, L., J. Minna, T. Sakamaki, R. Pestell and M. A. White (2002). "The RASSF1A tumor suppressor blocks cell cycle progression and inhibits cyclin D1 accumulation." Mol Cell Biol **22**(12): 4309-4318.

Sobell, H. M. (1985). "Actinomycin and DNA transcription." Proc Natl Acad Sci U S A **82**(16): 5328-5331.

Song, M. S., J. S. Chang, S. J. Song, T. H. Yang, H. Lee and D. S. Lim (2005). "The centrosomal protein RAS association domain family protein 1A (RASSF1A)-binding protein 1 regulates mitotic progression by recruiting RASSF1A to spindle poles." J Biol Chem **280**(5): 3920-3927.

Song, M. S., S. J. Song, N. G. Ayad, J. S. Chang, J. H. Lee, H. K. Hong, H. Lee, N. Choi, J. Kim, H. Kim, J. W. Kim, E. J. Choi, M. W. Kirschner and D. S. Lim (2004). "The tumour suppressor RASSF1A regulates mitosis by inhibiting the APC-Cdc20 complex." Nat Cell Biol **6**(2): 129-137.

Spizzo, R., M. I. Almeida, A. Colombatti and G. A. Calin (2012). "Long non-coding RNAs and cancer: a new frontier of translational research?" Oncogene.

Starlard-Davenport, A., V. P. Tryndyak, S. R. James, A. R. Karpf, J. R. Latendresse, F. A. Beland and I. P. Pogribny (2010). "Mechanisms of epigenetic silencing of the Rassf1a gene during estrogen-induced breast carcinogenesis in ACI rats." Carcinogenesis **31**(3): 376-381.

Struhl, K. (2007). "Transcriptional noise and the fidelity of initiation by RNA polymerase II." Nat Struct Mol Biol **14**(2): 103-105.

Strunnikova, M., U. Schagdarsurengin, A. Kehlen, J. C. Garbe, M. R. Stampfer and R. Dammann (2005). "Chromatin inactivation precedes de novo DNA methylation during the progressive epigenetic silencing of the RASSF1A promoter." Mol Cell Biol **25**(10): 3923-3933.

Sunwoo, H., M. E. Dinger, J. E. Wilusz, P. P. Amaral, J. S. Mattick and D. L. Spector (2009). "MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles." Genome Res **19**(3): 347-359.

Taft, R. J., E. A. Glazov, N. Cloonan, C. Simons, S. Stephen, G. J. Faulkner, T. Lassmann, A. R. Forrest, S. M. Grimmond, K. Schroder, K. Irvine, T. Arakawa, M. Nakamura, A. Kubosaki, K. Hayashida, C. Kawazu, M. Murata, H. Nishiyori, S. Fukuda, J. Kawai, C. O. Daub, D. A. Hume, H. Suzuki, V. Orlando, P. Carninci, Y. Hayashizaki and J. S.

Mattick (2009). "Tiny RNAs associated with transcription start sites in animals." Nat Genet **41**(5): 572-578.

Taft, R. J., K. C. Pang, T. R. Mercer, M. Dinger and J. S. Mattick (2010). "Non-coding RNAs: regulators of disease." J Pathol **220**(2): 126-139.

Taft, R. J., M. Pheasant and J. S. Mattick (2007). "The relationship between non-protein-coding DNA and eukaryotic complexity." Bioessays **29**(3): 288-299.

Tahira, A. C., M. S. Kubrusly, M. F. Faria, B. Dazzani, R. S. Fonseca, V. Maracaja-Coutinho, S. Verjovski-Almeida, M. C. Machado and E. M. Reis (2011). "Long noncoding intronic RNAs are differentially expressed in primary and metastatic pancreatic cancer." Mol Cancer **10**: 141.

Topisirovic, I., B. Culjkovic, N. Cohen, J. M. Perez, L. Skrabanek and K. L. Borden (2003). "The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth." EMBO J **22**(3): 689-703.

Trapnell, C., L. Pachter and S. L. Salzberg (2009). "TopHat: discovering splice junctions with RNA-Seq." Bioinformatics **25**(9): 1105-1111.

Tsai, M. C., O. Manor, Y. Wan, N. Mosammaparast, J. K. Wang, F. Lan, Y. Shi, E. Segal and H. Y. Chang (2010). "Long noncoding RNA as modular scaffold of histone modification complexes." Science **329**(5992): 689-693.

Tycowski, K. T., M. D. Shu and J. A. Steitz (1993). "A small nucleolar RNA is processed from an intron of the human gene encoding ribosomal protein S3." Genes Dev **7**(7A): 1176-1190.

van Bakel, H., C. Nislow, B. J. Blencowe and T. R. Hughes (2010). "Most "dark matter" transcripts are associated with known genes." PLoS Biol **8**(5): e1000371.

van der Weyden, L. and D. J. Adams (2007). "The Ras-association domain family (RASSF) members and their role in human tumourigenesis." Biochim Biophys Acta **1776**(1): 58-85.

Vire, E., C. Brenner, R. Deplus, L. Blanchon, M. Fraga, C. Didelot, L. Morey, A. Van Eynde, D. Bernard, J. M. Vanderwinden, M. Bollen, M. Esteller, L. Di Croce, Y. de Launoit and F. Fuks (2006). "The Polycomb group protein EZH2 directly controls DNA methylation." Nature **439**(7078): 871-874.

Vos, M. D., A. Dallol, K. Eckfeld, N. P. Allen, H. Donninger, L. B. Hesson, D. Calvisi, F. Latif and G. J. Clark (2006). "The RASSF1A tumor suppressor activates Bax via MOAP-1." J Biol Chem **281**(8): 4557-4563.

Vos, M. D., C. A. Ellis, A. Bell, M. J. Birrer and G. J. Clark (2000). "Ras uses the novel tumor suppressor RASSF1 as an effector to mediate apoptosis." J Biol Chem **275**(46): 35669-35672.

Wang, X., S. Arai, X. Song, D. Reichart, K. Du, G. Pascual, P. Tempst, M. G. Rosenfeld, C. K. Glass and R. Kurokawa (2008). "Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription." Nature **454**(7200): 126-130.

Wang, X., X. Song, C. K. Glass and M. G. Rosenfeld (2011). "The long arm of long noncoding RNAs: roles as sensors regulating gene transcriptional programs." Cold Spring Harb Perspect Biol **3**(1): a003756.

Wapinski, O. and H. Y. Chang (2011). "Long noncoding RNAs and human disease." Trends Cell Biol **21**(6): 354-361.

Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S. E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M. R. Brent, D. G. Brown, S. D. Brown, C. Bult, J. Burton, J. Butler, R. D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A. T. Chinwalla, D. M. Church, M. Clamp, C. Clee, F. S. Collins, L. L. Cook, R. R. Copley, A. Coulson,

O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K. D. Delehaunty, J. Deri, E. T. Dermitzakis, C. Dewey, N. J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D. M. Dunn, S. R. Eddy, L. Elnitski, R. D. Emes, P. Eswara, E. Eyras, A. Felsenfeld, G. A. Fewell, P. Flicek, K. Foley, W. N. Frankel, L. A. Fulton, R. S. Fulton, T. S. Furey, D. Gage, R. A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T. A. Graves, E. D. Green, S. Gregory, R. Guigo, M. Guyer, R. C. Hardison, D. Haussler, Y. Hayashizaki, L. W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D. B. Jaffe, L. S. Johnson, M. Jones, T. A. Jones, A. Joy, M. Kamal, E. K. Karlsson, D. Karolchik, A. Kasprzyk, J. Kawai, E. Keibler, C. Kells, W. J. Kent, A. Kirby, D. L. Kolbe, I. Korf, R. S. Kucherlapati, E. J. Kulbokas, D. Kulp, T. Landers, J. P. Leger, S. Leonard, I. Letunic, R. Levine, J. Li, M. Li, C. Lloyd, S. Lucas, B. Ma, D. R. Maglott, E. R. Mardis, L. Matthews, E. Mauceli, J. H. Mayer, M. McCarthy, W. R. McCombie, S. McLaren, K. McLay, J. D. McPherson, J. Meldrim, B. Meredith, J. P. Mesirov, W. Miller, T. L. Miner, E. Mongin, K. T. Montgomery, M. Morgan, R. Mott, J. C. Mullikin, D. M. Muzny, W. E. Nash, J. O. Nelson, M. N. Nhan, R. Nicol, Z. Ning, C. Nusbaum, M. J. O'Connor, Y. Okazaki, K. Oliver, E. Overton-Larty, L. Pachter, G. Parra, K. H. Pepin, J. Peterson, P. Pevzner, R. Plumb, C. S. Pohl, A. Poliakov, T. C. Ponce, C. P. Ponting, S. Potter, M. Quail, A. Reymond, B. A. Roe, K. M. Roskin, E. M. Rubin, A. G. Rust, R. Santos, V. Sapojnikov, B. Schultz, J. Schultz, M. S. Schwartz, S. Schwartz, C. Scott, S. Seaman, S. Searle, T. Sharpe, A. Sheridan, R. Shownkeen, S. Sims, J. B. Singer, G. Slater, A. Smit, D. R. Smith, B. Spencer, A. Stabenau, N. Stange-Thomann, C. Sugnet, M. Suyama, G. Tesler, J. Thompson, D. Torrents, E. Trevaskis, J. Tromp, C. Ucla, A. Ureta-Vidal, J. P. Vinson, A. C. Von Niederhausern, C. M. Wade, M. Wall, R. J. Weber, R. B. Weiss, M. C. Wendl, A. P. West, K. Wetterstrand, R. Wheeler, S. Whelan, J. Wierzbowski, D. Willey, S. Williams, R. K. Wilson, E. Winter, K. C. Worley, D. Wyman, S. Yang, S. P. Yang, E. M. Zdobnov, M. C. Zody and E. S. Lander (2002). "Initial sequencing and comparative analysis of the mouse genome." Nature **420**(6915): 520-562.

Wilusz, J. E., H. Sunwoo and D. L. Spector (2009). "Long noncoding RNAs: functional surprises from the RNA world." Genes Dev **23**(13): 1494-1504.

Wright, M. W. and E. A. Bruford (2011). "Naming 'junk': human non-protein coding RNA (ncRNA) gene nomenclature." Hum Genomics **5**(2): 90-98.

Yamamoto, T., S. Taya and K. Kaibuchi (1999). "Ras-induced transformation and signaling pathway." J Biochem **126**(5): 799-803.

Yu, H. (2007). "Cdc20: a WD40 activator for a cell cycle degradation machine." Mol Cell **27**(1): 3-16.