UNIVERSIDADE DE SÃO PAULO

INSTITUTO DE QUÍMICA

**Busca e identificação de genes regulados
pelo Antígeno MT do vírus Polioma
na transformação maligna de células Balb-3T3**

Leonardo de Oliveira Rodrigues

Orientadora: Profa. Dra. Mari Cleide Sogayar

Tese de doutoramento apresentada ao

Departamento de Bioquímica da USP

Data de depósito do trabalho na SPG: 12/03/2007

(Doze de Março de Dois mil e sete)

Aos meus pais por terem me dado todo o amor e apoio necessário para eu chegar aonde cheguei...

A minha esposa, amiga, colega de turma, companheira de trabalho e eterno amor da minha vida Fe, por estar sempre ao meu lado e fazer de cada dia da minha vida um dia mais feliz...

Ao meu irmão Pe e minhas irmãs Dani e Caro, pelo eterno amor, carinho, amizade, risadas, companheirismo e apoio incondicional...

AGRADECIMENTOS

À professora Mari, por ter me guiado desde a Iniciação Científica até hoje e pela eterna paciência, apoio e amizade.

À Profa. Lucile Winter, por ter me apresentado o mundo da Biologia Molecular.

Ao Mario Bengtson, pela amizade e pela paciência e empenho em me ensinar os primeiros passos da Biologia Molecular.

À Sheila, por ter me guiado nos primeiros passos deste projeto.

À Fernanda Festa, por ter me acompanhado e incentivado desde o dia em que descobri o que era um PCR até...

Aos amigos Wagner, Antero, Theri, Rita, Fefesta, Pipoca, Lu e Karin, por terem feito do LBCM, e da copinha, um dos lugares mais agradáveis e divertidos de se trabalhar.

Aos que já saíram do Lab, mas sempre presentes FeOrtis, Sheila, Maki, LuHe, Mario, Katlin, Colin, Wagner e Pips.

Às ICs Juliane e Letícia, por terem sido meu braço direito em vários pontos do projeto.

Aos amigos do laboratório Adriano, Carol, Fujita, Anna Carla, Antero, Cristiane, Débora, Fernanda, Fernando, Flavia, Gisella, Gustavo, Helena, Helene, Humberto, Icaro, Ilana, Isnar, Juan, Kairn, Letícia Terra, Letícia Labriola, Luciana Cruz, Luciana Gomes, Maic, Marica, Marcos, Mari, Mria das Graças, Maria Fernanda, Marina, Patrícia dos Santos, Patrícia Kossugue, Rafael, Renata, Ricardo Corrêa, Ricardo de Oliveira, Rita, Sandra, Tatiana, Tatiane, Theri, Thiago e Zizi pelo constante apoio e amizade.

Aos eternos amigos da ETE Vivian, Tic, Piu, Dr. Zunto, Ticiane, Gustavo, Daniela, Fabiana, Bruce, Denise, Lucio, Murilo e Rocinha pelo apoio, risadas e rummikubs nas horas vagas.

Aos amigos do CM Tie, Daniel, Bruno, Marcos e Fernanda, pela amizade mantida ainda que á distância.

À Theri/Ricardo, pelas pizzas, risadas e momentos de descontração.

Às *roomates* Mari e a Rapha, pela amizade e por terem ajudado a fazer a distância de casa algo agradável.

Ao sogro Paulo, sogra Marta e cunhada Rê, por terem sido uma segunda família por todos esses anos.

Ao Prof. Pio Colepicolo e Adriana, pelo apoio e colabroação durante o projeto.

À minha família, por estar sempre ao meu lado ajudando no que fosse necessário.

Ao Instituto de Química-USP pelo apoio institucional, e a todos os seus funcionários pelo serviços prestados

Aos programadores que desenvolvem ferramentas de bioinformática, por fazerem desse mundo um lugar melhor para se viver.

APOIO FINANCEIRO

FAPESP: Fundação de Amparo à Pesquisa do Estado de São Paulo

CNPq: Conselho Nacional de Desenvolvimento Científico e Tecnológico

FINEP: Financiadora de Estudos e Projetos

PRP-USP: Pró-Reitoria de Pesquisa - USP

ÍNDICE

RESUMO.....	I
ABSTRACT.....	III
ABREVIATURAS	V
1. INTRODUÇÃO.....	1
1.1. Vírus.....	1
1.1.1. Histórico.....	1
1.1.2. Propriedades das partículas virais	2
1.1.3. Ciclo de vida viral	4
1.2. Vírus e Câncer.....	8
1.2.1. Controle da proliferação celular	8
1.2.2. Oncogenes e Genes Supressores de Tumor.....	10
1.2.3. Transformação celular por vírus tumorais	12
1.3. Vírus Polioma (Py)	15
1.3.1. Antígeno LT	19
1.3.2. Antígeno ST	19
1.3.3. Antígeno MT	20
1.4. Busca de genes diferencialmente expressos	29
2. OBJETIVOS	31
2.1. Objetivo Geral.....	31
2.2. Objetivos Específicos.....	31
3. MATERIAIS E MÉTODOS.....	32
3.1. Materiais	32
3.1.1. Plasmídeos	32
3.1.2. Linhagens celulares	33
3.1.3. Soluções e Meio de Cultura para células de mamíferos	34
3.1.4. Soluções e Meio de Cultura para bactérias	34
3.1.5. Isótopos radioativos.....	34
3.1.6. Soluções.....	34
3.2. Cultura de linhagens celulares.....	36

3.3.	Extração e purificação de RNA.....	36
3.4.	Síntese de cDNA.....	37
3.4.1.	Síntese da primeira fita do cDNA	37
3.4.2.	Controle da qualidade do cDNA	37
3.4.3.	Purificação do mRNA.....	38
3.4.4.	Síntese de cDNA dupla fita (<i>Smart cDNA Synthesis Kit</i>).....	38
3.5.	Sequenciamento de DNA	39
3.6.	Subtração dos genes diferencialmente expressos.....	40
3.6.1.	Geração das bibliotecas subtraídas através do RDA.....	40
3.6.2.	Clonagem das bibliotecas subtraídas	43
3.6.3.	Rastreamento dos clones diferencialmente expressos.....	43
3.6.4.	Síntese das sondas radioativas.....	44
3.6.5.	Hibridização dos macroarranjos de cDNA.....	44
3.6.6.	Análise da densitometria	45
3.7.	Membranas comerciais	45
3.7.1.	Hibridização da membrana	45
3.8.	Confirmação da expressão diferencial por RT-PCR.....	46
3.8.1.	Padronizações	46
3.8.2.	Confirmação da expressão diferencial.....	47
3.9.	Confirmação da expressão diferencial por <i>Real-Time PCR</i>	47
3.9.1.	Padronização da concentração dos <i>primers</i>	47
3.9.2.	Padronização da eficiência dos <i>primers</i>	48
3.9.3.	Confirmação da expressão diferencial.....	49
3.9.4.	Análise estatística dos dados.....	50
3.10.	<i>CodeLink bioarrays</i>	51
3.10.1.	Preparação da sonda.....	51
3.10.2.	Hibridização e lavagem	53
3.11.	Amplificação do cDNA completo do gene MN1	56
3.11.1.	Amplificação a partir de <i>primers</i> baseados em ESTs.....	56
3.11.2.	Amplificação utilizando a metodologia SMART RACE.....	57
3.12.	Inibição da expressão do gene MN1	59
3.12.1.	Preparação de DNA plasmideal em média escala.....	60
3.12.2.	Transfecção para produção de partículas virais.....	60
3.12.3.	Coleta e concentração dos sobrenadantes virais	60
3.12.4.	Infecção com lentivírus	60

3.12.5. Confirmação do RNAi	61
3.12.6. Curva de crescimento	61
3.12.7. Formação de colônia em meio semi-sólido	61
3.13. Amplificação do cDNA completo do gene Dlk1	62
3.13.1. Ligação do produto amplificado ao vetor.....	62
3.13.2. Transformação por choque térmico	63
3.13.3. Confirmação da clonagem.....	63
3.13.4. Preparação de DNA plasmideal em média escala.....	64
3.13.5. Transfecção de células empacotadoras	64
3.13.6. Coleta dos sobrenadantes virais	64
3.13.7. Infecção com retrovírus.....	64
3.13.8. Curva de crescimento	65
4. RESULTADOS	66
4.1. Busca de genes diferencialmente expressos em celular MTWT através de membranas comerciais	66
4.1.1. Densitometria da membrana	68
4.2. Geração de bibliotecas subtraídas através do RDA	70
4.2.1. Geração das representações	70
4.2.2. Busca dos genes diferencialmente expressos	72
4.2.3. Seqüenciamento e análise dos resultados	75
4.3. <i>CodeLink bioarrays</i>	76
4.3.1. Marcação das sondas, hibridização e lavagens.....	76
4.3.2. Aquisição e análise dos dados	76
4.3.3. Análise dos genes selecionados	86
4.4. Análise da expressão dos genes candidatos por RT-PCR	89
4.4.1. Síntese da primeira fita do cDNA	89
4.4.2. Controle da qualidade do cDNA	89
4.4.3. Padronização das reações de RT-PCR	90
4.4.4. Confirmação da expressão diferencial.....	92
4.5. Confirmação dos genes por <i>Real-Time PCR</i>	95
4.5.1. Padronização da concentração dos <i>primers</i>	95
4.5.2. Análise da eficiência dos <i>primers</i>	96
4.5.3. Padronização do gene <i>housekeeping</i>	97
4.5.4. Análise da expressão diferencial	99

4.6.	Análise funcional do gene MN1	108
4.6.1.	Clonagem do cDNA completo do gene MN1	109
4.6.2.	Inibição de MN1 através de RNA de interferência	116
4.7.	Clonagem do cDNA completo correspondente ao gene Dlk1	121
4.7.1.	Amplificação do transcrito completo de Dlk1	122
5.	DISCUSSÃO	130
5.1.	Busca de genes regulados por MT através das membranas comerciais e do RDA	130
5.2.	Análise dos genes regulados por MT através do CodeLink <i>Bioarrays</i> .	135
6.	CONCLUSÕES	142
7.	BIBLIOGRAFIA	144
8.	APÊNDICES	163

ÍNDICE DE FIGURAS

Figura 1 - Esquema da classificação dos principais vírus animais.....	4
Figura 2 – Formação do mRNA a partir do material genético viral	6
Figura 3 – Ciclo celular	9
Figura 4 - Camundongo infectado com Polioma Vírus	16
Figura 5 - Mapa representativo do genoma do Vírus Polioma	18
Figura 6 - Principais vias de transdução de sinal mediada por MT.....	22
Figura 7 - Esquema do RDA	42
Figura 8 - Saída gráfica do programa ABI-Manager SDS	49
Figura 9 - Aplicação da solução de hibridização do <i>bioarray</i>	54
Figura 10 - Representação esquemática do RACE-5' e RACE-3'	57
Figura 11 - Membranas comerciais	67
Figura 12 – Gráfico de dispersão do sinal das membranas comerciais.....	69
Figura 13 - Padronização da amplificação do cDNA de PLJ e MTWT.....	72
Figura 14 – Macroarranjos de cDNAs proveniente do RDA	73
Figura 15 – Gráfico de dispersão dos macroarranjos do RDA	74
Figura 16 – Lâminas do <i>CodeLink Bioarrays</i>	77
Figura 17 - Telas de saída do <i>CodeLink System Software</i>	78
Figura 18 – Gráficos de dispersão após diferentes métodos de normalização	82
Figura 19 - Gráfico MA do <i>CodeLink</i>	85
Figura 20 - Tela de saída do programa <i>PubGene</i>	86
Figura 21 – Tela de saída do <i>FatiGO</i>	87
Figura 22 - Gel de agarose para confirmação da qualidade dos cDNAs.....	90
Figura 23 – Padronização da amplificação do cDNA de GAPDH para o RT-PCR...	91
Figura 24 - Resultado final da padronização de GAPDH	92
Figura 25 – RT-PCR dos genes selecionados a partir do RDA.	93
Figura 26 - Densitometria em três orientações diferentes.	94
Figura 27 - Curvas de amplificação dos genes GAPDH, Dlk1 e UNR	95
Figura 28 – Diluição da amostra em função Ct para o gene Npdc1	97
Figura 29 – <i>Real-Time PCR</i> utilizando diferentes controles internos	98
Figura 30 - Resultado do <i>Real-Time PCR</i> dos genes analisados	104
Figura 31 - Tela de saída do programa para análise das correlações.....	105
Figura 32 - Mapas de correlação dos genes analisados.	107

Figura 33 – Real Time PCR do gene MN1	108
Figura 34 - Padronização da amplificação do gene MN1	109
Figura 35 – Condições selecionadas para amplificação de MN1	110
Figura 36 - Tela de saída do programa BLAT	111
Figura 37 - Esquema da disposição dos <i>primers</i> desenhados para amplificação de MN1	113
Figura 38 - Testes de amplificação de MN1	114
Figura 39 - Resultado da amplificação da porção 3' do gene MN1.....	116
Figura 40 – Células PLJ infectadas com lentivírus	118
Figura 41 – <i>Real-Time PCR</i> para confirmação do RNAi de MN1.....	119
Figura 42 – Curva de Crescimento de células expressando o RNAi de MN1	119
Figura 43 - Crescimento em suspensão de agarose.....	120
Figura 44 - Real Time PCR do genes Dlk1	121
Figura 45 - Amplificação de Dlk1	122
Figura 46 – Análise da seqüência clonada do gene Dlk1	124
Figura 47 – <i>Real-Time PCR</i> da superexpressão de Dlk1.....	126
Figura 48 – Curva de crescimento – Dlk1.....	128
Figura 49 – Primeiro mapa de correlação identificado	137

ÍNDICE DE TABELAS

Tabela 1 – Gênes selecionados nas membranas comerciais	70
Tabela 2 - Genes identificados pelo RDA	75
Tabela 3 – Número de genes selecionados em cada método de normalização ...	83
Tabela 4 - Genes selecionados a partir dos dados obtidos no <i>CodeLink</i>	88
Tabela 5 – Ct do <i>primer</i> de GAPDH em função da sua concentração final	96
Tabela 6 - Tempo de dobramento – Dlk1	128

RESUMO

O Vírus Polioma (Py), induz a formação de múltiplos tumores em camundongos, e causa transformação maligna em cultura, levando à menor dependência de soro e ancoragem para crescimento. Dentre as proteínas codificadas pelo vírus polioma, o antígeno MT (*Middle T*), isoladamente, é capaz de transformar linhagens celulares imortalizadas e gerar tumores quando injetadas em camundongos, sendo que este caráter transformante está relacionado à sua capacidade de se ligar e modular a atividade de diversas proteínas, muitas das quais relacionadas com o controle do ciclo celular. Para melhor compreender os mecanismos moleculares da transformação maligna mediada por MT e do controle do ciclo celular, o objetivo deste trabalho foi identificar e caracterizar novas vias de transdução de sinal e novos genes regulados por MT.

Para a busca de genes regulados pela atividade de MT, comparou-se o perfil da expressão gênica da linhagem que expressa MT (MTWT) com o da linhagem controle (PLJ) através de três metodologias: membranas de nylon comerciais (*Atlas Mouse Cancer 1.2* - Clontech), bibliotecas subtrativas de cDNA construídas através da técnica de RDA (Análise de Diferença Representacional) e lâminas comerciais de microarrays (CodeLink Bioarrays – GE Healthcare). As três metodologias permitiram a identificação de cerca de 1.200 genes-candidatos a serem regulados por MT, dos quais 84 genes foram selecionados para terem sua expressão diferencial analisada por *Real-Time PCR*. Para identificar possíveis vias de transdução de sinal relacionadas à regulação dos genes analisados utilizou-se, como *template*, além do cDNA proveniente das linhagens PLJ e MTWT, cDNAs de linhagens que expressam a proteína MT mutada nas tirosinas 315, 322, 250 e 315/322. Após confirmação por *Real-Time PCR*, dois genes foram selecionados para análise funcional: Dlk1 (*Delta Like-1 homolog*) que, apesar de não ter relação direta com o controle do ciclo celular, está ligado à formação de alguns tumores, e MN1 (Meningioma 1), um gene pouco caracterizado que já foi relacionado à formação de leucemias mielóides e meningiomas.

Para caracterizar o papel de MN1 na linhagem PLJ, foram geradas linhagens PLJ que superexpressam RNAi de MN1. A análise dos parâmetros de crescimento, em substrato sólido e semi-sólido, da linhagem PLJ parental e PLJ expressando o RNAi de MN1 não apresentou alterações estatisticamente significativas. Para o estudo do efeito de Dlk1 na transformação mediada por MT, foram geradas linhagens MTWT que superexpressam Dlk1, o que levou a um aumento estatisticamente significativo na taxa de divisão celular, quando comparada com a linhagem controle (vetor vazio), indicando um possível papel de Dlk1 no sinal desencadeado por MT e no controle do ciclo celular.

Visando identificar novas vias de regulação entre os genes analisados, os níveis de expressão gênica dos demais genes foram avaliados por *Real-Time PCR* e analisados através de correlação de Pearson, aplicada a cada par de genes. Após a análise, foram identificadas 64 possíveis vias de regulação, sendo que 13 destas foram descritas pela primeira vez no presente trabalho e as demais, por já estarem descritas na literatura, permitiram validar o modelo estudado.

Palavras-chave: Poliovírus, antígeno MT, controle do ciclo celular, RDA, CodeLink Bioarrays, vias de sinalização

ABSTRACT

Polyomavirus (Py) induces multiple tumors in mice and malignant transformation in culture, leading to a lower serum and anchorage dependence. Among the proteins coded by the polyomavirus genome, the MT antigen (Middle T) alone is able to transform immortalized cell lines and to generate tumors when injected into animals. This transforming ability is related to its capacity of bind to and modulate several proteins, many of which are related to cell cycle control. Aiming at a better understanding not only of the molecular mechanisms of malignant transformation mediated by MT, but, also, of the processes underlying cell cycle control, the objectives of this work were to identify and characterize novel genes and signal transduction pathways regulated by MT.

For screening of the MT regulated genes, the gene expression profiles of the MT expressing MTWT cell line and of the PLJ control cell line were compared using three methodologies: commercial nylon membranes (Atlas Mouse Cancer 1.2 – Clontech), subtracted cDNA libraries constructed using the RDA methodology (Representational Difference Analysis) and commercial microarrays (CodeLink Bioarrays – GE Healthcare). These three methodologies allowed us to identify about 1,200 candidate genes as being regulated by MT, from which 84 genes were selected to have their differential expression analyzed by *Real-Time PCR*. In an attempt to identify possible signal transduction pathways related to these genes, we used as templates, in addition to cDNAs from the PLJ and MTWT cell lines, cDNAs from cell lines which express MT mutated in tyrosines 315, 322, 250 and 315/322. After confirmation by *Real-Time PCR*, two genes were selected for further functional analysis: Dlk1 (Like-1 Delta homolog), which is related to the differentiation process and to formation of some tumors, and MN1 (Meningioma 1), a newly identified gene, which can induce leukemia and meningioma tumor formation.

To characterize the functional role of MN1, we generated PLJ cell lines in which MN1 was knocked-down by RNAi. The growth characteristics in solid and semi-solid substrates of control PLJ and of

MN1 knocked down PLJ cells did not present any statistically significant alteration. To study the effect of Dlk1 in MT mediated cell transformation, we generated MTWT cell lines overexpressing Dlk1, which induced a significant increase in cell division rate when compared to the empty vector control cell line, indicating a possible role of Dlk1 in the MT induced signals and in cell cycle control.

Aiming at identifying new regulatory pathways among the genes analyzed, the *Real-Time PCR* results were subjected to analysis by Pearson correlation, applied to each pair of genes, allowing us to identify 64 possible regulatory pathways, 13 of which were described, for the first time, in the present work. The other pathways, which had previously been described in the literature, allowed us to validate the model studied.

Key-words: Poliomavirus, MT antigen, cell cycle control, RDA, CodeLink Bioarrays, signaling pathways

ABREVIATURAS

36B4	- Fosfoproteína Ribossômica Ácida P0
Adam12	- <i>ADAM Metallopeptidase Domain 12</i>
ARF	- <i>Alternative Reading Frame of p16^{INK4}</i>
ATP	- Adenina Trifosfato
BLAST	- <i>Basic Local Alignment Search Tool</i>
Cdc2a	- <i>Cell Division Cycle 2 Homolog A</i>
CDK	- Proteína Quinase Dependente de Ciclina
CDKI	- Inibidor de Proteína Quinase Dependente de Ciclina
cDNA	- DNA Complementar
Chk2	- <i>CHK2 Checkpoint Homolog</i>
Csk	- <i>Src C-Terminal Kinase</i>
DAG	- Diacilglicerol
DIK1	- <i>Delta-Like 1 Homolog</i>
DMEM	- <i>Dulbecco's Modified Eagle Médium</i>
DMSO	- Dimetilsulfóxido
DNA	- Ácido Desoxirribonucleico
dNTP	- Desoxinucleotídeo trifosfato
ds	- <i>Double Strand</i>
DTT	- Ditionotrieto
EBV	- Vírus Epstein-Barr
EDTA	- Ácido Etilenodiaminotetracético
EGFP	- <i>Enhanced Green Fluorescence Protein</i>
EST	- <i>Expressed Sequence Tag</i>
Fanca	- <i>Fanconi Anemia, Complementation Group A</i>
FDR	- <i>False Discovery Rate</i>
Fzd2	- <i>Frizzled Homolog 2</i>
GAPDH	- Gliceraldeído 3-Fosfato Desidrogenase
GDP	- Guanosina Difosfato
Grb2	- <i>Growth Factor Receptor-Bound Protein 2</i>
GSK	- Glicogênio Sintase Quinase-3
GTP	- Guanosina Trifosfato
HBV	- Vírus da Hepatite B

HCV	- Vírus da Hepatite C
HIV	- Vírus da Imunodeficiência Humana
HPV	- Papilomavírus Humano
ICTV	- <i>International Committee on Taxonomy of Viruses</i>
Ier5l	- <i>Immediate Early Response 5-like</i>
IGF-1	- <i>Insulin-Like Growth Factor 1</i>
INSIG1	- <i>Insulin Induced Gene 1</i>
IPTG	- Isopropil- β -D-Tio-Galactopiranosídeo
JCV	- JV Vírus
Ksr	- <i>Kinase Suppressor of Ras</i>
Lamb2	- <i>Laminin, Beta 2</i>
LB	- Meio de Luria-Bertani
LIMMA	- <i>Linear Models for Microarray Data</i>
LOWESS	- <i>Locally-Weighted Scatter plot Smoother</i>
LT	- <i>Large T</i>
LTR	- <i>Long Terminal Repeat</i>
Mad2l1	- <i>MAD2 Mitotic Arrest Deficient-like 1</i>
MAPK	- <i>Mitogen-Activated Protein Kinase</i>
MAPKK	- <i>Mitogen-Activated Protein Kinase Kinase</i>
MAPKKK	- <i>Mitogen-Activated Protein Kinase Kinase Kinase</i>
MCM3	- <i>Minichromosome Maintenance Deficient 3</i>
MDM2	- <i>Transformed 3T3 Cell Doublé Minute 2</i>
MLH1	- <i>MutL Homolog 1</i>
MN1	- Meningioma 1
MOI	- <i>Multiplicity of Infection</i>
mRNA	- RNA mensageiro
MT	- <i>Middle T</i>
Nek2	- <i>NIMA-Related Kinase 2</i>
NOTCH2	- <i>Notch Gene Homolog 2</i>
NS5A	- <i>Nn-Structural Protein 5A</i>
Oip5	- <i>Opa Interacting Protein 5</i>
ORF	- <i>Open Reading Frame</i>
PBSA	- <i>Phosphate Buffered Saline</i>
PCR	- Reação de Polimerase em Cadeia

PI-3K	- Fosfatidinositol 3-Quinase
PIP ₂	- Fosfatidinositol 4,5 Bisfosfato
PIP ₃	- Fosfatidinositol 3,4,5 Trifosfato
PKB	- Proteína Quinase B
PKC	- Proteína Quinase C
PLC-γ	- Fosfolipase C-γ
PTB	- Domínio de Ligação a Fosfotirosinas
RAC	- <i>A and C Related Protein Kinase</i>
RACE	- <i>Rapid Amplification of cDNA Ends</i>
Rb	- Retinoblastoma
RDA	- Análise de Diferença Representacional
Reck	- <i>Reversion-Inducing-Cysteine-Rich Protein with Kazal Motifs</i>
RISC	- <i>RNA-Induced Silencing Complex</i>
RNA	- Ácido Ribonucléico
SDS	- <i>Dodecil Sulfato de Sódio</i>
SFB	- Soro Fetal Bovino
SH2	- <i>Src Homolog Domain 2</i>
SMA	- <i>Statistics for Microarray Analysis</i>
Sos	- <i>Son of Sevenless Homolog</i>
ss	- <i>Single Strand</i>
ST	- <i>Small T</i>
SV40	- <i>Simian Vírus 40</i>
TAE	- Tampão Tris-Acetato-EDTA
TFI	- <i>Transcription Finishing Initiative</i>
TM	- <i>Melting Temperature</i>
Tmem158	- <i>Transmembrane Protein 158</i>
TMV	- <i>Tobacco Mosaic Vírus</i>
Uhr1	- <i>Ubiquitin-like Containing PHD and RING Finger Domains1</i>
UNR	- <i>Upstream of NRAS</i>
VP	- <i>Viral Protein</i>
X-gal	- 4-Bromo-5-Cloro-Indolil-Galactopiranosídeo

1. INTRODUÇÃO

1.1. Vírus

1.1.1. Histórico

O primeiro relato sobre vírus data de 1882, quando Adolf E. Mayer supôs que a doença do mosaico do tabaco era causada por um agente contagioso semelhante a uma enzima. Posteriormente, em 1892, o biólogo Dimitrii Ivanovsky observou que o fluído das folhas atacadas pelo mosaico do tabaco mantinham sua capacidade infectante mesmo depois de ser filtrado (Ivanovsky, 1892). Entretanto, o marco histórico na Virologia ocorreu em 1898, quando Martinus W. Beijerinck, ao estudar a causa da doença do mosaico do tabaco, descreveu um agente ao qual designou de *contagium vivum fluidum*, ou vírus, um termo que era utilizado para designar agentes “venenosos” que causavam doenças. Entre as precisas observações de Beijerinck tem-se que o extrato de plantas doentes continuava a ser infeccioso após ser filtrado, que o extrato filtrado não apresentava nenhum crescimento após plaqueado e mantido em estufa por três meses e que o agente necessitava de um tecido vivo para se multiplicar (Beijerinck, 1898).

Após o trabalho de Beijerinck, seguiram diversos trabalhos identificando vírus capazes de infectar os mais variados organismos e associando sua ação com diversas patologias, tais como a febre amarela em humanos (Reed, Carroll *et al.*, 1901) e tumores em animais (Ellerman e Bang, 1908; Rous, 1983). Entretanto, foram necessários quase 40 anos até que as primeiras metodologias e técnicas para caracterizar os vírus fossem desenvolvidas. As mudanças começaram em 1935 quando Wendel M. Stanley obteve sucesso em cristalizar o vírus do mosaico do tabaco (Stanley, 1935). Posteriormente, em 1937, Bawden & Pirie caracterizaram o vírus como um complexo de proteína e ácidos nucléicos e, em 1939, o vírus do mosaico do tabaco (TMV) foi o primeiro vírus a ser observado ao microscópio eletrônico (Creager, Scholthof *et al.*, 1999).

Principalmente após a caracterização da estrutura do DNA (Watson e Crick, 1953), os vírus passaram a ser utilizados como ferramentas para estudar os mais variados mecanismos celulares. O TMV, por exemplo, teve o seu material genético utilizado para demonstrar que a informação genética

poderia ser armazenada no RNA (Fraenkel-Conrat, Singer *et al.*, 1957), o sinal de poliadenilação do mRNA foi descrito utilizando o vírus SV40 (Kaiser, 1957) e o antígeno MT (*Middle T*) do Vírus Polioma foi utilizado para caracterizar a proteína PI-3K (*Phosphatidylinositol 3-kinase*) (Whitman, Downes *et al.*, 1988; Serunian, Auger *et al.*, 1990), que regula importantes processos celulares e está relacionada a diversas patologias, incluindo o câncer (Samuels e Ericson, 2006).

1.1.2. Propriedades das partículas virais

Vírus são elementos genéticos que necessitam de uma célula hospedeira para se multiplicar. Uma vez na célula hospedeira, são capazes de replicar seu material genético independentemente do cromossomo celular. Os vírus também possuem um estado extracelular que permite que se locomovam de um hospedeiro para outro. Dependendo do ambiente e do hospedeiro, a infecção viral pode levar a uma patologia ou a alterações genéticas na célula infectada.

No seu estado extracelular (vírion), a partícula viral consiste de material genético (DNA ou RNA), que pode ser tanto fita dupla (ds) como fita simples (ss). Nos vírus de fita simples, esta pode estar no sentido *sense* (+) ou *antisense* (-), sendo que os vírus que apresentam o simples-fita (-) precisam codificar para a sua própria polimerase, uma vez que as células hospedeiras não possuem a polimerase necessária para produzir mRNA a partir de RNA(-).

O genoma viral é revestido por uma estrutura protéica denominada capsídeo. Esta capa protéica pode ser constituída por um único ou por vários tipos de proteínas, que são distribuídas de uma maneira organizada e repetitiva. De acordo com a distribuição dessas proteínas, elas podem apresentar dois tipos de simetria, helicoidal ou cúbica. Alguns vírus são constituídos por diversas partes que apresentam simetrias distintas, denominada simetria complexa.

Diversos vírus, principalmente os vírus animais, também são revestidos por uma bicamada lipídica envolvida por proteínas, o envelope. A bicamada lipídica é proveniente da célula hospedeira e as proteínas do envelope são codificadas pelo próprio vírus. Assim como a membrana celular, o envelope é impermeável e bloqueia a entrada de drogas ou enzimas na

partícula viral, além de ser o meio pelo qual o vírus interage com a membrana da célula.

Para classificação dos vírus, foi estabelecido em 1966, pelo Comitê Internacional de Taxonomia de Vírus (ICTV), um sistema taxonômico que difere dos sistemas utilizados para os demais organismos. Neste sistema, um conjunto de propriedades que descreve um determinado vírus é comparado com o de outros vírus, sendo que não existe uma hierarquia nestas propriedades e não há um conjunto fixo de propriedades a ser utilizado. Geralmente são empregadas características relacionadas à morfologia viral e, entre essas, são comumente utilizadas: natureza do ácido nucléico (DNA ou RNA), simetria do capsídeo, presença ou ausência de envelope e dimensão do vírion e do capsídeo (White e Fenner, 1994; Flint, Enquist *et al.*, 2000; Fields, Knipe *et al.*, 2001). Algumas dessas características estão esquematizadas na Figura 1.

Figura 1 - Esquema da classificação dos principais vírus animais

Classificação dos principais vírus que infectam animais vertebrados e um vírus que infecta insetos (Baculovírus) quanto a: tipo de ácido nucléico - DNA ou RNA; simetria do capsídeo - icosaédrica, cúbica ou complexa; presença ou não de envólpe e arquitetura do genoma - dupla fita (ds)/simples fita (ss), com genoma segmentado (seg.) ou contínuo (cont.) e fita no sentido do mRNA (+) ou no sentido complementar (-). *Imagem adaptada de Flint et al., 2000.*

1.1.3. Ciclo de vida viral

O ciclo de vida dos vírus pode ser dividido em sete estágios: ligação na célula hospedeira (adsorção), penetração do material genético viral na célula (injeção), síntese precoce de proteínas, replicação do ácido nucléico, síntese tardia de proteínas, reunião dos componentes virais e liberação das partículas virais.

Na Adsorção, receptores presentes na superfície das células são reconhecidos por partículas virais através de proteínas de seu envelope ou capsídeo. Na etapa seguinte (injeção), os vírus se valem de diferentes estratégias para liberar o seu material genético no interior da célula. Nos vírus com envelope, alguns mecanismos são relativamente simples como a fusão do envelope com a membrana plasmática, permitindo que o ácido nucléico seja liberado no citoplasma celular. Contudo, outros vírus envelopados utilizam estratégias mais complexas, como a endocitose do vírus (invaginação da membrana plasmática colocando o vírus numa vesícula endocítica) e posterior liberação do material genético, ou endocitose seguida da lise da vesícula endocítica. Nos vírus sem envelope, o mecanismo de entrada ainda não foi completamente elucidado, havendo vírus que realizam a entrada por endocitose, que pode ser passiva ou mediada por receptores, e outra classe de vírus que induzem a formação de poros na membrana para a sua entrada (Smith e Helenius, 2004; Marsh e Helenius, 2006).

Uma vez que o material genético se encontre no interior da célula, novas cópias do genoma viral precisam ser replicadas e proteínas específicas do vírus precisam ser sintetizadas para que ocorra multiplicação. Para a síntese de proteínas, independente do material genético do vírus, DNA ou RNA, o primeiro passo deste processo na maioria das vezes consiste na síntese do RNA mensageiro (mRNA).

A forma como o vírus sintetiza o mRNA varia de acordo com a estrutura do seu genoma. Nos vírus de DNA simples fita, primeiramente é feita a síntese da fita de DNA complementar e posteriormente o mRNA, como ocorre com o parvovírus (Deleu, Pujol *et al.*, 1999; Bashir, Horlein *et al.*, 2000). Nos vírus de RNA com fita negativa (-) ou dupla fita é realizada a transcrição da fita negativa (-) para gerar o mRNA (+) e, como normalmente a célula não possui uma RNA polimerase capaz de realizar essa transcrição, o vírus carrega no próprio vírion uma RNA polimerase dependente de RNA. Como exemplo pode-se citar o vírus da raiva (Jayakar, Jeetendra *et al.*, 2004) e o reovírus (Roner e Joklik, 2001), que são ssRNA (-) e dsRNA, respectivamente. Nos vírus ssRNA(+), alguns vírus, como o da poliomielite, utilizam o próprio material genético como mRNA (Franco, Pathak *et al.*, 2005) e outros, como o HIV, utilizam sua transcriptase reversa para gerar um

dsDNA intermediário que será utilizado na síntese do mRNA (Freed, 2001; Harrich e Hooker, 2002) (Figura 2).

Figura 2 – Formação do mRNA a partir do material genético viral

Mecanismos já descritos de formação de mRNA a partir do material genético viral. O sentido dos RNAs está indicado entre parênteses. Por convenção, o mRNA é considerado como positivo (+). Assim, os ácidos nucleicos cuja informação está no mesmo sentido do mRNA estão indicados por (+) e aqueles que estão no sentido complementar por (-).

Uma vez que o mRNA foi sintetizado, é iniciada a tradução das proteínas virais as quais, na maior parte dos vírus, podem ser divididas em dois grandes grupos, o das proteínas necessárias para a replicação do vírus e que são sintetizadas logo após a infecção, denominadas proteínas precoces, e o das proteínas que codificam para o capsídeo viral e que são sintetizadas posteriormente, denominadas proteínas tardias. Como as proteínas precoces são enzimas, são produzidas em pequenas quantidades enquanto que as proteínas tardias, por possuírem função estrutural, são produzidas em altas quantidades.

Para replicar seu material genético, os vírus de RNA utilizam duas abordagens possíveis, ou sintetizam o RNA a partir de outro RNA ou geram um DNA intermediário que funcionará como *template* para a síntese do RNA genômico. Para as duas sínteses, geralmente é necessário a presença de uma enzima codificada pelo vírus, uma vez que as células hospedeiras raramente possuem RNA ou DNA polimerases RNA-dependente. Os vírus de DNA utilizam a maquinaria de replicação da célula hospedeira, sendo que alguns destes vírus são capazes de induzir a replicação viral, o que confere o caráter transformante de muitos vírus de DNA.

Após a replicação viral, o principal mecanismo para a reunião dos componentes virais é através dos genomas virais, que contêm sinais de empacotamento (*packaging signal*), podendo ser pequenas seqüências ou estruturas secundárias do material genético viral, que direcionam a formação do capsídeo ao redor do genoma (Frolova, Frolov *et al.*, 1997; Clever, Mirandar *et al.*, 2002; Gordon-Shaag, Ben-Nun-Shaul *et al.*, 2002; Hibbert, Mirro *et al.*, 2004).

Na etapa final da infecção, a liberação do vírus maduro, cada classe de vírus apresenta uma estratégia. Nos vírus sem envelope, ocorre a destruição da célula para liberação dos vírus. Nos vírus com envelope, a saída do vírus é realizada por brotamento. Quando o processo de brotamento é acentuado ocorre a lise e conseqüente morte da célula; entretanto, quando o processo é mais brando, ocorre uma infecção persistente, na qual a célula continua a produzir o vírus por um longo período.

Os ciclos de vida virais que culminam com a lise celular são denominados ciclo lítico. Há casos onde os eventos relacionados ao ciclo lítico não se completam e o material genético do vírus se integra ao genoma da célula hospedeira, onde fica em uma forma latente, sendo que os sintomas da infecção somente aparecem quando o vírus abandona o estado latente (White e Fenner, 1994; Flint, Enquist *et al.*, 2000; Fields, Knipe *et al.*, 2001; Lodish, 2004). Entretanto, essa infecção latente pode levar à transformação celular e posteriormente à formação de um tumor. O vírus papiloma (HPV), por exemplo, causa verrugas genitais em infecções líticas e pode gerar um câncer cervical em infecções latentes (Doorbar, 2005).

1.2. Vírus e Câncer

O processo de transformação de uma célula de normal em tumoral é o resultado de uma série de alterações que levam a um desbalanço entre os processos de divisão, diferenciação e morte celular, culminando com uma divisão celular descontrolada. Essas alterações podem ser devidas a mutações, alterações epigenéticas, agentes internos/externos ou infecção viral. Nos próximos itens serão discutidas as principais bases moleculares do controle do ciclo celular e a relação que existe entre alguns vírus e o processo tumoral.

1.2.1. Controle da proliferação celular

O ciclo celular consiste de uma série de eventos que preparam a célula para o seu crescimento e divisão. Em uma célula eucariótica o ciclo celular possui duas fases principais: a fase S (síntese), única fase onde ocorre a duplicação do DNA, e a fase M (mitose), onde os cromossomos duplicados se separam e são geradas duas novas células a partir da célula mãe. As fases M e S são separadas por dois intervalos, G1 (*Gap 1*) antes da fase S e G2 (*Gap 2*) após a fase S. No início da fase G1, a célula pode entrar em uma fase quiescente denominada G0, que pode durar desde algumas horas até toda a vida da célula (Figura 3) (Tessema, Lehmann *et al.*, 2004).

O principal mecanismo de controle do ciclo celular ocorre através de uma família de proteínas serina/treonina quinases cuja atividade regula vias de transdução de sinal relacionadas à proliferação celular. Estas quinases são heterodímeros com uma subunidade regulatória, a ciclina, e uma subunidade catalítica, a proteína quinase dependente de ciclina (CDK) (Murray, 2004). Em humanos, já foram identificadas 11 tipos ciclinas e 9 tipos de CDKs, que se combinam de acordo com os substratos que serão fosforilados (Meijer, 2000).

Como diferentes vias de transdução de sinal devem ter sua atividade regulada em diversos pontos do ciclo celular, diferentes complexos ciclinas-CDKs também se encontram ativados ao longo do ciclo e, para que ocorra essa oscilação da atividade, são utilizados quatro mecanismos: fosforilação/desfosforilação da CDK em resíduos de serina e treonina, degradação das ciclinas, síntese de ciclinas/CDKs e utilização de inibidores

específicos para as CDKs (CKIs). Os principais complexos ciclina-CDK bem como as fases do ciclo celular em que eles atuam estão representados na Figura 3 (Murray, 2004).

Figura 3 – Ciclo celular

Representação das fases do celular e dos principais complexos ciclina-CDK envolvidos no controle do ciclo celular, *adaptado de Vermeulen, 2003*.

Como exemplo de regulação por um complexo ciclina CDK tem-se a proteína retinoblastoma (pRb) que quando se encontra fosforilada devido, por exemplo, à ação do complexo ciclina E-CDK2, libera o fator de transcrição E2F, que ativa a transição da fase G1 para S (Figura 3) (Hinds, Mitnacht *et al.*, 1992). Quando ocorre algum dano no DNA da célula, para que esse dano não seja propagado para a célula filha ocorre a ativação de p53, a qual ativa a CKI denominada p21^{Cip1/Waf1} que, ao se ligar e inativar diversos complexos, como o complexo ciclina E/CDK2, impede a fosforilação de pRb e, consequentemente, impede a liberação de E2F, bloqueando a transição G1 → S até que o dano no DNA seja reparado. Se o dano não for reparado,

p53 ativa uma via de morte celular programada denominada apoptose (Harper, Adami *et al.*, 1993; El-Deiry, Harper *et al.*, 1994; Jackson, Chevalier *et al.*, 1995).

Quando ocorre uma falha nos mecanismos que regulam a progressão do ciclo celular e a morte celular, ocorre a transformação celular e, posteriormente, a formação de um tumor. Devido a esse descontrole de diversos mecanismos, a célula tumoral passa a apresentar diferenças fenotípicas em relação às células normais como aumento da densidade de saturação, menor necessidade de fatores de crescimento e alteração na morfologia (Cooper, 1995).

1.2.2. Oncogenes e Genes Supressores de Tumor

As mutações que ocorrem nos genes relacionados à divisão celular podem resultar em uma proliferação descontrolada. Essas mutações ocorrem principalmente em duas classes de genes, os proto-oncogenes e os genes supressor de tumor.

Em células normais, os produtos dos proto-oncogenes agem em diferentes níveis ao longo do ciclo celular como proliferação, vias apoptóticas ou vias de reparo de DNA; entretanto, as versões alteradas dos proto-oncogenes, conhecidas como oncogenes, podem promover o crescimento de um tumor, sendo que a mutação de somente um dos alelos do proto-oncogene pode ser suficiente para gerar um oncogene. Os proto-oncogenes podem passar a ter sua atividade aumentada por superexpressão, ativação constitutiva ou mutação (Weinberg, 1994).

Os oncogenes foram inicialmente identificados como genes presentes em retrovírus que causavam transformação na célula hospedeira. Posteriormente, foi descrito que esses oncogenes virais foram obtidos por recombinação com genes da célula hospedeira e que, portanto, possuíam um análogo celular. O primeiro oncogene descrito, gene *src*, codifica uma proteína tirosina quinase responsável pela ativação de diversas vias de transdução de sinal que culminam, entre outros, com a progressão do ciclo celular (Stehelin, Varmus *et al.*, 1976).

Ao contrário dos oncogenes onde o ganho de função leva a um aumento da proliferação celular, este efeito pode ser causado por um mecanismo oposto, ou seja, pela perda da função de um determinado gene.

Esses genes cujos produtos são necessários para o funcionamento de uma célula normal e cuja perda de função leva à formação de tumores são classificados como genes supressores de tumor. Os genes supressores de tumor geralmente codificam para proteínas que inibem a progressão em estágios específicos do ciclo celular, proteínas que levam à parada do ciclo celular, caso ocorra algum dano no DNA, e proteínas que promovem a apoptose (Macleod, 2000).

Diferentemente dos proto-oncogenes, onde a mutação em um alelo é suficiente para estimular a proliferação, os dois alelos de um gene supressor de tumor devem ser perdidos ou inativados para que um tumor possa ser iniciado. Na maior parte dos cânceres, há um gene supressor de tumor deletado ou mutado, impedindo a produção de uma proteína ou levando à produção de uma proteína não funcional. Existem também relatos de que a inativação dos genes supressores de tumor pode ser devida à alterações epigenéticas. Os dois membros dessa classe que já foram melhor caracterizados são os genes que codificam para as proteínas pRb e p53. A proteína Rb foi inicialmente encontrada como estando mutada em retinoblastomas e depois foi caracterizada como sendo um importante regulador da progressão do ciclo celular. A proteína p53 consiste no mais importante regulador do ciclo celular, com funções que vão desde a parada do ciclo celular até indução da apoptose, sendo estimado que pelo menos 50% dos tumores humanos possuem o gene p53 mutado (Harris, 1996).

Os produtos dos proto-oncogenes e dos genes supressores de tumor atuam de forma coordenada no controle da proliferação de células normais e agem em conjunto na evolução de um processo neoplásico. Assim, um único gene alterado não é suficiente para a tumorigênese, sendo necessária a ocorrência de uma combinação de alterações que afetam a função de genes pertencentes às duas classes (Hunter, 1991). Em média, entre vias de transdução de sinal reguladas por oncogenes e genes supressor de tumor, é necessária a alteração em pelo menos seis vias para que ocorra a transformação maligna de células humanas (Rangarajan, Hong *et al.*, 2004).

1.2.3. Transformação celular por vírus tumorais

O primeiro relato que se tem de um vírus tumoral data de 1911 quando Rous isolou o vírus do sarcoma de Rous (RSV) como o agente transmissível responsável por sarcomas em galinhas. Em humanos, o primeiro vírus tumoral descrito foi o Epstein-Barr (EBV), em 1964, através de observações em microscópio eletrônico de células derivadas do linfoma de Burkitt (Epstein, Achong *et al.*, 1964), e foram necessários mais 15 anos para que diversos vírus começassem a ter sua atividade relacionada ao processo tumoral. De acordo com um estudo epidemiológico realizado em 2002, cerca 12,1% dos tumores humanos estão relacionados a infecções virais (Parkin, 2006).

Nas últimas décadas, os vírus tumorais foram vitais na descoberta dos mecanismos moleculares que controlam o ciclo celular, na caracterização de diversos conceitos da Biologia Celular e no reconhecimento da etiologia de diversos cânceres humanos. Eles revelaram as bases da genética do câncer e permitiram a construção de uma base de dados aplicada não só aos cânceres induzidos por vírus, como também a todas neoplasias. Os oncogenes *ras*, *myc*, *fos*, *abl*, e *src* e os genes tumor supressores *p53*, *erbB2* e *Wnt-APC* são alguns exemplos dos genes e vias que foram descobertos através dos vírus tumorais. Nos próximos itens serão descritos os principais mecanismos pelos quais os vírus tumorais levam a transformação celular.

1.2.3.1. Retrovírus carregando oncogenes celulares

Logo após a infecção de um retrovírus, a transcriptase reversa viral gera um DNA dupla fita que é integrado ao cromossomo celular. A cópia integrada, denominada provírus, é similar a um gene celular, com exceção de que é controlada pelas LTR (*Long Terminal Repeat*), que são seqüências virais capazes de funcionar como um promotor forte (Yamamoto, De Crombrughe *et al.*, 1980). Quando ocorre a integração de um provírus a montante de um proto-oncogene, a transcrição do provírus pode produzir uma quimera entre o genoma viral e o transcrito celular, levando à incorporação do gene celular no genoma viral. Estima-se que cerca de 15% dos transcritos virais possuem fragmentos do genoma celular incorporado (Coffin, Hughes *et al.*, 1997).

A captura de um oncogene funcional consiste num evento raro, e como parte do genoma viral tende a ser perdido durante a recombinação, o vírus resultante geralmente não é capaz replicar sem que tenha a ajuda da maquinaria de um vírus auxiliar. Se ocorrer a incorporação do proto-oncogene no genoma viral, o oncogene não responderá mais aos controles da célula e, devido à ação da LTR, ocorrerá uma superexpressão constitutiva do gene. Como resultado, retrovírus com oncogenes incorporados são capazes de transformar células em cultura e induzir a formação de tumores em animais após um curto período de latência.

1.2.3.2. Retrovírus sem oncogenes virais

Muitos retrovírus, mesmo sem possuírem oncogenes incorporados, podem induzir a formação de tumores *in vivo* através de uma mutagênese insercional do provírus, que ocorre quando o provírus se integra próximo a um proto-oncogene e a LTR viral ativa o proto-oncogene (Kung, Boerkoel *et al.*, 1991; Uren, Kool *et al.*, 2005). Como essa integração é um evento ao acaso, é necessária uma replicação contínua do vírus a fim de aumentar as chances de que ocorra a integração próxima a um proto-oncogene.

Quando ocorrer a integração próxima ao proto-oncogene e a LTR ativar a sua expressão, esta simples ativação geralmente não é suficiente para induzir a transformação, devendo ocorrer outros eventos igualmente raros para que a célula perca o controle sobre o ciclo celular. Assim, esses tumores

surgem após um longo período de latência e são monoclonais, devido à série de eventos necessários para levar à transformação celular.

O processo de infecção mediado por retrovírus também pode levar à formação de um tumor por criar um ambiente propício para a infecção de um segundo vírus tumoral. Como exemplo, tem-se o vírus da imunodeficiência humana (HIV) que, apesar de não possuir um oncogene e não ter sido comprovada sua capacidade de ativar um proto-oncogene, gera um quadro de imunossupressão severa que facilita a indução de um tumor por outras infecções, principalmente pelo herpesvírus humano tipo 8 (HHV-8), vírus Epstein-Barr (EBV) e papilomavírus humano (HPV). Portanto, o HIV contribui indiretamente para a carcinogênese humana (Bonnet e Morlat, 2006).

1.2.3.3. Vírus tumorais de DNA

Diferente dos oncogenes retrovirais, que tiveram sua origem nas células hospedeiras e que não exercem papel na replicação viral, os oncogenes dos vírus tumorais de DNA não possuem homólogos celulares, o que sugere uma origem viral, e são fundamentais tanto para a replicação viral como para a transformação celular (Talbot e Crawford, 2004).

A fim de proporcionar um ambiente propício à replicação do DNA viral, as proteínas não estruturais codificadas pelos vírus de DNA estimulam células quiescentes a entrarem na fase S. Entretanto, por possuírem um genoma pequeno, estes vírus ativam o ciclo celular e previnem uma apoptose prematura utilizando um pequeno conjunto de proteínas ou uma única proteína multifuncional.

Os vírus tumorais de DNA foram fundamentais na descoberta dos genes supressores de tumor, uma vez Rb e p53 são alvos comuns das proteínas codificadas por eles. Para bloquear a ação da proteína p53, por exemplo, o antígeno LT do vírus SV40 inibe a ligação de p53 ao DNA, E6 do papilomavírus (HPV) leva à degradação de p53 por ubiquitinação e a proteína E1B de adenovírus inibe a capacidade de transativação de p53. Para inibir a interação de Rb com E2F, LT de SV40, E7 de HPV e E1A do adenovírus se ligam a Rb, liberando o fator de transcrição E2F. Outras proteínas também são alvos dos vírus tumorais de DNA como, por exemplo, a proteína fosfatase 2A (PP2A), um supressor de tumor, que é regulada pela proteína ST de SV40

(Yang, Lickteig *et al.*, 1991) e ativada por NS5A de HCV e E4 de adenovírus (Georgopoulou, Tsitoura *et al.*, 2006).

Os vírus tumorais de DNA são responsáveis por cerca de 65% dos tumores humanos com origem infecciosa, que corresponde a quase 12% dos tumores ao redor do mundo, sendo que os principais vírus são o EBV, responsável por 10% dos carcinomas gástricos, o HBV/HCV responsáveis por 85% dos tumores de fígado e o HPV responsável por 100% dos cânceres de cérvix (Parkin, 2006). Portanto, o estudo dos vírus tumorais de DNA é importante não só para a busca de alvos terapêuticos como, também, para uma melhor compreensão das bases moleculares do ciclo celular e da transformação maligna, uma vez que seus alvos celulares são, em sua grande maioria, os mesmos que se encontram alterados em uma célula tumoral onde não houve infecção.

1.3. Vírus Polioma (Py)

O Vírus Polioma (Py) pertencente à família *Papovaviridae* e subfamília *Polyomavirinae*, que foi primeiramente descrita por Ludwig Gross em 1953 ao descrever que o filtrado de células provenientes de leucemias espontâneas em camundongos continha um contaminante que causava tumores quando re-inoculado em camundongos (Gross, 1953). Posteriormente, foram estabelecidas técnicas para produzir este agente em cultura e, como ele era capaz de causar diversos tumores quando inoculado no animal, recebeu o nome de polioma (Eddy, Stewart *et al.*, 1958).

Os Poliomavírus são vírus não envelopados, com capsídeo de simetria icosaédrica e material genético constituído por um DNA dupla fita circular. O genoma dos poliomavírus possui duas regiões: a região precoce, que codifica para proteínas não estruturais ou enzimas que direcionam a síntese do ácido nucléico, e são sintetizadas antes da replicação viral, e a região tardia, que codifica para as VPs (*Viral Protein*), que são proteínas do capsídeo viral e são sintetizadas após a replicação viral. São capazes de infectar uma gama variada de animais, sendo que os mais estudados são os vírus SV40 que infecta macacos, JCV que infecta humanos e Py que infecta camundongos.

O SV40, por ter sido um dos principais vírus utilizado para caracterizar o ciclo celular e o processo neoplásico, é o poliomavírus mais bem estudado. Entre os genes que devem sua descoberta ao SV40 está o gene supressor de

tumor p53, que foi encontrado por se ligar fortemente ao antígeno LT (*Large T*) de SV40 (Chang, Simmons *et al.*, 1979; Kress, May *et al.*, 1979; Lane e Crawford, 1979). Apesar do hospedeiro natural de SV40 ser representado por algumas espécies de macacos asiáticos, há indícios de que possa estar ligado à neoplasias humanas (Poulin e Decaprio, 2006).

O JC vírus (JCV) infecta cerca de 80% da população humana mundial, sendo responsável pela leucoencefalopatia multifocal (Berger, Levy *et al.*, 1998), uma desmielinização do sistema nervoso central que ocorre principalmente em pacientes imuno-deprimidos, como aqueles infectados por HIV (Berger, Pall *et al.*, 1998). A capacidade do JCV de causar tumores em modelos animais, transformar células em cultura e o fato de ter sido encontrado em diversos tumores humanos do sistema nervoso central, fornece indícios de que JCV possa estar relacionado à alguns tumores humanos, entretanto esta relação ainda não foi comprovada (White, Gordon *et al.*, 2005; Shah, 2007).

O Vírus Polioma (Py) foi um dos primeiros vírus que teve o seu potencial tumorigênico estudado e o primeiro poliomavírus descoberto (Gross, 1953). O Vírus Polioma é altamente oncogênico no seu hospedeiro natural, levando ao crescimento de diversos tumores sólidos após um curto período de inoculação do vírus em camundongos recém-nascidos (Gross, 1983), como pode ser observado na Figura 4.

Figura 4 - Camundongo infectado com Polioma Vírus

Imagem da secção de um camundongo de 10 dias de idade, infectado com o Vírus Polioma após o nascimento, hibridizado com DNA viral marcado com S³⁵ (Benjamin, 2001). As setas indicam os principais focos de infecção.

Além de levar à formação de tumores em camundongos, células em cultura infectadas por este vírus apresentam menor dependência de soro para crescimento, aumento da densidade de saturação, capacidade de formar colônias em meio semi-sólido e capacidade de induzir a formação de tumores quando inoculadas em animal. Assim, sua capacidade de reproduzir o processo de gênese tumoral *in vitro* permite a utilização deste sistema como modelo da transformação de fibroblastos em cultura (Oliveira, Brochado *et al.*, 1999).

Três fatores atuam em conjunto para permitir que o Vírus Polioma tenha uma alta capacidade transformante. Primeiramente, o vírus é capaz de infectar mais de 40 tipos celulares distintos através da sua ligação a ácidos siálicos, presentes no sacarídeo terminal de diversas glicoproteínas da membrana celular (Dawe, Freund *et al.*, 1987; Chen e Benjamin, 1997). Uma vez na célula, o vírus possui uma região reforçadora (*enhancer*) multivalente no qual podem se ligar mais de 30 fatores de transcrição diferentes, o que permite a tradução das proteínas do vírus em diversos tipos celulares (Campbell e Villarreal, 1988; Rochford, Campbell *et al.*, 1990; Amalfitano, Martin *et al.*, 1992; Rochford, Moreno *et al.*, 1992). Por fim, a região precoce do seu genoma codifica três proteínas que, além de auxiliarem na replicação viral, ativam diversas vias de sinalização relacionadas à indução do ciclo celular, os chamados antígenos tumorais (antígenos T), sendo eles: *Large T* (LT), *Middle T* (MT) e *Small T* (ST) (Figura 5) (Treisman, Cowie *et al.*, 1981; Rassoulzadegan, Cowie *et al.*, 1982) .

Os três antígenos tumorais são gerados a partir do processamento diferencial do mesmo mRNA. Assim, a sequência codificadora da região N-terminal destas proteínas é idêntica, enquanto que a região C-terminal difere em cada uma, já que o processamento diferencial do mRNA gera diferentes fases de leitura. Um processo semelhante ocorre com as três proteínas estruturais (Liu e Carmichael, 1993).

Figura 5 - Mapa representativo do genoma do Vírus Polioma

Os transcritos codificados pelas regiões precoce (LT, MT e ST) e tardia (VP2, VP2 e VP3) estão representados pelos círculos externos. As linhas pontilhadas indicam os íntrons e o sentido dos éxons (⇒) a extremidade 3' do transcrito. *Adaptado de Ichaso and Dilworth, 2001a.*

O papel de cada antígeno T durante a transformação celular *in vitro* foi analisada através de mutantes virais (Schlegel e Benjamin, 1978). Estes estudos mostraram que MT e LT são necessários para a manutenção do fenótipo transformado, enquanto que vírus mutantes que expressam apenas ST não apresentam fenótipo transformante. Através do uso de vetores de expressão para mamíferos, contendo o cDNA de cada um dos três antígenos T, foi possível demonstrar que a expressão combinada de LT e MT é necessária e suficiente para transformar culturas primárias (Rassoulzadegan, Cowie *et al.*, 1982), e que MT sozinho é capaz de transformar linhagens celulares imortalizadas e gerar tumores quando injetadas em camundongos (Treisman, Novak *et al.*, 1981).

1.3.1. Antígeno LT

O antígeno LT é uma proteína de 98KDa, que se localiza no núcleo das células infectadas e é capaz de se ligar a seqüências específicas presentes na origem de replicação do DNA viral, atuando como um importante regulador negativo da transcrição precoce (Farmerie e Folk, 1984) e um regulador positivo da transcrição tardia (Kern, Pellegrini *et al.*, 1986).

Através da sua porção N-terminal, LT é capaz de se ligar à proteína Rb (Dilworth, 1990), o que impede o seqüestro de E2F e induz a progressão do ciclo celular. Para contribuir com a atividade de E2F, LT também se liga a co-ativadores com atividade histona-acetiltransferase, promovendo alterações epigenéticas necessárias para a atividade de E2F (Nemethova, Smutny *et al.*, 2004). O antígeno LT também ativa genes relacionados à progressão do ciclo celular como Ciclina A (Schuchner, Nemethova *et al.*, 2001) e timidina quinase (Ogris, Rotheneder *et al.*, 1993) e, ao contrário do que ocorre com a maior parte dos poliomavírus, LT do Vírus Polioma não é capaz de inibir o gene p53, papel este desempenhado pelo antígeno ST.

Devido principalmente à sua capacidade de se ligar a pRb, LT é capaz de imortalizar fibroblastos primários em cultura. Vírus que possuem LT mutado, porém MT íntegro, apesar de não imortalizarem culturas primárias (Freund, Bronson *et al.*, 1992), são capazes de transformar linhagens previamente estabelecidas, indicando que apesar de LT ser necessário para o processo de transformação celular induzida pelo Vírus Polioma, é a ação de MT que consiste no passo crítico para essa transformação.

1.3.2. Antígeno ST

A proteína ST possui 22KDa, localiza-se tanto no núcleo quando no citoplasma de células infectadas pelo vírus e, assim como MT, é capaz de se ligar à proteína serina/treonina fosfatase 2A (PP2A). Em geral, ST não é necessário nem suficiente para a transformação celular; entretanto, ST do Vírus Polioma substitui uma das subunidades de PP2A, o que leva a célula a um aumento da densidade de saturação, reorganização do citoesqueleto e desregulação das *tight junctions*, que têm papel na adesão celular e na metástase tumoral (Liang, Carmichael *et al.*, 1984; Pallas, Shahrik *et al.*, 1990; Nunbhakdi-Craig, Craig *et al.*, 2003; Janssens, Goris *et al.*, 2005).

Uma das regulações de p53 é através da proteína MDM2 (*Transformed 3T3 Cell Double Minute 2*), que se liga a p53 inativando-o ou induzindo sua degradação. Um mecanismo de controle dessa inativação de p53 é através da proteína ARF (*Alternative Reading Frame of p16^{INK4}*) que seqüestra MDM2. Apesar do mecanismo ainda não estar esclarecido, foi demonstrado que células expressando ST apresentam um alto nível de ARF e um baixo nível de p53, dando indícios de que o complexo ST/PP2A inibe a ativação de p53 através de ARF. Estes dados sugerem que ao invés de inibir PP2A, como era esperado, ST confere uma nova função à proteína PP2A (Lomax e Fried, 2001; Moule, Collins *et al.*, 2004).

1.3.3. Antígeno MT

No ciclo lítico natural do Vírus Polioma, o antígeno MT promove um ambiente favorável para a replicação viral e parece estar relacionado com a montagem do capsídeo, uma vez que induz a fosforilação da proteína VP1 em resíduos de treonina (Li, Lyon *et al.*, 1995). Entretanto, em condições experimentais, devido à sua capacidade de transformar eficientemente linhagens celulares estabelecidas, MT tem sido estudado em nosso laboratório não só para se compreender seus mecanismos de ação, mas também como uma poderosa ferramenta utilizada para caracterizar as bases moleculares dos mecanismos de controle da proliferação celular e da transformação maligna (Rameh e Armelin, 1991; Oliveira, Roberts *et al.*, 1998; Winnischofer, De Oliveira *et al.*, 2003).

O antígeno MT é uma fosfoproteína de 421aa codificada através de *splicing* alternativo da região precoce do Vírus Polioma. A proteína possui uma massa molecular de 55KDa e, através de um região de 22aa hidrofóbicos próximo da sua porção C-terminal, se ancora na parte interna da membrana plasmática ou em compartimentos perinucleares, sendo que a deleção desta região leva à perda da capacidade transformante de MT (Novak e Griffin, 1981). A localização de MT na região perinuclear parece estar associada ao rearranjo do citoesqueleto de actina e à condensação dos endossomos tubulares, enquanto que sua ligação à membrana celular está ligada à capacidade transformante de MT (Novak e Griffin, 1981; Carmichael, Schaffhausen *et al.*, 1982; Zhu, Eicher *et al.*, 1998).

MT possui múltiplos sítios de fosforilação como as tirosinas (Tyr) dos aminoácidos 250, 315 e 322 e das serinas (Ser) 257 e 283. Entretanto, esta atividade quinásica não é intrínseca de MT, uma vez que MT sintetizado por tradução *in vitro*, e expresso por bactérias ou células de insetos, não apresenta atividade de quinase. Além disso, MT não apresenta os domínios de ligação a ATP ou a sequência comum aos domínios quinásicos (Hanks, Quinn *et al.*, 1988).

Através da análise de imunoprecipitados de MT por eletroforese mono ou bi-dimensional (Courtneidge e Smith, 1983; Pallas, Cherington *et al.*, 1988), juntamente com a geração de inúmeros mutantes (Morgan, Kaplan *et al.*, 1988; Druker e Roberts, 1991), foi demonstrado que o caráter transformante de MT está relacionado com sua capacidade de se ligar e modular a atividade de diversas proteínas celulares através, principalmente, de interações entre as tirosinas fosforiladas e domínios SH2 (Src homolog domain 2). A fosforilação das tirosinas 250 (Tyr-250), 315 (Tyr-315) e 322 (Tyr-322) por Src, por exemplo, criam sítios que regulam a atividade de Shc, PI-3K e PLC- γ , respectivamente (Figura 6) (Gottlieb e Villarreal, 2001; Dilworth, 2002). A seguir, serão descritas as principais proteínas que têm sua atividade modulada por MT, bem como sua relação com o fenótipo transformante de MT.

Figura 6 - Principais vias de transdução de sinal mediada por MT

Uma vez ligado à membrana, devido a atividade de Src, a oncoproteína MT é fosforilada em sítios de tirosina, sendo que as tirosinas fosforiladas se ligam e ativam proteínas relacionadas com a progressão do ciclo celular.

1.3.3.1. Proteína fosfatase 2A (PP2A)

No metabolismo das células humanas, existem cerca de 530 proteínas quinases cujas atividades regulam os mais variados processos celulares. Até o momento, são conhecidas poucas fosfatases que antagonizam este efeito, aproximadamente 180 fosfatases (Arena, Benvenuti *et al.*, 2005). Entre estas fosfatases, a atividade da proteína serina/tirosina fosfatase 2A (PP2A) controla a atividade de pelo menos 50 quinases diferentes, que estão relacionadas a mecanismos como transdução de sinal, progressão do ciclo celular, transcrição, tradução e replicação do DNA, fazendo com que PP2A seja considerada um dos principais genes supressores de tumor (Janssens, Goris *et al.*, 2005).

A proteína PP2A é formada por uma subunidade estrutural A (65KDa), uma catalítica C (36Kda) e, facultativamente, uma regulatória B (55KDa,

61KDa ou 72KDa). Cada subunidade possui pelo menos duas isoformas que podem se combinar das mais variadas formas (Van Hoof e Goris, 2004).

Uma vez que ST e a porção N-terminal de MT apresentam uma alta homologia, MT também é capaz de formar complexos estáveis com as subunidades A e C de PP2A (Pallas, Shahrik *et al.*, 1990; Walter, Ruediger *et al.*, 1990). Embora ainda não se conheça o mecanismo envolvido, a ligação de PP2A a MT é necessária para que ocorra uma associação estável de MT com a família Src de proteínas quinases, sendo que na ausência da ligação entre PP2A e MT, não ocorre a fosforilação das tirosinas mediada por Src e, consequentemente, não são ativadas as vias de sinalização para o crescimento (Ichaso e Dilworth, 2001). Entretanto, mutantes da subunidade C de PP2A, cataliticamente inativos, são capazes de formar complexos com o antígeno MT e ainda ativar Src, demonstrando que a atividade de PP2A não é essencial para a formação do complexo MT funcional (Ogris, Mudrak *et al.*, 1999). Como a ausência de PP2A influencia na capacidade de MT se ligar à membrana (Brewster, Glover *et al.*, 1997), uma possível explicação é que a interação com PP2A leve a alterações conformacionais em MT que possibilitam sua ligação a Src.

Assim como o complexo ST/PP2A, o complexo MT/PP2A também é capaz de ativar vias de MAPK por um mecanismo independente da ligação de MT a Shc (descrito posteriormente no item 1.3.3.5) (Rodriguez-Viciano, Collins *et al.*, 2006). Portanto, o fato da ligação de PP2A a MT terem uma possível função estrutural não descarta o fato de que exista uma atividade enzimática de PP2A, ainda que essa atividade não seja necessária para a transformação mediada por MT.

1.3.3.2. Família Src de proteínas tirosina quinase

A família Src de proteínas tirosina quinases possui 11 membros descritos em humanos e sua atividade quinásica exerce papéis fundamentais em processos como diferenciação, motilidade, proliferação e sobrevivência celular. As proteínas da família Src possuem uma estrutura semelhante que consiste num domínio N-terminal, pelo qual se ancoram a membranas, um domínio SH3, que se liga a seqüências ricas em prolina, um domínio SH2, que se liga a fosfotirosinas, um domínio quinásico e uma pequena cauda C-terminal com função regulatória (Brown e Cooper, 1996).

Na família Src de proteínas quinases, pp60^{c-src} (Src) é expresso ubiquamente no organismo e consiste no membro mais estudado. O principal mecanismo de controle de Src é através da fosforilação de dois resíduos de tirosina, a Tyr-416 e a Tyr-527. Em sua conformação inativa, o resíduo Tyr-416 se encontra desfosforilado e interagindo diretamente com o domínio quinásico. Quando ocorre sua auto-fosforilação, este resíduo é deslocado do domínio quinásico permitindo o acesso do seu substrato. Um dos mecanismos de inativação de Src é através da proteína Csk (*Src C-Terminal Kinase*) que, através da fosforilação do resíduo Tyr-527 de Src, leva a uma interação intramolecular entre o resíduo fosforilado e o domínio SH2. Ainda que Src fosforilada no resíduo Tyr-416 também seja substrato para Csk, a fosforilação de Tyr-416 se sobrepõe a de Tyr-527, mantendo Src ativa. Em condições basais, cerca de 95% da Src se encontra fosforilada no resíduo Tyr-527, sendo necessário a atividade de uma fosfatase, como a proteína tirosina fosfatase- α (PTP α), para desfosforilação de Tyr-527, auto-fosforilação de Tyr-416 e conseqüente ativação de Src (Abram e Courtneidge, 2000; Roskoski, 2004). Os outros membros da família Src também apresentam sítios análogos para o controle da sua atividade.

Entre os membros da família Src, somente pp60^{c-src} (Src) (Courtneidge e Smith, 1983), e em menor escala, pp59^{c-fyn} (Fyn) (Cheng, Harvey *et al.*, 1988) e pp62^{c-yes} (yes) (Kornbluth, Sudol *et al.*, 1987) interagem com MT. Como descrito anteriormente, a interação prévia de PP2A com MT é necessária para a ligação de Src (Ogris, Mudrak *et al.*, 1999), havendo indícios de que a proteína quinase C (PKC) também seja necessária para essa ligação (Matthews e Benjamin, 1986).

A região de MT responsável pela interação com a proteína Src foi mapeada na porção N-terminal da molécula de MT (aminoácidos 185 a 210), sendo que nesta região também ocorre a interação MT/PP2A (Pallas, Shahrik *et al.*, 1990; Campbell, Auger *et al.*, 1995; Brewster, Glover *et al.*, 1997). Normalmente, as interações de Src ocorrem através dos seus domínios SH2 e SH3, entretanto, com MT é necessário somente o domínio quinásico e uma pequena região da cauda (aminoácidos 518 a 525 de Src) (Dunant, Senften *et al.*, 1996). Uma vez ligada à MT, o resíduo Tyr-527 Src é desfosforilado e o Tyr-416 fosforilado, levando à ativação de Src (Cartwright, Kaplan *et al.*, 1986). Como a ligação de PP2A funcional não é necessária para atividade de

Src e o resíduo Tyr-527 não interage com MT, possivelmente a desfosforilação de Tyr-527 é mediada por uma outra fosfatase recrutada por MT.

Assim que é ativada por MT, Src leva à fosforilação de diversos resíduos de tirosina em MT, sendo que os principais sítios de fosforilação são os resíduos Tyr-250, Tyr-315 e Tyr-322. Além dessa atividade em MT, ainda não foram descritas outras proteínas celulares como sendo substratos da Src ativada por MT.

1.3.3.3. Fosfatidilinositol-3 quinase (PI-3K)

A fosforilação da Tyr-315 de MT é essencial para a indução da transformação celular e resulta na sua ligação à proteína fosfatidinositol-3 quinase (PI-3K) (Figura 6). A proteína PI-3K é um heterodímero composto de duas subunidades, uma catalítica de 110KDa (p110) e outra regulatória de 85KDa (p85) (Carpenter, Duckworth *et al.*, 1990), sendo que p85 possui um domínio SH2 (domínios que se ligam a tirosinas fosforiladas) que permite sua ligação a MT (Courtneidge e Heber, 1987; Elis, Lessmann *et al.*, 2006).

O substrato preferencial da PI-3K ativada é o fosfatidilinositol 4,5 bisfosfato (PIP₂), o qual é fosforilado na posição 3 do anel de inositol gerando moléculas de fosfatidilinositol-3,4,5-trifosfato (PIP₃), um potente segundo mensageiro capaz de se ligar e ativar diversas proteínas (Serunian, Auger *et al.*, 1990).

Entre as proteínas que se ligam a PIP₃ está a proteína quinase Akt, cuja ativação consiste na principal via de sinalização decorrente da ativação de PI-3K. Akt é uma Serina/Treonina quinase de aproximadamente 57kDa, e também é conhecida por PKB (Proteína quinase B) e RAC (*A and C related protein kinase*). Com a produção de PIP₃, Akt se transloca para a membrana onde se liga a PIP₃ e sofre alterações conformacionais que expõem seus sítios de fosforilação (Tyr-308 e Ser-473). Depois de ativado (fosforilado), Akt é translocado para o citoplasma e para o núcleo, passando a modular a atividade de diversas proteínas (Vivanco e Sawyers, 2002; Osaki, Oshimura *et al.*, 2004).

Entre os efeitos da atividade de Akt, a inibição da glicogênio sintase quinase-3 (GSK-3) leva à indução da transcrição de diversos genes, como a

Ciclina D (Diehl, Cheng *et al.*, 1998), e a fosforilação das CDKIs p21^{WAF1/Cip1} e p27^{Kip2}, inibindo suas atividades anti-proliferativas (Liang e Slingerland, 2003). Akt também possui atividade anti-apoptótica ao inibir proteínas apoptóticas como Bad, que controla a saída de citocromo C da mitocôndria, inibir a transcrição de genes pró-apoptóticos como FasL, IGFBP-1 e Bim, e ativar cAMP, que leva à transcrição de genes anti-apoptóticos como Bcl-2, Mcl-1 e o próprio Akt (Franke, Hornik *et al.*, 2003).

Além de Akt, PI-3K também leva à ativação de outras proteínas como Rac (Hawkins, Eguinoa *et al.*, 1995), da via de transdução de sinal de Ras (Rodriguez-Viciano, Warne *et al.*, 1994; Hu, Klippel *et al.*, 1995) e de diferentes isoformas de PKC (Toker, Meyer *et al.*, 1994).

Estudos recentes têm mostrado que PI-3K também é capaz de se ligar não apenas ao resíduo Tyr-315 fosforilado de MT, mas também ao resíduo Tyr-250, o qual era visto como um sítio específico para Shc (Hong, Mikami *et al.*, 2003).

1.3.3.4. PLC- γ (Fosfolipase C - γ)

A proteína PLC- γ (fosfolipase C - γ) pertence à família C de fosfolipases que, de acordo com os seus domínios catalíticos X e Y, são subdivididas em três tipos, β , γ e δ . Entre estes tipos, somente a PLC- γ está envolvida no controle da proliferação celular, possivelmente devido à presença de dois domínios SH2 e um domínio SH3 entre os domínios X e Y, o que permite que ela se ligue a tirosinas fosforiladas e a seqüências ricas em prolina, respectivamente (Katan, 1998).

A interação de PLC- γ com MT se dá através de um dos seus domínios SH2 (Su, Liu *et al.*, 1995) com Tyr-322 fosforilada (Figura 6), levando à sua fosforilação em resíduos específicos de tirosina, principalmente Tyr-783 e Tyr-1254, aumentando assim sua atividade catalítica. A PLC- γ atua em PIP₂, cuja hidrólise leva à formação de inositol 1,4,5 trifosfato (IP3) e diacilglicerol (DAG), que são responsáveis, respectivamente, pela mobilização de íons Ca²⁺ intracelular de estoques internos, por meio da ligação a receptores específicos no retículo endoplasmático (Berridge e Irvine, 1984), e ativação de membros da família PKC, que leva à fosforilação de diversas proteínas celulares como GSK-3, a qual é ativada por PI-3K, e Raf (Morrison e Cutler, 1997), um membro da via de MAPK (Boyle, Smeal *et al.*, 1991;

Shin, Yoon *et al.*, 2002). O íon Ca^{2+} , por sua vez, é co-fator na ativação de PKC por DAG (Berridge, 1984). Mutações na proteína MT que impedem a sua associação à PLC- γ levam a uma proteína MT defectiva para a transformação maligna (Oliveira, Roberts *et al.*, 1998).

1.3.3.5. Shc (proteína que contém o domínio 2 de homologia a Src)

Algumas moléculas não possuem atividade catalítica intrínseca, mas funcionam como adaptadores que conectam um receptor a moléculas efetoras. Shc é uma proteína adaptadora de 52KDa que contém um domínio de ligação a fosfotirosinas (PTB), um domínio central de homologia a colágeno (CH1) e um domínio SH2 na porção C-terminal (Blaikie, Fournier *et al.*, 1997). Shc contém três sítios de fosforilação em tirosinas, Tyr-239, Tyr-240 e Tyr-317. Uma vez ligado a Tyr-250 fosforilada de MT, ocorre a fosforilação em uma das três tirosinas do domínio PTB, que então se torna sítio de interação para Grb2 (growth factor receptor-bound protein 2). Grb2 é uma outra molécula adaptadora que possui um domínio SH2, com o qual se liga a Shc, e dois domínios SH3, responsável pela interação de Grb2 com Sos (Son of sevenless homolog), um fator de troca de nucleotídeo de Ras que irá converter Ras da forma inativa, Ras-GDP, para a forma ativa, Ras-GTP (Figura 6) (Li, Batzer *et al.*, 1993).

Ras é um membro da família de proteínas codificadas pela família de oncogenes *ras*, homólogos celulares dos genes transformantes dos vírus de sarcoma de Harvey e Kirstein, que desempenha um papel importante na proliferação celular. Algumas das vias ativadas por Ras são denominadas vias de MAPK (*Mitogen-Activated Protein Kinase*) que envolvem uma cascata de quinases, as quais enviam o sinal mitogênico para o núcleo. As MAPKs são ativadas por fosforilação em resíduos de treonina e tirosina por uma quinase de especificidade dual, conhecida como MAPK quinase (MAPKK) que, por sua vez, é ativada por fosforilação em seus resíduos de serina por outra quinase, a MAPKK quinase (MAPKKK). O sinal $\text{MAPKKK} \rightarrow \text{MAPKK} \rightarrow \text{MAPK}$ é então passado adiante por translocação da MAPK para o núcleo e fosforilação de fatores de transcrição específicos. Entre as vias de MAPKs, a melhor conhecida é a via $\text{Ras} \rightarrow \text{RAF (MAPKKK)} \rightarrow \text{MEK (MAPKK)} \rightarrow \text{Erk (MAPK)}$, descrita a seguir:

O primeiro passo da via Ras/Erk é a ligação de Ras-GTP à Serina/Treonina quinase Raf (homólogo de v-Raf – produto do gene transformante *v-raf* do vírus de sarcoma murino 3611) (Zhang, Settleman *et al.*, 1993); entretanto, esta associação, por si só, não promove ativação de Raf, mas permite a sua localização na membrana plasmática onde será fosforilada devido a mecanismos de ativação que envolvem a quinase Ksr (*Kinase suppressor of Ras*), PP2A e uma proteína da família 14-3-3 (Raabe e Rapp, 2003). Através da fosforilação de resíduos específicos de serina, Raf ativa MEK-1 ou MEK-2 (quinases de MAPKs/Erks), as quais ativam Erk-1 e Erk-2 (quinases reguladas por sinais extracelulares, pertencentes à família de MAPKs) por fosforilação em resíduos de treonina e tirosina. No seu estado ativado, as ERKs são capazes de passar para o núcleo onde catalisam a fosforilação de fatores de transcrição como o fator de transcrição Elk-1 (Campbell, Khosravi-Far *et al.*, 1998).

Um outro exemplo de indução da via de MAPKs por Ras é dado pela ativação de uma MAPKKK (que não Raf) resultando na fosforilação de JNK (Quinase da região N-terminal de Jun, pertencente à família das MAPKs) (Hibi, Lin *et al.*, 1993; Minden, Lin *et al.*, 1994) a qual regula a atividade de Jun, um dos componentes do fator de transcrição AP-1.

A associação de MT a Shc é essencial para a transformação celular por MT, já que mutação no resíduo de Tyr-250, que impede esta associação, torna o antígeno MT completamente defeutivo para transformação celular, enquanto que a re-introdução do domínio restaura sua capacidade transformante (Oliveira, Roberts *et al.*, 1998).

A associação de MT a PI-3K também parece ser necessária para um aumento na atividade de Erk, já que alguns mutantes de MT defeitivos para a ligação a PI-3K apresentam níveis baixos de Erk fosforilada (Oliveira, Roberts *et al.*, 1998). O *cross-talk* entre as vias de Ras e PI-3K já foi descrito, entretanto, os resultados parecem variar de acordo com o tipo celular estudado (Rodriguez-Viciano, Warne *et al.*, 1994; Hu, Klippel *et al.*, 1995; Campbell, Khosravi-Far *et al.*, 1998).

Assim, como MT apresenta a capacidade de influenciar no controle proliferativo, o estudo das suas vias de ação representa uma poderosa

ferramenta na caracterização dos mecanismos de controle do ciclo celular e da transformação maligna.

1.4. Busca de genes diferencialmente expressos

Uma das maneiras utilizadas para descrever as bases moleculares de um determinado evento biológico consiste na busca de genes diferencialmente expressos entre duas condições que se queira estudar. Portanto, visando uma melhor compreensão do processo tumoral desencadeado pelo antígeno MT, o principal foco deste projeto consistiu em identificar e caracterizar novos genes regulados pela ação do antígeno MT.

Na busca dos genes regulados por MT foram utilizadas três abordagens: a) hibridização de membranas (macroarranjos de cDNA) comerciais; b) geração de bibliotecas subtraídas de cDNA, através da técnica de Análise de Diferença Representacional (RDA); c) hibridização de lâminas de microarranjos de oligonucleotídeos (*CodeLink Bioarrays*® - Amersham Biosciences)

As membranas comerciais são membranas de nylon que apresentam diferentes cDNAs, a maioria dos quais previamente descritos, imobilizados de maneira ordenada. Assim, cada uma de duas membrana idênticas é hibridizada com uma das duas populações de cDNAs marcados de forma radioativa ou quimioluminescente. Em geral, estas duas populações de cDNAs são provenientes de mRNA extraído de células mantidas em duas condições diferentes como, por exemplo, células que expressam ou não MT. Essa abordagem permite a identificação rápida dos genes que apresentam expressão diferencial nas populações analisadas.

A metodologia de Análise de Diferença Representacional (RDA) foi originalmente desenvolvida para ser utilizada como um método que permitisse encontrar diferenças entre dois genomas e, posteriormente, foi adaptada para a subtração de cDNA. O princípio da técnica consiste em gerar duas populações de cDNAs derivados a partir de mRNA de células em duas condições diferentes e, após etapas de subtração acopladas a amplificações por PCR, os fragmentos comuns às duas populações têm sua representação drasticamente reduzida, resultando em uma biblioteca enriquecida para genes diferencialmente expressos em uma das populações analisadas (Hubank e Schatz, 1994; Braun, Frieden *et al.*, 1995; Welford, Gregg *et al.*,

1998). Construída a biblioteca, a seleção dos cDNAs candidatos a serem diferencialmente expressos pode ser realizada através das metodologias de macro e/ou microarranjos de DNA (Wang e Brown, 1991). Devido à redundância dos resultados provenientes do RDA, após a busca é realizado o sequenciamento dos clones selecionados e uma posterior confirmação da sua expressão diferencial por *Real-Time PCR*.

A plataforma *CodeLink* consiste em microarranjos (*microarrays*) de oligonucleotídeos pré-fabricados, onde as lâminas são previamente cobertas com um polímero hidrofóbico formando uma matriz 3D que orienta os oligonucleotídeos imobilizados para o sentido oposto ao do vidro. A característica hidrofóbica do polímero dificulta a ligação de sondas inespecíficas tanto ao DNA imobilizado quanto à lâmina de vidro, diminuindo assim o *background*. Juntamente com os oligonucleotídeos, é dispensado um fluoróforo derivado de fluoresceína que permite uma futura inspeção da qualidade de cada *spot* dispensado. Após a ligação dos oligonucleotídeos ao polímero, sítios adicionais para a ligação são bloqueados, o *bioarray* é lavado, seco e então é colado sobre a lâmina um adesivo de polipropileno que irá formar a câmara de hibridização (Ramakrishnan, Dorris *et al.*, 2002).

O *CodeLink* possui algumas vantagens em relação às outras plataformas de microarranjos. Primeiramente, o fato dos oligos estarem ligados em uma matriz tridimensional aumenta o número de oligos por área de lâmina, além de facilitar o acesso das sondas aos oligos, aumentando assim a sensibilidade da técnica. Para dispensar as amostras, é utilizado um robô acoplado a um dispensador pizoelétrico, que garante maior precisão do volume dispensado, além de diminuir as chances de contaminação da lâmina, uma vez que o mesmo não toca a lâmina. Como os oligos são aplicados juntamente com um fluoróforo é realizado, após a confecção da lâmina, um escaneamento do *bioarray*, a fim de identificar *spots* que apresentaram alguma irregularidade, quase que impossibilitando a utilização de *spots* que apresentem um falso sinal negativo (Dorris, Nguyen *et al.*, 2003).

2. OBJETIVOS

2.1. Objetivo Geral

Compreender as bases moleculares da transformação maligna induzida pela oncoproteína MT do Vírus Polioma, através da busca e identificação de genes diferencialmente expressos em células que super-expressam MT.

2.2. Objetivos Específicos

a) Identificar novos genes e novas vias de sinalização relacionados à transformação maligna induzida pelo antígeno MT de polioma vírus.

b) Verificar o nível de expressão dos genes selecionados em células expressando o MT selvagem bem como células expressando a proteína MT mutante para determinadas vias, utilizando as metodologias de RT-PCR e *Real-Time PCR*.

c) Identificar novas vias de sinalização intracelular utilizando, como ferramenta, células que expressam MT selvagem e diferentes mutantes.

3. MATERIAIS E MÉTODOS

3.1. Materiais

3.1.1. Plasmídeos

pBlueScript SK - Vetor de clonagem, de elevado número de cópias. (Stratagene).

pLPCX - Vetor de expressão retroviral contendo elementos do vírus Moloney de leucemia murina (MoMuLV) e do vírus Moloney de sarcoma murino (MoMuSV) e um promotor interno do citomegalovírus humano, além de um cassete que confere resistência à puomicina (Clontech).

pLPCX-EGFP - Vetor de expressão retroviral pLPCX contendo o cDNA codificante para EGFP (*Enhanced Green Fluorescent Protein*) construído e cedido pelo Dr. Marcos A. Demasi (pós-doutorando em nosso laboratório).

pLPCX-MN1 - Vetor de expressão retroviral contendo a região codificadora de NRP/B selvagem de rato, construído neste trabalho 38

As seguintes construções foram geradas em colaboração com o Dr. Ricardo Corrêa, que então se encontrava no Salk Institute for Biological Studies (Tiscornia, Singer *et al.*, 2003):

pMDL - Vetor codificante das seqüências gag e pol.

pREV - Vetor codificante das seqüências do envelope viral, que regula a expressão lentiviral ao nível pós-transcricional.

pVSVG - Vetor codificante das seqüências do envelope viral heterólogo, derivado de *G glycoproteins of vesicular stomatitis vírus*.

pEGFP - Vetor lentiviral que expressa o gene EGFP (*Enhanced Green Fluorescence Protein*).

siMN1 - Vetor lentiviral contendo a sequência correspondente ao RNAi para MN1 e a sequência correspondente à proteína EGFP, possibilitando a fácil confirmação da transfecção e infecção celular sob microscopia de fluorescência.

3.1.2. Linhagens celulares

A31 - clone celular obtido a partir de células Balb-3T3 (célula mesenquimal de embrião de camundongo Balb-c; (Aaronson and Todaro, 1968) cedido pelo laboratório do Dr. Charles Stiles, Dana Faber Cancer Institute, Harvard Medical School, Boston, MA, USA.

PLJ - A31 infectada com vetor retroviral PLJ, que contém um marcador de resistência à geneticina (G418).

MTWT - A31 infectada com o vetor retroviral PLJ-MTWT, que codifica a proteína MT selvagem do Vírus Polioma.

MT315 - A31 infectada com o vetor retroviral PLJ-MT315, que codifica a proteína MT do Vírus Polioma mutada na posição 315 (tyr→phe).

MT322 - A31 infectada com o vetor retroviral PLJ-MT322, que codifica a proteína MT do Vírus Polioma mutada na posição 322 (tyr→phe).

MT2F - A31 infectada com o vetor retroviral PLJ-MT2F, que codifica a proteína MT do Vírus Polioma mutada nas posições 315 (tyr→phe) e 322 (tyr→phe).

250phe - A31 infectada com o vetor retroviral PLJ-250phe, que codifica a proteína MT do Vírus Polioma mutada na posição 250 (tyr→phe).

293T - Linhagem de rim embrionário humano, transformada com a proteína LT de SV40 (Dubridge, Tang *et al.*, 1987).

Φnx-Ampho - Linhagem celular empacotadora anfotrópica, derivada da linhagem de rim embrionário humano transformada 293T, fornecidas pelo Dr. Gary Nolan - Universidade de Stanford - CA - USA.

3.1.3. Soluções e Meio de Cultura para células de mamíferos

DMEM - Meio de Cultura DMEM (*Dulbecco's Modified Eagle Medium* - Gibco BRL-Life Technologies, Inc., USA).

SFB - Soro Fetal Bovino (Cultilab Materiais para Cultura de Células, Brasil).

PBSA - Solução salina (*phosphate buffered saline*), sem cálcio ou magnésio, tamponada em pH 7,2 composta por NaCl 140mM, KCl 2,7mM, Na₂PO₄ 8mM e KH₂PO₄ 1,4mM.

Tripsina - Solução 0,1% diluída em PBSA, contendo EDTA 1mM (ICN Pharmaceuticals Inc; Gibco Limited).

Antibióticos - puromicina (Sigma).

Hidrocortisona - polibreno (Sigma).

3.1.4. Soluções e Meio de Cultura para bactérias

LB - Meio de Luria-Bertani para cultivo bacteriano, composto por triptona 10g/L, extrato de levedura 5g/L e NaCl 10g/L, pH 7,5.

SuperBroth - Meio para cultivo bacteriano, composto por triptona 32g/L, extrato de levedura 20g/L, NaCl 5g/L, pH 7,5.

Ampicilina - Solução estoque 100mg/ml, em água (Sigma, USA).

Kanamicina - Solução estoque 50mg/mL, em etanol (Sigma, USA).

3.1.5. Isótopos radioativos

[α ³²P] dCTP 3.000 Ci/mmol (Amersham Biosciences).

3.1.6. Soluções

Todas as soluções foram preparadas utilizando-se reagentes de grau de pureza para análise, seguindo especificações de manuais de laboratório (Ausubel, Brent *et al.*, 2001; Sambrook e Russell, 2001).

Agarose para ensaio de crescimento celular em suspensão - Agarose Medium EEO (Fisher Scientific, New Jersey, USA).

Gel de agarose - UltraPure TM Agarose (Invitrogen).

Tampão MOPS - Acetato de sódio 10mM; EDTA 1mM; MOPS 40mM.

SSC 20X - NaCl 3M; citrato trissódico dihidratado 300mM; pH 7,0.

Solução I para lavagem de membrana de Northern blot - SSC 2X; SDS 0,1%.

Solução II para lavagem de membrana de Northern blot -SSC 0,1X; SDS 0,1%.

TBE (5X) - 446mM Tris-Base; 445mM ácido bórico; 10mM EDTA: pH 8,0.

Tampão de amostra para DNA - formamida deionizada 98%; EDTA 10mM; xileno cianol 0,025%; azul de bromofenol 0,025%.

Tampão de lise para extração de RNA - isotiocianato de guanidina 4M; citrato de sódio 25mM pH 7,0; β -mercaptoetanol 0,1M.

Solução de cloreto de cézio - cloreto de cézio 5,7M; acetato de sódio 25mM; pH 5,0.

TE - Tris-HCl 10mM, EDTA 1mM, pH 8,0.

TBS - NaCl 150mM, Tris-HCl 50mM, pH 7,5. 42

TBST - NaCl 150mM, Tris-HCl 50mM, pH 7,5; Tween 20® 0,05%.

RIPA+ - Tris-HCl 10mM pH7,5, desoxicolato de sódio 1%*m/v*, Nonidet P40 1%*v/v*, NaCl 150mM, SDS 0,1%*m/v*, PMSF 1mM, DTT 1mM, Aprotinina 1 μ g/mL, Leupeptina 5 μ g/mL, pepstatina 1 μ g/mL, NaF 10mM.

Tampão de amostra para proteína (3x) - Tris-HCl 150mM pH6,8; SDS 6%*m/v*, glicerol 30%*v/v*, β -mercaptoetanol 15%*v/v*, azul de bromofenol 0,3%*m/v*.

Tampão de transferência - Glicina 39mM, Tris-HCl 48mM pH8,3, SDS 0,037%*m/v*, metanol 20%*v/v*.

Corante de Ponceau - Ponceau 0,1%*m/v*, ácido acético glacial 10%.

Tampão hipotônico - Hepes 10mM pJ7,9, KCl 1mmM, EGTA 0,1mM, EDTA 0,1mM, 1X Protease Inhibitor *Mix* (Amersham, NJ, EUA).

Tampão de extração nuclear - Hepes 20mM pH7,9, NaCl 400mM, EGTA 1mM, EDTA 1mM, 1X Protease Inhibitor *Mix* (Amersham, NJ, EUA).

Tampão CSK - NaCl 100mM, sacarose 300mM, Pipes 10mM pH6,8, MgCl₂ 3mM, EGTA 1mM, Triton X-100® 0,5%v/v, 1X Protease Inhibitor *Mix* (Amersham, NJ, EUA).

Tampão RSB-Majic - NaCl 10mM, MgCl₂ 3mM, Tris-HCl 10mM pH7,4, Tween®20 1%v/v, deoxicolato de sódio 0,5%v/v, 1X Protease Inhibitor *Mix* (Amersham, NJ, EUA).

3.2. Cultura de linhagens celulares

As linhagens celulares foram cultivadas em frascos plásticos adequados para o cultivo de células, contendo meio de Eagle modificado por Dulbecco (DMEM) suplementado com 10% de soro fetal bovino (SFB). O pH e a temperatura foram mantidos na faixa fisiológica através da incubação em estufa a 37°C e atmosfera de 2% de CO₂.

O sinal mediado por MT mimetiza a ação de fatores de crescimento presentes no SFB. Para que estes fatores não mascarassem a ação de MT, antes de cada ensaio as culturas foram carenciadas, para soro, por 48h em DMEM suplementado com 0,5% de SFB.

Estoques celulares foram congelados em meio contendo 10% de dimetilsulfóxido (DMSO), e mantidos em reservatório contendo nitrogênio líquido.

3.3. Extração e purificação de RNA

As linhagens celulares foram plaqueadas em placas de 150mm de diâmetro (P150) e, após carenciamento de 48h para soro, o meio de cultura foi removido e as culturas foram lavadas com PBS. O PBS foi removido e adicionou-se 2ml de solução de lise (4M isotiocianato de guanidina; 25mM citrato de sódio pH 7,0 e 0,1M beta-mercaptoetanol). O lisado celular foi cuidadosamente aplicado sobre um colchão de Cloreto de Césio (5,7M CsCl e 25mM NaAc pH 5,0) e as amostras centrifugadas a 37.000g, 20°C por 17h (Chirgwin, Przybyla *et al.*, 1979).

Após a centrifugação, o sobrenadante foi coletado com uma pipeta Pasteur até sua completa remoção. O precipitado (RNA total) foi eluído em 50µl de H₂O-MilliQ. O RNA foi quantificado através da absorbância da amostrada quando atingida por luz de comprimento de onda de 260nm. O

grau de pureza foi determinado através da razão das medidas nos comprimentos de onda a 260 e 280nm.

3.4. Síntese de cDNA

3.4.1. Síntese da primeira fita do cDNA

RNA total, resultante do item 3.3, foi utilizado como *template* para a síntese de cDNA em uma reação da transcrição reversa. Assim, foram preparadas as reações contendo 4µg de RNA total, extraído das linhagens utilizadas (item 3.2), 1µl de Oligo dT 500ng/ml (Amersham Biosciences), 1µl de *Random Primer* 100ng/µl e H₂O para um volume final de 24µl. As amostras foram incubadas a 75°C por 10min e, em seguida, adicionou-se 14µl de uma solução contendo 4µl de Buffer 5x para enzima Super Script (Invitrogen), 4µl de DTT (Ditiotrietol) e 2µl de RNase OUT 40U/µl. Após uma nova incubação de 25°C por 10min, a 42°C por 2min foi adicionado 2µl de enzima SuperScript 200U/µl (Invitrogen).

A reação se deu durante uma incubação de 2h a 42°C. Em seguida, para inativação da enzima, a amostra foi incubada a 72°C por 10min. Visando degradar o RNA, foi adicionado 1µl de RNase H 5U/µl em cada tubo e os tubos foram incubados a 37°C por 30min, e em seguida incubados a 72°C por 10min para inativação da RNase H. Posteriormente, as amostras foram diluídas em 60µl de TE.

3.4.2. Controle da qualidade do cDNA

Para confirmar a qualidade do cDNA sintetizado foram realizadas reações de PCR controle. O primeiro controle consistia na amplificação de GAPDH (gliceraldeído 3-fosfato desidrogenase), um gene *housekeeping* com expressão ubíqua. Este era o controle interno da reação de transcrição reversa. O segundo controle era contra um gene de baixa expressão, p53, para checar se genes pouco expressos estavam representados na amostra. O terceiro controle era contra a contaminação com DNA genômico utilizando oligos desenhados nos íntrons do gene MLH1 (*mutL homolog 1*). O quarto controle checava a presença de transcritos grandes na amostra, amplificando

um fragmento de um gene de 12Kb do gene NOTCH2 (*Notch gene homolog 2*).

Portanto, utilizando os quatro diferentes conjuntos de *primers* (GAPDH, p53, MLH e NOTCH2), cada reação foi montada de forma a apresentar uma concentração final de 1x para o Tampão de Reação (Biolase), 1,5mM de MgCl₂, 0,2mM de dNTP, 0,2μM de cada um dos *primers* (*Forward* e *Reverse*), enzima Taq (Biolase) e *template*.

Nas reações onde foram utilizados os *primers* p53, GAPDH e MLH foi utilizado um programa de 94°C por 2min, 35 ciclos de 94°C por 30 segundos, 56°C por 30 segundos e 72°C por 45 segundos e, após os ciclos, 72°C por 6min, enquanto que para NOTCH2 o programa de amplificação foi alterado para 94°C por 2min, 40 ciclos de 94°C por 30 segundos, 56°C por 30 segundos e 72°C por 1 minuto e, após os ciclos, 72°C por 10min (Sogayar, Camargo *et al.*, 2004).

Os produtos das reações de PCR foram posteriormente analisados em um gel de agarose 1% contendo Brometo de Etídeo.

3.4.3. Purificação do mRNA

Para purificação do mRNA a partir do RNA total (item 3.4.1) foi utilizado o Kit comercial *PolyAtract® mRNA Isolation System* (Promega). No kit, o *primer* oligo(dT) biotinilado se liga ao mRNA e, após a adição de estreptavidina associada a um partícula magnética, o complexo mRNA-oligo(dT)-biotina-estreptavidina-partícula magnética é separado pela exposição da amostra a um campo magnético.

3.4.4. Síntese de cDNA dupla fita (*Smart cDNA Synthesis Kit*)

Inicialmente, foi sintetizada a primeira fita do cDNA através do Kit *Smart PCR cDNA Synthesis Kit* (Clontech). Brevemente, foi montada uma reação contendo 1μg de mRNA e os *primers* "CDS primer IIA – Sal 3A1" e "Smart IIA Oligo – Sau3A1". A amostra foi aquecida a 72°C, para anelamento do *primer* "CDS primer IIA" e posteriormente resfriada a 4°C. Para a reação de transcrição reversa, foi adicionado o *mix* contendo Tampão, DTT, dNTP, Transcriptase Reversa e Inibidor de RNase, nas concentrações recomendadas pelo fabricante. Como uma das limitações do kit é a de não sintetizar cDNA a partir de mRNAs muito longos, foi adicionado ao *mix*

Trealose (2M) e Betaína (0,6M). A reação foi incubada a 42°C por 35min, para a síntese da 1ª fita da maior parte dos transcritos, 60°C por 1 hora e 42°C, para síntese da 1ª fita dos transcritos ricos em GC, e por 45min 42°C, para finalizar a síntese das primeiras fitas. Terminada a reação, para inativar a Transcriptase Reversa foi adicionado TE e a amostra foi aquecida a 72°C por 7min.

A uma diluição da amostra foi adicionado o *mix* para a síntese da segunda fita e amplificação dos transcritos (Tampão, MbSO₄, dNTPs, *primer* "PCR Smart IIA Óligo Sau 3A1 ", Taq HiFi - Invitrogen). O programa utilizado pelo termociclador foi alterado, em relação ao sugerido pelo fabricante, devido à adição da Trealose e Betaína. Assim, após a desnaturação inicial a 95°C, foram realizados 15 ciclos de 95°C por 5 segundos, 65°C por 5 segundos e 68°C por 8min. Uma alíquota foi retirada para análise em gel. pós três ciclos, foi retirada outra alíquota e esse procedimento foi repetido até um total de 31 ciclos.

As amostras foram amplificadas e então purificadas em colunas do kit *GFX PCR cDNA and Gel Band Purification Kit* (Amersham), conforme recomendações do fabricante. As colunas GFX são feitas em matriz de fibra de vidro à qual se liga o DNA na presença de agentes caotrópicos, permitindo sua purificação como remoção de proteínas, nucleotídeos e sais. Fragmentos muito pequenos, tais como *primers*, se ligam permanentemente à coluna, não sendo eliminados durante a purificação.

3.5. Sequenciamento de DNA

Nas reações de sequenciamento foi utilizado o Kit *ABI PRISM® BigDye® Terminator v3.1 Cycle Sequencing Kit* (Applied Biosystems). O Kit baseia-se na metodologia do seqüenciamento com dideoxinucleotídeos e nucleotídeos marcados acoplado a PCR com apenas um *primer*.

Para as reações, foi utilizado 1µl do produto de PCR de colônia, 3,2pmol de *primer* 10µM, Big Dye e água. Decorridos 40 ciclos, de 96°C por 10seg, 52°C por 20seg e 60°C por 4min, foi adicionado 20µl de água a cada reação, seguido de uma nova adição de 60µl de isopropanol.

As amostras foram homogeneizadas, mantidas à temperatura ambiente por 15min e, em seguida, centrifugadas a 3.000g, por 30min a 20°C. O sobrenadante foi descartado e adicionou-se 150µl de isopropanol

70%. Após uma nova centrifugação a 3.000g e 20°C por 10min o sobrenadante foi novamente descartado e a placa foi colocada para secar por 10min a 37°C. Esta reação foi ressuspensa e fracionada no aparelho de seqüenciamento automático ABI 377 (Perkin Elmer).

3.6. Subtração dos genes diferencialmente expressos

3.6.1. Geração das bibliotecas subtraídas através do RDA

Na produção das bibliotecas subtraídas, a população de cDNAs que continha os genes diferencialmente expressos foi denominada *Tester* enquanto os cDNAs da população controle foi denominada *Tester*.

Para diminuir a complexidade das amostras, 40µl de cDNA dupla fita foi digerido por 16h em uma reação contendo 50U de *Mbo*I (Fermentas) e 1X Tampão R (Fermentas). Os fragmentos resultantes da digestão foram purificados em colunas GFX e posteriormente amplificados em uma reação contendo 1,2µg do cDNA digerido com *Mbo*I, Tampão de T4 DNA ligase 1X (Fermentas), 16,5µM do adaptador R12 e 16,5µM do adaptador R24. Os oligos R12 e R24 pareiam entre si formando pontas coesivas e complementares aos fragmentos digeridos pela enzima *Mbo*I. As amostras foram aquecidas a 50°C por 1 minuto e resfriadas a 1°C/min até atingirem 10°C. Após a estabilização da temperatura a 10°C foi adicionado 1µl de T4 DNA ligase (Fermentas) e a temperatura foi elevada a 14°C por 16h. Posteriormente, a ligase foi inativada por temperatura, 65°C por 10min, e as amostras diluídas com 40µl de TE.

Após outra purificação com GFX, as amostras foram amplificadas numa reação contendo 1µl de amostra, Tampão 1X, MgCl₂ 1,5mM, dNTPs 0,32mM e oligo R24 1,3µM. As reações foram incubadas a 72°C por 3 minutos, adicionou-se de 5U de Taq DNA polimerase (Biolase) e, após uma nova incubação a 72°C por mais 5min, foram realizados 5 ciclos de 95°C por 30seg e 72°C por 3min. Retirou-se alíquotas do 8º ao 15º ciclo para a identificar os ciclos onde a amplificação estava em sua fase exponencial. Escolhido o número ideal de ciclos, uma nova reação de 1ml foi realizada e, em seguida, as amostras foram purificadas com GFX.

Para formar a população *Driver*, o adaptador R24 foi removido através de uma digestão contendo 5U de *Mbo*I (Fermentas), tampão R+ (Fermentas)

1X e 535µl de amostra por 16h a 37°C, seguida de purificação em coluna GFX. Para formar a população *Tester* foi ligado ao produto da digestão o adaptador J12 e J24, seguindo o mesmo protocolo que a ligação do adaptador R12 e R24.

Adicionou-se um excesso de *Driver* (10mg) a população *Tester* (0,1mg) e a amostra foi destanurada a 98°C por 5min e, em seguida, reduzida a 67°C, adicionou-se 1 ml de NaCl 5M e a amostra foi incubada por 20h a 67°C, possibilitando uma completa desnaturaçõ para hibridização. O excesso de *Driver* foi colocado a fim de favorecer a formação de híbridos *Driver-Tester* entre os genes presentes nas duas populações, sendo que os fragmentos existentes somente na população *Tester* ou *Driver* formaram as conformações *Tester-Tester* e *Driver-Driver*, respectivamente.

Ao final da hibridização, os adaptadores foram preenchidos e um novo PCR foi realizado. Para este PCR utilizou-se 20µl da hibridização diluída 1:200, Tampão 1X, MgCl₂ 1,50mM, dNTP 0,32mM e oligo J24 1,3µM. As amostras foram incubadas a 72°C por 7 minutos, adicionou-se 8U de Taq DNA polimerase (Biolase), e me seguida 11 ciclos de 95°C por 1min e 70°C por 3min. Uma vez que os híbridos *Tester-Tester* possuíam adaptadores nas duas extremidades, estes foram amplificados exponencialmente, enquanto os híbrido *Driver-Tester* foram amplificados linearmente, pois só possuíam adaptador em uma das extremidades. As conformações *Driver-Driver* não foram amplificadas, pois nenhuma das extremidades possuía adaptadores. Este primeiro PCR, onde já houve algum enriquecimento dos genes diferencialmente expressos em relação à população inicial, foi denominado primeiro produto diferencial (DP1).

As amostras foram purificadas por GFX e, visando digerir os adaptadores que estivessem presentes nos híbridos *Tester-Driver*, sítios para ligação de futuros adaptadores e cDNAs fitas simples, foi realizado um tratamento com a nuclease Mung Bean (MBN - Fermentas), que digere DNA simples fita. A reação de tratamento com Mung Bean foi realizada com 34µl de DPI purificado, Tampão MBN 1X e 80U de MBN (Amersham), por 35min a 37°C. As amostras foram incubadas novamente purificadas com o Kit GFX.

Em seguida, o adaptador J12/24 foi removido e os adaptadores R12/R24 ligados, com protocolos análogos as digestões e ligações anteriores. Assim, uma nova diluição do *Tester* foi colocada para desnaturar e hibridizar

com um novo excesso de *Driver* (800x o *Driver*). Após o preenchimento dos adaptadores e nova amplificação obtivemos o DP2 (Figura 7).

Foram realizados dois tipos de subtração, sendo que na primeira o cDNA proveniente da linhagem MTWT foi utilizado como *Tester*, enquanto aqueles da linhagem PLJ como *Driver*. Na segunda subtração o cDNA de MTWT foi utilizado como *Driver* e o de PLJ como *Tester*. Essas subtrações, onde temos os genes induzidos e reprimidos por MT, foram denominadas subtração *Forward* e *Reverse*, respectivamente.

Figura 7 - Esquema do RDA

Representação esquemática das etapas do RDA (*Representational Difference Analysis*)

3.6.2. Clonagem das bibliotecas subtraídas

Para clonagem, os produtos do PCR resultantes da subtração foram tratados com *Mbo*I para a remoção do adaptador e recuperação das pontas coesivas. O vetor de clonagem pBluescript II SK - foi digerido com BamH1, cujas extremidades são coesivas com aquelas geradas por *Mbo*I, e desfosforilado. Os cDNAs foram clonados em vetor de clonagem pBluescript constituindo as bibliotecas subtraídas. Os plasmídeos foram utilizados para transformar bactérias DH10B eletrocompetentes, as quais foram plaqueadas em meio seletivo contendo ampicilina, X-gal (4-bromo-5-cloro-indolil-galactopiranosídeo) e IPTG (isopropil- β -D-tio-galactopiranosídeo). Em seguida foram incubadas a 37°C por 12 horas para o crescimento das colônias transformantes.

3.6.3. Rastreamento dos clones diferencialmente expressos

Bactérias crescidas e congeladas em placas de 96 poços foram repicadas (aproximadamente 1 μ l) para placas de PCR contendo 14 μ l de *mix* de PCR. Este *mix*, para 100 reações, continha 1226 μ l de H₂O, 150 μ l de tampão de PCR 10x, 45 μ l de MgCl₂ 50mM, 19 μ l de dNTPs 10mM, 25 μ l de cada *primer* (M13F/M13R) 10 μ M e 10 μ l de *Taq* 5U/ μ l. A reação foi submetida à 95°C para lise das bactérias, 40 ciclos de 95°C por 45seg, 57°C por 45seg, 72°C por 1min e, após os ciclos, 72°C por 5min. O produto da reação foi fracionado em gel de agarose 1%.

As amostras foram aplicadas em membranas de nylon Hybond-XL (Amersham) através de um sistema de aplicação manual adaptado a partir daquele proposto por (Schummer, Ng *et al.*, 1997). O sistema, *ExpreCISION™ DNA Arraying System* (Bioinformática), permite a aplicação de 768 amostras, em duplicata, em cada membrana, sendo transferido 0,1 μ l de amostra por aplicação. A concentração do produto do PCR utilizada era em média 225ng/ μ l e, como foram realizadas duas aplicações, foi transferido cerca de 50ng de cada amostra. Entre cada aplicação, o *spotter* foi lavado com *Bleach Solution* (Água sanitária 10%), em seguida com H₂O-MilliQ e, por fim, flambada com etanol 90%, para diminuir e homogeneizar a tensão superficial dos pinos.

Terminadas as aplicações, as membranas foram desnaturadas em solução NaOH 0,6N por 10 minutos e este neutralizado em Tris-HCl 0,5M pH

7,5 por cinco minutos. A fixação de DNA às membranas foi realizada através da incubação em forno a 80°C por duas horas.

3.6.4. Síntese das sondas radioativas

Para a marcação radioativa das sondas, utilizou-se o kit *Ready-to-go DNA labelling beads (-dCTP)* (Amersham Biosciences), através da reação de *random primer extension* (Feinberg e Vogelstein, 1984).

Neste kit, aproximadamente 45µl de solução contendo 50ng do cDNA a ser marcado é desnaturado e utilizado para ressuspender *beads* contendo todos os componentes necessários para a reação de marcação, como os nucleotídeos (com exceção do [α^{32} P]dCTP que é adicionado em seguida), *random primers*, enzima Klenow e tampão. Durante a incubação a 37°C, os *primers* anelam em sítios randômicos no DNA, permitindo que a enzima Klenow sintetize as fitas complementares, incorporando o nucleotídeo radioativamente marcado durante o processo. As sondas foram, posteriormente, purificadas em colunas *Microspin S-300* (Amersham Biosciences) de acordo com as especificações do fabricante.

Após a purificação, a atividade específica da sonda (CPM incorporado por µg de DNA) foi determinada em cintilador líquido. Assim, a mesma quantidade de CPMs foi utilizada na hibridização de cada par de membranas a serem comparadas.

3.6.5. Hibridização dos macroarranjos de cDNA

As membranas foram incubadas por 1h a 42°C em solução de pré-hibridização (SSPE 5x, Formamida 50%, Denhardt 5x, SDS 0,1% e DNA de esperma de salmão desnaturado 100µg/ml). A seguir, as amostras foram hibridizadas em solução de hibridização (solução de pré-hibridização acrescida da sonda radioativa desnaturada) por 16h a 42°C.

Após a hibridização, as membranas foram lavadas por 2x em solução SSC 2x e SDS 0,1%, 15 minutos cada lavagem, e 4x em solução SSC 20x e SDS 0,1%, 30 minutos cada lavagem, sendo todas as lavagens a 56°C.

As membranas foram expostas a uma tela *PhosphorImager* de 2-15 dias (dependendo da qualidade da marcação) e, em seguida, lidas em um *scanner* a laser de alta definição (*Storm* – Molecular Dynamics).

3.6.6. Análise da densitometria

Após exposição das membranas à tela de *PhosphorImager* e digitalização das telas pelo Storm, os sinais presentes nas membranas foram densitometrados pelo programa *ImageQuant* (Molecular Dynamics).

O sinal da densitometria foi normalizado de acordo com o princípio de que a reta que descreve a intensidade dos sinais de uma membrana em função da outra possui coeficiente angular igual a 1. Baseado nesse conceito, foi utilizado como fator de correção a razão das medianas dos sinais apresentados em cada membrana.

Após a normalização, para cada conjunto de membranas foi escolhido um fator de corte (*cut-off*), ou seja, o cDNA com a razão dos sinais superior ao fator de corte seria considerado como candidato a ser diferencialmente expressos no sistema analisado.

Após a análise, as amostras selecionadas foram seqüenciadas conforme descrito no item 3.5 e, em seguida, as seqüências obtidas foram comparadas com aquelas depositadas no *GeneBank*, utilizando-se o programa *BLAST* (<http://www.ncbi.nlm.nih.gov/BLAST>).

3.7. Membranas comerciais

3.7.1. Hibridização da membrana

As sondas radioativas foram preparadas da mesma maneira descrita anteriormente (item 3.6.6). As membranas foram incubadas por 1h a 68°C na solução de pré-hibridização *Rapid-hyb buffer* (Amersham Biosciences) complementada com DNA de esperma de salmão desnaturado (100µg/ml). A seguir, a sonda foi adicionada à amostra e estas hibridizadas por 16h a 68°C.

Após a hibridização, as membranas foram lavadas por 30min em solução SSC 2x e SDS 1%, 30min em solução SSC 0,1x e SDS 0,5% e 5min em SSC 2x, sendo todas as lavagens a 68°C.

As membranas foram expostas a uma tela *PhosphorImager* por um período de 2-20 dias (dependendo da qualidade da marcação) e, em seguida, lidas em um scanner laser de alta definição (*Storm* – Molecular Dynamics).

A análise posterior das membranas foi a mesma realizada com os *macroarrays* provenientes do RDA (item 3.6.6).

3.8. Confirmação da expressão diferencial por RT-PCR

3.8.1. Padronizações

Para confirmação dos genes candidatos a serem regulados por MT, através da metodologia de RT-PCR, foram desenhados *primers* para os genes MN1, UNR, Ferritina e GAPDH, com o auxílio do programa *Primer3* (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi).

Inicialmente, foi realizada uma reação de PCR, utilizando-se GAPDH (Gliceraldeído 3-fosfato desidrogenase) como controle, com a finalidade de padronizar o número de ciclos onde a amplificação de GAPDH estivesse na sua fase exponencial, visto que se a reação de PCR chegar ao platô de amplificação não é possível analisar as diferenças de expressão. Foi utilizado como template cDNAs provenientes das linhagens PLJ e MTWT e um programa de 94°C por 2 minutos, 35 ciclos de 94°C por 30 segundos, 56°C por 30 segundos e 72°C por 45 segundos e, após os ciclos, 72°C por 6 minutos. Foi recolhida uma alíquota das amostras nos ciclos 10, 15, 20, 25 e 30 para sua posterior análise em gel de agarose 1%. Em seguida, padronizou-se a quantidade de cDNA que seria utilizada em cada reação de RT-PCR de modo que, após a amplificação, observássemos no gel de agarose o mesmo sinal de amplificação do gene GAPDH para todos os cDNAs.

Após a reação, as amostras foram corridas em um gel de agarose 1% corado com Brometo de Etídeo. Três abordagens foram utilizadas para a análise dos géis. Na primeira, a imagem era fotografada, a fotografia escaneada e a densitometria das bandas realizada através do programa *ImageQuant* (Molecular Dynamics). Na segunda abordagem, a imagem do gel era escaneada diretamente pelo aparelho *ImageMaster VDS* (Amersham Biosciences) e essa imagem era analisada pelo programa *ImageQuant*. Em uma última abordagem, que no final se mostrou a mais fiel aos dados observados nos géis, a imagem era escaneada pelo *ImageMaster* e as bandas densitometradas pelo programa do aparelho (*ImageMaster VDS software*).

Baseando-se nos valores obtidos na densitometria, pode-se estimar a quantidade de DNA que seria necessária para obtermos o mesmo sinal em todas as bandas. Assim, as quantidades de *template* foram alteradas, uma nova reação foi realizada e esse processo se repetiu até que todas as

amostras apresentassem o mesmo sinal, quando analisadas em um gel de agarose.

Para os demais genes, foi realizada a padronização do número de ciclos onde a amplificação do gene estivesse em sua fase exponencial. Como *template* foram utilizados cDNAs provenientes das linhagens estudadas, nas quantidades definidas previamente na padronização com GAPDH.

3.8.2. Confirmação da expressão diferencial

Uma vez estipulado o número de ciclos e a quantidade de *template* que seria utilizada para amplificar o cDNA proveniente de cada mutante, foram realizadas reações de PCR utilizando *primers* desenhados para os genes selecionados como candidatos a serem regulados por MT. Assim, como descrito anteriormente, as reações foram analisadas em gel de agarose 1% corado com brometo de etídeo e, nessa etapa, as quantificações foram feitas pelo programa ImageMaster VDS software após a imagem do gel ter sido escaneada pelo aparelho ImageMaster. A diferença de expressão entre as linhagens foi analisada em função da intensidade das bandas obtidas pela quantificação do gel.

3.9. Confirmação da expressão diferencial por *Real-Time PCR*

Nos experimentos de *Real-Time PCR* utilizou-se os mesmos cDNAs sintetizados para os experimentos de RT-PCR (3.4.1). Todas as reações foram realizadas em duplicata, contendo 3µl do cDNA da amostra analisada, 3µl de *primer*, na concentração definida no item 3.9.1, e 6µl do *SYBER Green Dye*. As reações de amplificação e quantificação dos produtos formados foram realizadas no termociclador *ABI 5700* (Roche), do nosso Departamento de Bioquímica, que trabalha com placas de 96wells. Todas reações foram montadas tomando-se precauções para que a placa não ficasse suja com pó, talco ou outras partículas que pudessem interferir na leitura da fluorescência.

3.9.1. Padronização da concentração dos *primers*

Para a quantificação do produto formado durante a reação de *Real-Time PCR* foi utilizado o reagente *SYBR® Green Dye* (Applied Biosystems). O *SYBER Green dye* possui afinidade pela *minor groove* da dupla fita de DNA.

Quando não está ligado ao DNA, emite uma pequena fluorescência no comprimento de onda de 520nm, entretanto, quando se liga à dupla fita de DNA a fluorescência é aumentada cerca de 100 vezes, permitindo então a detecção do produto do PCR em tempo real (Lekanne Deprez, Fijnvandraat *et al.*, 2002).

Para padronização da concentração de *primer* que seria utilizada na amplificação dos genes selecionados, foi utilizado como *template* um *pool* de cDNAs provenientes de todas as linhagens celulares estudadas. Para cada par de *primers* foram testadas as concentrações de 200nM, 600nM e 800nM. Assim, em cada reação foi adicionado 5µl do *pool* de cDNA, 5µl de *primers* na concentração desejada e 10µl do *SYBER Green Dye*. Todas as reações foram realizadas em duplicata. Para esta, e todas as amplificações que virão a seguir, foi utilizado o programa de 1 ciclo de 10min a 95°C e 40 ciclos de 30 segundos a 95°C, 1 minuto a 60°C e 1 minuto a 72°C.

O gerenciamento do termociclador e a coleta dos dados gerados durante a amplificação foi realizado pelo programa *ABI-Manager SDS* (Roche). Utilizando o mesmo programa, foi escolhida a concentração de *primers* que resultou no menor nível de formação de *primer-dimer*, ou inexistente, e a menor variação entre as duplicatas.

3.9.2. Padronização da eficiência dos *primers*

Para calcular a eficiência do *primer* (E), foram realizadas várias reações de *Real-Time PCR* contendo diluições seriadas do *template*. Posteriormente, foi elaborado o gráfico dos Cts obtidos em função do logaritmo da respectiva diluição, sendo que a eficiência do *primer* foi dada de acordo com o coeficiente angular da reta obtida, conforme a formula abaixo:

$$E = 10^{-\left(\frac{1}{\text{coeficiente angular}}\right)}$$

Para cada reação foram adicionados 3µl do *pool* de cDNA nas diluições desejadas (1:40, 1:80, 1:160, 1:320, 1:640), 3µl de *primers* na concentração desejada e 6µl do *SYBER Green Dye*. Todas as reações foram realizadas em duplicata e utilizando um programa de 1 ciclo de 10 minutos a 95°C e 40 ciclos de 30 segundos a 95°C, 1 minuto a 60°C e 1 minuto a 72°C. O

gerenciamento do termociclador e a coleta dos dados gerados durante a amplificação foram realizados pelo programa *ABI-Manager SDS* (Roche).

3.9.3. Confirmação da expressão diferencial

Terminada a reação de *Real-Time PCR*, através do programa *ABI-Manager SDS* foi definido manualmente um *threshold* que estivesse na fase exponencial de amplificação do gene, sendo que a intersecção da curva de amplificação com o *threshold* forneceu o *Ct* (*Threshold cycle* - o ciclo onde a fluorescência se encontra estatisticamente acima do *background*) da amostra analisada (Figura 8-A).

Figura 8 - Saída gráfica do programa ABI-Manager SDS

A) Tela de saída do programa ABI-Manager SDS fornecendo a fluorescência detectada em função do número de ciclos. No exemplo mostrado, para um *threshold* de 0,4 é obtido um *Ct* de 19,6. B) Representação do ΔCt , que é dada pela diferença entre o *Ct* das duas curvas que estão sendo comparadas.

Foi utilizado o método de quantificação relativa para analisar os dados. Assim, para cada cDNA analisado eram realizadas duas reações, sendo uma utilizando *primers* para o gene analisado e a outra utilizando *primers* para um controle interno, e a diferença entre os *Cts* dos dois genes foi definida como ΔCt (Figura 8-B)

Teoricamente, em uma reação de PCR a quantidade de produto formado é duplicada a cada ciclo, portanto, a diferença entre uma amostra analisada e uma amostra de referência é dada pelo ganho em cada ciclo (2 vezes) elevado a diferença de ΔCt das duas amostras, ou seja,

$2^{\Delta C_{t_{amostra\ de\ referência}} - \Delta C_{t_{amostra\ analisada}}}$ (Pfaffl, 2001). Entretanto, na prática a eficiência dos *primers* não é sempre 100%, ou seja, ocorre um ganho menor que 2x entre cada ciclo de PCR, sendo que o valor da eficiência do *primer* (item 3.9.2) deve ser considerado.

Para acrescentar a eficiência dos *primers* nos cálculos realizados, inicialmente a fórmula inicial foi desmembrada:

$$ratio_{relativo\ a\ amostra\ de\ referência} = 2^{-(\Delta C_{t_{amostra\ analisada}} - \Delta C_{t_{amostra\ de\ referência}})}$$

$$ratio_{relativo\ a\ amostra\ de\ referência} = \frac{2^{-\Delta C_{t_{amostra\ analisada}}}}{2^{-\Delta C_{t_{amostra\ de\ referência}}}}$$

Em seguida, substituiu-se 2 pela eficiência do *primer* da amostra de referência (E_r) e pela eficiência do *primer* da amostra estudada (E_a), respectivamente, sendo assumido que $\Delta CP = -\Delta CT$. Portanto, a fórmula final para o cálculo da diferença de expressão foi dada por:

$$ratio_{relativo\ a\ amostra\ de\ referência} = \frac{E_a^{\Delta CP_{amostra\ analisada}}}{E_r^{\Delta CP_{amostra\ de\ referência}}}$$

3.9.4. Análise estatística dos dados

Os resultados obtidos foram submetidos a uma análise estatística com auxílio do programa *Prism v4.02* (GraphPad). Para análise da diferença de expressão de cada gene, entre as diversas linhagens analisadas, foi inicialmente realizado um teste de D'Agostino e Pearson para constatar que se tratava de uma distribuição normal. Em se tratando de uma distribuição normal, para analisar a significância estatística entre as diferenças de expressão encontradas, foi utilizada uma análise de variância unidirecional (*one-way ANOVA*) seguida de um teste de Tukey, que realiza uma comparação dois a dois das médias dos valores de níveis de expressão. Para identificar possíveis relações entre os níveis de expressão dos genes estudados foi utilizada uma análise de correlação de Pearson.

3.10. CodeLink bioarrays

3.10.1. Preparação da sonda

3.10.1.1. Purificação do RNA

O RNA total foi extraído das linhagens celulares através da centrifugação dos lisados celulares em colchão de cloreto de cézio (item 3.3). Seguindo recomendações do fabricante dos *Bioarrays* (Amersham Biosciences), o RNA foi novamente purificado através do kit *RNeasy* (Qiagen). O kit se baseia em uma membrana de sílica-gel que, mediante alta concentração de sal, se liga a moléculas de RNA maiores que 200bp. Uma vez que o RNA se encontra ligado à membrana, basta uma eluição com H₂O-MilliQ para que este seja recuperado.

Assim, para a purificação do RNA, foram adicionados 350µl de tampão de lise RLT (fornecido pelo kit) a 4µg de RNA total e, após homogeneização, adicionou-se 1 volume de etanol 70%. A Coluna *RNeasy* foi colocada em um tubo coletor e então a amostra foi adicionada à coluna. Após centrifugação (8.000g por 15 segundos), o tubo coletor foi trocado, adicionou-se 700µl do tampão de lavagem RW1 (fornecido pelo kit) e a coluna foi centrifugada novamente (15 segundos a 8.00g). O tubo coletor foi trocado novamente, adicionou-se o tampão RPE (fornecido pelo kit) e a amostra foi centrifugada (8.000g por 15s). Por fim, foi realizada uma nova troca do tubo coletor, adicionou-se 500µl de tampão RPE e a amostra foi centrifugada (8.000g por 2min). Para eluir a amostra, a coluna foi colocada em um tubo limpo de 1,5µl, adicionou-se 30µl de H₂O-MilliQ e a coluna foi centrifugada por 1 minuto a 8.000g. As amostras foram quantificadas em um espectrofotômetro utilizando comprimento de onda $\lambda=260\text{nm}$.

3.10.1.2. Síntese do cDNA

Para a síntese da primeira fita do cDNA, adicionou-se à reação diferentes concentrações de mRNAs de genes bacterianos (concentração final de 1:200.000 dos genes *araB* e *entF*, 1:600.000 dos genes *fixB* e *gnD* e 1:200.000 dos genes *hisB* e *leuB*) que consistem em controles da marcação das sondas, hibridização das lâminas e sensibilidade do experimento. Para cada sonda foi montada uma reação contendo 2µg de RNA total, 2µl do mRNA

de controle (bacteriano), 1µl de T7oligo(dT) *primer* e H₂O-MilliQ para um volume final de 12µl. As amostras foram colocadas a 70°C por 10min para desnaturar possíveis estruturas secundárias dos RNAs e, logo em seguida, colocadas no gelo. Adicionou-se 2µl do tampão de primeira fita, 4µl de dNTP 5mM, 1µl de inibidor de RNase e 1µl de transcriptase reversa. Para a síntese da primeira fita, as amostras foram incubadas a 42°C por 2h.

Terminada a reação de síntese da primeira fita do cDNA, montou-se a reação para síntese da segunda fita. Foram adicionados a reação 63µl de H₂O-MilliQ, 10µl de tampão para segunda fita (10x), 4µl de um *mix* de dNTP 5mM, 2µl do *mix* de DNA polimerase e 1µl de RNaseH. O tubo foi gentilmente agitado e incubado por 2h a 16°C.

3.10.1.3. Purificação do cDNA dupla-fita

Foi utilizado o kit *QIAquick Spin* (Qiagen) para purificação do cDNA dupla-fita, que assim como o kit *RNeasy* (descrito anteriormente), e baseado em uma membrana de sílica-gel que se liga ao DNA quando em alta concentração de sal.

Adicionou-se 500µl do tampão de ligação PB (fornecido pelo kit) ao cDNA sintetizado anteriormente. A coluna foi colocada em um tubo de coleta, a amostra foi transferida para a coluna e então a coluna foi centrifugada a 10.000g por 45s. Para lavar a membrana, trocou-se o tubo de coleta, sendo adicionado 700µl do tampão PE (fornecido pelo kit) e a coluna centrifugada a 10.000g. Após uma nova centrifugação a 10.000g com um novo tubo de coleta, a coluna foi transferida para um tubo limpo de 1,5ml, adicionou-se 30µl do tampão de eluição EB (fornecido pelo kit) e a coluna foi centrifugada por 1 minuto a 10.000g. Adicionou-se mais 30µl de tampão de eluição e a coluna foi centrifugada novamente.

Após a eluição, a amostra foi concentrada em um SpeedVac por aproximadamente 1 hora a 40°C.

3.10.1.4. Síntese do cRNA

Os pellets do item anterior foram ressuspensos em 13,5µl de uma solução contendo 9,5µl de H₂O-MilliQ e 4,0µl de tampão de reação T7. Assim que os pellets foram ressuspensos, adicionou-se a cada amostra 4µl de solução T7 ATP, 4µl de T7 GTP, 4µl de T7 CTP, 4µl de T7 UTP, 7,5µl de

biotina-11-UTP (10mM) e 4 μ l de *m*ix da enzima T7 (10x). Para a síntese pela enzima T7 a amostra foi incubada por 14h a 37°C.

3.10.1.5. Purificação do cRNA marcado

O volume da solução de cRNA foi ajustado para 100 μ l através da adição de H₂O-MilliQ. Adicionou-se 350 μ l do tampão RLT (fornecido pelo fabricante) e, após homogeneização, foi adicionado 250 μ l de etanol 100%. As amostras foram então purificadas em colunas *RNeasy* (Qiagen) utilizando um protocolo semelhante àquele descrito em 3.10.1.1. Assim, a amostra foi aplicada em uma coluna *RNeasy* em um tubo de coleta. Centrifugou-se a coluna a 8.000g por 15 segundos, o tubo de coleta foi alterado e então foi adicionado 500 μ l de tampão RPE (fornecido pelo kit). A coluna foi centrifugada a 8.000g por 15 segundos e, após trocarmos o tubo de coleta, a coluna foi centrifugada por mais 2min a 8.000g, para secar a membrana. Transferiu-se a coluna para um tubo limpo de 1,5ml e pipetou-se 50 μ l de H₂O-MilliQ sobre a membrana. Após uma incubação de 10min à temperatura ambiente, o cRNA foi eluído por centrifugação de 1 minuto a 8.000g. Esse procedimento de eluição foi repetido mais uma vez.

A concentração do cRNA foi determinada posteriormente em um espectrofotômetro medindo-se a absorbância a $\lambda=260\text{nm}$ e a sua pureza foi determinada pela razão das medidas da absorbância em $\lambda=260\text{nm}$ e $\lambda=280\text{nm}$.

3.10.2. Hibridização e lavagem

3.10.2.1. Fragmentação do cRNA

Para cada uma das amostras, 10 μ g de cRNA, preparados conforme descrito no item anterior, foram ajustados para um volume de 10 μ l através da adição de H₂O-MilliQ. À amostra foram adicionados 5 μ l de tampão de fragmentação 5x (fornecido pelo fabricante do *CodeLink*). Para que ocorresse a fragmentação do cRNA, as amostras foram colocadas em um termociclador para ficarem por 20min a 94°C e, em seguida, 5min a 0°C.

3.10.2.2. Hibridização dos *bioarrays*

Para cada *bioarray* hibridizado, preparou-se uma mistura contendo 260µl de uma solução de hibridização contendo 78µl do componente do tampão A (fornecido pelo fabricante), 130µl do componente do tampão B (fornecido pelo fabricante) e 27µl de H₂O-MilliQ, à qual adicionou-se 25µl do cRNA fragmentado. As soluções foram misturadas por 5 segundos e então incubadas a 90°C por 5min para desnaturação do cRNA. Até o momento de serem aplicadas nos *bioarrays*, as soluções de hibridização foram mantidas no gelo.

Em cada *bioarray*, foram aplicados 250µl da solução de hibridização utilizando-se uma micropipeta de 1ml conforme a Figura 9. As câmaras de hibridização foram então seladas com os adesivos fornecidos pelo fabricante. Os *bioarrays* foram colocados em um agitador por cerca de 20h, e mantidos a 37°C e 300rpm.

Figura 9 - Aplicação da solução de hibridização do *bioarray*

3.10.2.3. Lavagem e marcação

Após a hibridização, o adesivo, que formava a câmara de hibridização, foi removido e o *bioarray* colocado em um reservatório contendo uma solução de TNT 0,75x à temperatura ambiente. Posteriormente, foi transferido para um recipiente contendo solução de TNT 0,75x a 46°C, onde foi mantido por exatamente 1 hora.

O *bioarray* foi removido da solução de TNT 0,75x e colocado em um outro reservatório contendo uma solução de Cy5-Estreptavidina. Para evitar a degradação do fluoróforo, o reservatório foi coberto e os *bioarrays* foram mantidos nessa solução por 30min. Em seguida, o *bioarray* foi colocado em um novo reservatório contendo o tampão TNT 1x, onde foi mantido por 5min, também protegido da luz. Esse procedimento foi realizado mais 3vezes, sempre utilizando uma nova solução de TNT 1x a cada lavagem.

Terminada a quarta lavagem, o *bioarray* foi colocado em um reservatório contendo uma solução Tween20 0,05%, na qual o *bioarray* foi movido para cima e para baixo 5x e o procedimento foi repetido novamente com uma nova solução Tween20 0,05%. Por fim, os *bioarrays* foram secos através de centrifugação a 644g por 3min e então guardados em um recipiente apropriado para mantê-los protegidos da luz até que fosse escaneados.

3.10.2.4. Aquisição e análise dos dados

Os *bioarrays* foram escaneados pelo *GenePix 4000B Array Scanner* (Axon Instruments) e os dados relativos às intensidades de fluorescência emitida pelos *spots* coletadas pelo programa *CodeLink System Software* (Amersham Biosciences). Após a coleta dos dados, o programa invalidou automaticamente os dados referentes aos *spots* que haviam apresentado alguma irregularidade e forneceu os valores da intensidade de cada *spot* antes e após uma normalização baseada na mediana local da medida das intensidades.

Para a análise dos dados não normalizados, foi desenvolvido um algoritmo na linguagem R de programação (<http://www.r-project.org/>), com o auxílio das bibliotecas SMA (*Statistics for Microarray Analysis* - <http://www.stat.berkeley.edu/users/terry/zarray/Software/smacode.html>) e LIMMA (*Linear Models for Microarray Data* - <http://bioinf.wehi.edu.au/limma/>). Os dados foram normalizados utilizando-se duas metodologias distintas: a *Quantile*, onde todos os *arrays* são normalizados de modo que apresentem os mesmos quartis (Bolstad, Irizarry et al., 2003) , e o LOWESS (Locally-weighted scatter plot smoother) no qual, de acordo com um índice de alisamento pré-determinado, os dados

são normalizados em relação a uma curva de regressão robusta localmente ponderada (Cleveland, 1979).

Em seguida, para a seleção dos genes cujo nível de expressão entre os arrays era estatisticamente diferente, foi desenvolvido um algoritmo baseado em um método empírico de Bayes (e-Bayes), que consiste em um refinamento do teste-t clássico onde se procura evitar a ocorrência de falsos positivos e falsos negativos durante a análise de dados provenientes de arrays (Lonnstedt e Speed, 2002).

A lista dos genes selecionados foi submetida ao programa PubGene (<http://www.PubGene.org/>). Neste programa, para identificar possíveis interações entre os genes, é atribuída uma nota a cada par de genes que varia de 0 a 1, sendo 0 equivalente a 100% de chance de existir uma interação entre os genes e 1 a ausência de interação entre eles. Para o cálculo dessa nota é utilizado como base o número de publicações que os genes analisados possuem em comum. A partir das interações identificadas pelo PubGene foram desenhados *primers* para uma futura análise por *Real-Time PCR*.

3.11. Amplificação do cDNA completo do gene MN1

Para amplificar o transcrito de MN1 a partir do RNA total das células PLJ foram utilizadas três metodologias: 1) amplificação utilizando *primers* com seqüências baseadas em predições de ferramentas de Bioinformática e/ou ESTs (Expressed Sequence Tags) conhecidas que alinhavam com a região cromossomal onde se encontra o MN1; 2) *RACE (Rapid Amplification of cDNA Ends)*; 3) Metodologia *RACE* utilizando os próprios *primers* do *RACE* juntamente com os *primers* sintetizados para a seqüência predita.

3.11.1. Amplificação a partir de *primers* baseados em ESTs

O transcrito de MN1 foi amplificado a partir do cDNA de PLJ conforme descrito no item 3.4.1. Os *primers* foram desenhados tomando como base seqüências preditas por análise de Bioinformática e ESTs que alinham com a região cromossômica onde possivelmente se encontra o gene MN1 de camundongo. Para identificar as seqüências preditas por ferramentas de Bioinformática foi utilizado o programa *BLAST*

(<http://www.ncbi.nlm.nih.gov/Blast>), enquanto que para identificar a região cromossomal onde estaria o gene MN1 bem como as ESTs já descritas que alinham com essa região foi utilizado o programa BLAT (<http://genome.ucsc.edu/cgi-bin/hgBlat>).

3.11.2. Amplificação utilizando a metodologia SMART RACE

3.11.2.1. Síntese da 1ª fita do RACE

Para amplificar o transcrito de MN1 a partir do RNA total de células PLJ foi utilizada a metodologia de RACE (*Rapid Amplification of cDNA Ends*) adaptada em relação àquela descrita no SMART cDNA synthesis Kit (Amersham Biosciences - Figura 10).

Figura 10 - Representação esquemática do RACE-5' e RACE-3'

O RNA utilizado foi extraído e purificado conforme descrito no item 3.3. Foram preparadas duas reações, 5'-RACE e 3'-RACE. Na 5'-RACE foi utilizado

1µg de RNA total proveniente da linhagem PLJ, 0,7µl de Oligo dT 500ng/ml (Amersham Biosciences), 0,7µl de SMART II-A Oligo-Sal 3A1 20µM (Amersham Biosciences), e H₂O para um volume final de 5µl. Na 3'-RACE foi utilizado também 1µg de RNA total da linhagem PLJ, 0,7µl de 3' SMART CDS primer IIA-Sal3A1 20µM (Amersham Biosciences) e H₂O para um volume final de 5µl.

As amostras foram incubadas a 70°C por 10min e adicionou-se, a cada reação, 5µl de uma solução contendo 2µl de Buffer 5x para enzima Super Script (Invitrogen), 0,2µl de DTT (Ditiotrietol), 1µl de dNTP 10mM, 1µl de enzima *SuperScript* 200U/µl (Invitrogen) e H₂O para completar 5µl. Para amplificação dos transcritos as amostras foram colocadas a 45°C por 35min, 55°C por 50min e 42°C por 60min. A enzima foi inativada através de uma incubação a 72°C por 10min e, em seguida, as amostras foram diluídas em 40µl de TE. O controle da qualidade dos cDNAs foi realizado conforme descrito em 3.4.2.

3.11.2.2. Amplificação dos transcritos

Nas reações de amplificação foram utilizados, além do *primer* do RACE (*Smart PCR primer*), os *primers* sintetizados para realizar a amplificação de MN1 no *Real-Time PCR* (Real-MN1-F/Real-MN1-R) e os *primers* descritos no item 3.11.1.

Cada reação de amplificação foi montada de forma a apresentar uma concentração final de 1x para o Tampão de Reação (Fermentas), 1,5mM de MgCl₂, 0,2mM de dNTPs, 0,2µM do SMART PCR *primer*, 0,2µM do segundo *primer* (*primers Forward* para 3'-RACE e *primers Reverse* para 5'-RACE), 0,5U da enzima Taq (Fermentas) e Template.

Como não eram conhecidas as temperaturas ideais para a amplificação dos transcritos, foram realizados diversos testes variando o número de ciclos, a temperatura de anelamento do *primer* (que foi a mesma durante toda a amplificação), a temperatura de anelamento ao longo dos ciclos (que diminuía em média 1°C/ciclo), a temperatura de anelamento em cada ciclo (que diminuía em média 0,2°C/segundo) e, por fim, o tempo de extensão do *primer*. Portanto, a amplificação consistia em 94°C por 2 minutos, número de ciclos variável, temperatura e tempo e de anelamento conforme descrito acima, 72°C por diversos tempos e, terminado os ciclos, 72°C por 6 minutos.

As reações de RACE foram posteriormente analisadas em gel de agarose 1% contendo Brometo de Etídeo.

3.11.2.3. Purificação, clonagem e transformação

Os produtos das reações de RACE foram purificados em colunas GFX do kit *GFX PCR cDNA and Gel Band Purification Kit* (Amersham Biosciences). As amostras foram então concentradas no Speed Vac até um volume final de 25µl. A cada reação, foi acrescentada a seguinte mistura para reparo e fosforilação das pontas dos cDNAs: 1,0µl da enzima *Klenow* (Fermentas), 1,5µl da enzima *PNK Polynucleotide Kinase*, 3,5µl do tampão Tango Y+ 10x, 3,5µl de APT (A) 10mM e 1,0µl de dNTP 10mM. As amostras foram incubadas a 37°C por 45min e, em seguida, a 65°C por 20min.

Após uma segunda purificação em colunas GFX, as amostras foram novamente concentradas no Speed Vac para um volume final de 5µl e acrescidas de 1µl de preparação plasmideal do vetor pUC18, 1µl ligase, 1µl de tampão de ligase e H₂O para um volume final de 2µl. As amostras foram incubadas a 16°C por um período de 12h.

Os plasmídeos foram usados para transformar bactérias *E. coli* DH10B por eletroporação. Após 2h, sob agitação a 37°C, as bactérias transformadas foram plaqueadas em placas de Petri contendo LB-Agar e ampicilina 30µg/ml.

3.12. Inibição da expressão do gene MN1

Para silenciamento do gene MN1 foi utilizada a metodologia de RNA de interferência (RNAi), que se baseia no fenômeno de que RNAs dupla fita podem atuar como silenciadores de genes ao nível pós-transcricional.

Observado inicialmente em *Caenorhabditis elegans*, o mecanismo se baseia no princípio de que quando um dsRNA (RNA dupla fita) é introduzido em uma célula, este RNA é processado em pedaços menores (20-25 nucleotídeos) por uma RNase denominada Dicer, dando origem aos siRNAs (*small interfering RNAs*). Em seguida, as duas fitas dos siRNAs são separadas pela nuclease RISC (*RNA-induced silencing complex*), que guia uma das simples fitas resultantes para se anelar com a sequência de mRNA complementar. Após o pareamento, a DICER cliva o mRNA com o qual houve o pareamento, silenciando assim sua atividade (Hannon, 2002).

3.12.1. Preparação de DNA plasmideal em média escala

Em 150 ml de meio LB contendo 100µg/ml de ampicilina, foram inoculadas as bactérias recombinantes contendo as construções pEGFP e pMN1. Tais construções constituem vetores que, quando transfectados em células empacotadoras 293T, produzem lentivírus que codificam para a proteína EGFP e para proteína EGFP mais a sequência dsRNA homóloga à MN1 (pMN1), sendo que as construções foram geradas em colaboração com o Dr. Ricardo G. Correa, que na época se encontrava no *Salk Institute for Biological Studies* (Tiscornia, Singer *et al.*, 2003). A preparação de cDNA plasmideal foi obtida utilizando-se o Kit *Qiagen Plasmid Midi*, conforme instruções do fabricante.

3.12.2. Transfecção para produção de partículas virais

Células 293T foram plaqueadas em placas p35 (10^7 células/placa) na presença de DMEM contendo 10% SFB (HyClone) sem antibióticos, 24h antes da transfecção. O DNA plasmideal (5µg) foi diluído em 250µl de DMEM (10% SBF e sem antibiótico) e, em outro tubo eppendorf, 10µl do reagente de lipofecção (Lipofectamine 2000 – Invitrogen) foi diluído em 250µl de meio de cultura. O conteúdo dos dois tubos foi misturado e, após incubação de 30min à temperatura ambiente, foi adicionado ao meio de cultura das células.

3.12.3. Coleta e concentração dos sobrenadantes virais

Após 24h da transfecção das células 293T, foi efetuada a troca do meio de cultura para início das coletas de 1ml de meio DMEM 10%SFB a cada 24h, até 72h após a transfecção. As amostras coletadas foram centrifugadas a 1.000g por 5 min para eliminação de debris celulares, sendo então congeladas a -70°C .

Antes da infecção, os sobrenadantes contendo as partículas virais foram descongelados e centrifugados a 20.000g por 4h a 4°C . O sobrenadante foi descartado e as partículas virais ressuspensas em 450µl de meio DMEM 10%SFB contendo 8µg/ml de polibreno.

3.12.4. Infecção com lentivírus

As células PLJ foram plaqueadas em placas de 12 poços (3×10^4 células/poço) e incubadas por 12h em meio DMEM 10% SFB. Para que

a infecção ocorresse, o meio de cultura foi substituído pelo sobrenadante viral concentrado.

Antes da infecção, os sobrenadantes contendo as partículas virais foram descongelados e centrifugados a 20.000g por 4 horas à 4°C. O sobrenadante foi descartado e as partículas virais foram ressuspensas em 450µl de meio DMEM 10%SFB contendo 8µg/ml de polibreno. Em seguida, o meio de cultura das células foi substituído pelo sobrenadante viral concentrado.

3.12.5. Confirmação do RNAi

Como o vetor utilizado para expressão do RNAi de MN1 co-expressava GFP, a eficiência de infecção foi realizada através da observação das linhagens infectadas em luz UV.

Para confirmação da inibição de MN1 pelo seu respectivo RNAi foi extraído RNA das linhagens infectadas e sintetizado seu respectivo cDNA. Posteriormente, a confirmação foi realizada através de uma reação de *Real-Time PCR* utilizando-se *primers* para MN1.

3.12.6. Curva de crescimento

Para estudar o efeito da inibição de MN1, 10⁴ células das linhagens PLJ, PLJ transfectada com vetor vazio e PLJ transfectada com vetor contendo RNAi para MN1 foram plaqueadas em placas descartáveis de 24wells. As placas foram mantidas em condições adequadas de cultura sendo o meio renovado a cada dois dias. Cada ponto da curva foi coletado através de tripsinização e, em seguida, as células foram fixadas em solução 3,7% formaldeído-PBSA e mantidas a 4°C até o momento da contagem. A contagem das suspensões de células foi realizada em contador eletrônico modelo CC530 (CELM-Companhia de Equipadora de Laboratórios Modernos).

3.12.7. Formação de colônia em meio semi-sólido

Poços de placas de 24 poços foram preenchidos com 500µl de agarose 0,6% em meio de DMEM contendo 10% SFB. Sobre esta camada, foram plaqueados 100µl de meio de cultura contendo 10³ ou 10⁴ células em meio de cultura, em duplicata. Rapidamente, para evitar a adesão das células, foram adicionados 500µl de solução semi-sólida de agarose 0,3% ou 0,9% em meio

DMEM contendo 10% SFB. Após 24h, foram adicionados 500µl de meio DMEM 10%SBF sendo que esta camada superior de meio líquido foi trocada a cada 3 dias. Vinte e um dias após o plaqueamento as colônias foram fixadas com a adição de 100µl de formaldeído 37% e fotografadas em microscópio invertido (Nikon Diaphot, modelo Eclipse TE 300).

3.13. Amplificação do cDNA completo do gene Dlk1

A sequência de Dlk1 foi obtida através dos bancos de dados do NCBI (National Center for Biotechnology Information - <http://www.ncbi.nlm.nih.gov>) e, com base nessa sequência, foi identificada a ORF (*Open Reading Frame*) do gene através do programa *ORF Finder* (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>). Foram desenhados *primers* que amplificavam a ORF e continham adaptadores com sítios para a enzima *EcoRI*.

Cada reação foi montada de forma que apresentasse uma concentração final de 1x para o Tampão de Reação (Fermentas), 1,5mM de MgCl₂, 0,2mM de dNTP, 0,2µM de cada um dos *primers* (*Forward* e *Reverse*), enzima Taq (Fermentas) e cDNA proveniente da linhagem PLJ como *template*. O programa utilizado foi de 94°C por 2min, 39 ciclos de 94°C por 30seg, 56°C por 30seg e 72°C por 1min e, após os ciclos, 72°C por 6min.

3.13.1. Ligação do produto amplificado ao vetor

Para prevenir uma possível interferência entre o promotor do vetor PLJ e o promotor de vetor onde seria clonado o gene Dlk1, a clonagem foi realizada em três vetores distintos, pcDNA3 (Invitrogen) e pCXN2 (Niwa, Yamamura *et al.*, 1991) e o vetor retroviral pLPCX (Clontech).

O produto da amplificação foi purificado em colunas do kit *GFX PCR cDNA and Gel Band Purification Kit* (Amersham) conforme recomendações do fabricante. As amostras foram purificadas e os vetores selecionados forem digeridos com a enzima *EcoRI* por 2h a 37°C.

Para evitar o auto-anelamento dos vetores durante a etapa de ligação, após uma nova purificação utilizando o kit GFX as pontas dos vetores foram

desfosforiladas utilizando a fosfatase CIAP (*Calf Intestine Alkaline Phosphatase*) durante 1h a 37°C.

O vetor e os transcritos amplificados foram purificados utilizando o kit GFX e, após quantificação medindo-se a absorbância a 260nm, foi realizada a ligação dos transcritos aos vetores em uma reação contendo 0,1µg do vetor, 0,3µg do inserto, 1µl de ATP, 1µl de Tampão, 1µl de Ligase a 37°C durante 16h.

3.13.2. Transformação por choque térmico

Bactérias *Escherichia coli* DH10B foram crescidas em meio LB a 37°C por 12h, sob agitação. A pré-cultura foi diluída 1:10 e incubada a 37°C, sob agitação, até atingir uma absorbância a 600nm (luz visível) de 0,4 a 0,7. A cultura de bactérias foi transferida para o gelo e, em seguida, submetida a uma centrifugação de 4.000rpm por 5min. O sedimento de bactérias foi ressuspenso em tampão de competência gelado (acetato de sódio 20mM pH 6,5 adicionado de CaCl₂ 0,1M).

Após incubação no gelo por 20min, as bactérias foram coletadas por uma centrifugação a 2.000rpm por 5min. O sedimento foi ressuspenso novamente em tampão de competência gelado e mantido no gelo por 40min. Foi adicionado 0,5µg do produto da ligação a 100µl de bactéria. Após uma nova incubação no gelo por 20min, foi realizado o choque térmico a 42°C por 2min. As bactérias transformadas foram incubadas em meio LB a 37°C por 1h, sob agitação, e então plaqueadas em placas contendo LB Agar-ampicilina (50µg/ml). Após incubação por 12h, as colônias transformantes foram isoladas em uma placa de 96poços contendo LB-Ampicilina (50µg/ml). As placas foram estocadas a -70°C em meio contendo LB-Glicerol 15%.

3.13.3. Confirmação da clonagem

Para confirmação da clonagem foi realizada uma reação de PCR montada de forma a conter uma concentração final de 1x para o Tampão de Reação (Fermentas), 1,5mM de MgCl₂, 0,2mM de dNTP, 167nM de cada um dos *primers* (M13-Forward e M13-Reverse), enzima Taq (Fermentas) e meio LB contendo as bactérias transformadas como *template*. Para amplificação, as reações foram submetidas a um programa de 94°C por 4min, 40 ciclos de 94°C por 45seg, 54°C por 30seg e 72°C por 2min e, após os ciclos, 72°C por

5min. O resultado foi analisado em gel de agarose 1% contendo brometo de etídeo.

3.13.4. Preparação de DNA plasmideal em média escala

Em 150 ml de meio LB contendo 100µg/ml de ampicilina, foram inoculadas as bactérias recombinantes contendo as construções pLPCX e pLPCX+pMN1. Tais construções constituem vetores que, quando transfectados em células empacotadoras ϕ nx-Ampho, produzem retrovírus capazes de infectarem células com o vetor pLPCX somente (pLPCX) ou com o vetor pLPCX com a seqüência que codifica para o RNAi de MN1 clonada (RNAi-MN1). A preparação de cDNA plasmideal foi obtida utilizando o Kit Qiagen Plasmid Midi, conforme instruções do fabricante.

3.13.5. Transfecção de células empacotadoras

Células ϕ nx-Ampho foram plaqueadas em placas p60 (2×10^6 células/placa) na presença de DMEM sem antibióticos contendo 10% SFB (HyClone), 16h antes da transfecção. O DNA plasmideal (5µg) foi diluído em 250µl de DMEM (10% SBF e sem antibiótico) e, em outro tubo eppendorf, 12µl do reagente de lipofecção (Lipofectamine 2000 – Invitrogen) foi diluído em 288µl de meio de cultura. O conteúdo dos dois tubos foi misturado e, após incubação de 30 minutos à temperatura ambiente, foi adicionado ao meio de cultura das células. Após 5h foi efetuada a troca do meio para início das coletas.

3.13.6. Coleta dos sobrenadantes virais

Após 12h da transfecção das células ϕ nx-Ampho, foi efetuada a troca do meio de cultura para início das coletas de 3ml de meio DMEM 10%SFB a cada 12h, até 72h após a transfecção. As amostras coletadas foram centrifugadas à 1.000g por 5 min para eliminação de debris celulares, sendo então congeladas à -70°C.

3.13.7. Infecção com retrovírus

As células MTWT foram plaqueadas em placas de 12 poços (10^4 células/poço) e incubadas por 12h em meio DMEM 10% SFB.

Para infecção, alíquotas congeladas do sobrenadante contendo as partículas virais foram descongeladas e centrifugadas a 20.000g por 4h a 4°C. O sobrenadante foi descartado e as partículas virais ressuspensas em volume apropriado de meio DMEM 10%SFB contendo 8µg/ml de polibreno. Em seguida, o meio de cultura das células foi substituído pelo sobrenadante viral concentrado ajustado para cerca de 100 MOIs (*Multiplicity of Infection*). Após incubação por 24h, as células foram tripsinizadas e plaqueadas e após 24h foi repetido o procedimento de infecção. Após 24h da segunda infecção as células foram tripsinizadas, plaqueadas novamente e após 12h do plaqueamento iniciou-se a seleção com puromicina (1,50 µg/ml), a qual foi feita por 4 dias.

3.13.8. Curva de crescimento

Para estudar o efeito da superexpressão de Dlk1, 10⁵ células das linhagens MTWT, MTWT infectada com o vetor pLPCX vazio e pLPCX expressando o gene Dlk1 e a isoforma de Dlk1 foram plaqueadas em placas descartáveis de 6 poços. As placas foram mantidas em condições adequadas de cultura sendo o meio renovado a cada dois dias. Cada ponto da curva foi coletado através de tripsinização e, em seguida, as células foram fixadas em solução 3,7% formaldeído-PBSA e mantidas a 4°C até o momento da contagem. A contagem das suspensões de células foi realizada em contador eletrônico modelo CC530 (CELM - Companhia de Equipadora de Laboratórios Modernos).

4. RESULTADOS

4.1. Busca de genes diferencialmente expressos em celular MTWT através de membranas comerciais

Uma das estratégias utilizadas paara busca dos genes regulados por MT foi utiliar membranas *Hybond Atlas Mouse Cancer 1.2 Array* (Clontech), nas quais estão imobilizados 1.176 cDNAs cujos genes são associados ao processo tumoral. As sondas utilizadas para hibridizar as membranas foram sintetizadas a partir de mRNA de células PLJ e MTWT carenciadas para soro, SFB 0,5%, por 48h. Em seguida, uma tela de *PhosphorImager* foi exposta as membranas e, após um tempo de exposição de 14 dias, a leitura foi realizada pelo scanner *Storm*.

A

B

Figura 11 - Membranas comerciais

Imagem obtida através do escaneamento da tela do *PhosphorImager* após exposição de 14 dias. As membranas foram hibridizadas com cDNAs marcados radioativamente, provenientes de células MTWT (A) e PLJ (B). Os clones na mesma posição, em membranas diferentes, são equivalentes. Para hibridização foi utilizada a mesma quantidade de sonda marcada em ambas as membranas

4.1.1. Densitometria da membrana

Um dos problemas relacionados aos macroarranjos de cDNAs consiste na diferença de marcação entre as sondas que hibridizaram cada uma das membranas. Assim, para amenizar essa diferença, as sondas marcadas foram previamente quantificadas em um cintilador líquido, a mesma quantidade de sonda marcada foi utilizada em ambas as membranas.

Como esta abordagem não é precisa, após a hibridização e coleta dos dados de intensidade, foi realizado um tratamento matemático (normalização) nos valores da densitometria da membrana. Entre as várias abordagens possíveis para normalizar a intensidade do sinal entre as duas membranas, pode-se utilizar genes *housekeeping*, ou seja, cuja expressão é a mesma nas duas condições analisadas, ou se basear no fato de que a razão das distribuições tende a ser igual a um.

Em relação aos genes *housekeeping*, é feito um ajuste nos valores da densitometria de acordo com a diferença de sinal que eles apresentam nas duas membranas. Ainda que houvesse genes *housekeeping* na membrana utilizada, foi observado que havia uma variação entre 5% e 80% entre os sinais dos genes *housekeeping* quando as duas membranas eram comparadas. Como não era possível identificar qual dos genes *housekeeping* realmente não se alterava mediante a ação de MT, não foi possível utilizar essa abordagem. Foi utilizada então a normalização de acordo com o perfil de dispersão dos dados referentes às intensidades (Figura 12).

Figura 12 – Gráfico de dispersão do sinal das membranas comerciais

Gráfico de dispersão, em escala logarítmica, dos valores de hibridização das membranas *Hybond Atlas Mouse Cancer 1.2 Array*, com sondas provenientes das linhagens PLJ (ordenada) e MTWT (abscissa). As duas retas, paralelas à diagonal central, representam o fator de corte utilizado na membrana (*cut-off*=3). Em A está representada a distribuição antes da aplicação do fator de correção e em B após o fator de correção. Os pontos que se distanciam da reta central indicam os genes candidatos a serem diferencialmente expressos.

Como pode ser notado na Figura 12, após a normalização os pontos apresentaram uma distribuição ao redor da diagonal (Quackenbush, 2001), o que reduz a chance de falso-positivos serem selecionados para análises futuras.

O gráfico de dispersão mostrou um considerável espalhamento dos pontos em relação à diagonal principal, o que ocorreu devido a diversas manchas presentes na membrana (*background*), provavelmente por problemas ocorridos no processo de lavagem. Como esse *background* pode comprometer a análise matemática dos dados, para aumentar o grau de confiabilidade dos genes selecionados, optou-se por trabalhar com os genes que apresentavam uma maior diferença de expressão entre as duas membranas (*ratio*), sendo constatado que era possível confiar em um *ratio* maior que três (*cut-off*=3). Portanto, os pontos que estavam abaixo deste valor não foram considerados como candidatos a futuros ensaios de confirmação. A lista dos genes que foram selecionados, bem como seus nomes e classificação molecular, está descrita na Tabela 1.

Tabela 1 – Gênes selecionados nas membranas comerciais

Lista dos genes selecionados a partir da membrana *Atlas Mouse Cancer 1.2 Array Hybond* como regulados pela atividade de MT. Na tabela consta a localização dos genes na membrana, o nome do gene, sua classificação molecular e se o gene é induzido ou reprimido por MT.

Posição	Nome	Classificação	MT
A11e	<i>disintegrin and metalloproteinase domain 4 (ADAM4)</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
A11l	<i>catenin beta</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
A12d	<i>desmoglein 2</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
A12e	<i>delta-like homolog 1 (DLK1)</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
A13a	<i>epithelial membrane protein 3 (EMP3)</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
A13e	<i>CD49C antigen</i>	<i>Cell-Cell Adhesion Receptors</i>	Reprimido
B10f	<i>collagen I alpha 1 subunit (COL1A1)</i>	<i>Extracellular Matrix Proteins</i>	Reprimido
B10h	<i>procollagen III alpha 1 subunit (COL3A1)</i>	<i>Extracellular Matrix Proteins</i>	Reprimido
B10i	<i>procollagen, type IV, alpha 1</i>	<i>Extracellular Matrix Proteins</i>	Reprimido
B10l	<i>procollagen V alpha 2 subunit (COL5A2)</i>	<i>Extracellular Matrix Proteins</i>	Reprimido
B11e	<i>thrombospondin 2</i>	<i>Extracellular Matrix Proteins</i>	Reprimido
C04j	<i>H19 fetal liver mRNA</i>	<i>Turnover & Transport Proteins</i>	Reprimido
C06i	<i>fibroBlast growth factor receptor1 (FGFR1)</i>	<i>Protein Kinase Receptors</i>	Reprimido
C10h	<i>angiotensin II receptor type 2</i>	<i>G Protein-Coupled Receptors</i>	Reprimido
C14h	<i>retinoid X receptor alpha (RXR-α)</i>	<i>Intracellular Transducers</i>	Reprimido
C14i	<i>retinoid X receptor beta (RXR-β)</i>	<i>Transcription Activators/Repressors</i>	Reprimido
F05c	<i>stromal cell derived factor 3</i>	<i>Protease Inhibitors</i>	Reprimido
A05a	<i>cyclin F (S/G2/M-specific)</i>	<i>Cyclins</i>	Induzido
A06a	<i>NIMA-related protein kinase 2 (NEK2)</i>	<i>Cell Cycle-Regulating Kinases</i>	Induzido
A06b	<i>NIMA-related protein kinase 3 (NEK3)</i>	<i>Cell Cycle-Regulating Kinases</i>	Induzido
B04i	<i>transcription termination factor 1 (TTF1)</i>	<i>Oncogenes & Tumor Suppressors</i>	Induzido
F04a	<i>inter-alpha trypsin inhibitor heavy chain 3</i>	<i>Protease Inhibitors</i>	Induzido

4.2. Geração de bibliotecas subtraídas através do RDA

4.2.1. Geração das representações

Para a síntese da primeira fita, foi utilizado o kit *Smart PCR cDNA Synthesis Kit* (Clontech) que, para gerar cDNAs completos (*full-length*), se baseia no princípio de que a Transcriptase Reversa acrescenta três guaninas ao final do cDNA assim que termina de sintetizá-lo. Portanto, após a síntese da primeira fita do cDNA, um *primer* contendo três citosinas na ponta 3' é utilizado para a síntese da segunda fita.

Devido à necessidade da primeira fita ser sintetizada até o final para que exista o sítio de ligação do *primer*, que permitirá a síntese da segunda fita, em geral os transcritos longos estão pouco representados na população de cDNA, geralmente devido a falhas da transcriptase reversa ou porque estes mRNAs longos apresentam estruturas secundárias que impedem a ação da transcriptase reversa.

Procurando impedir a formação de estruturas secundárias, que consiste na principal limitação para síntese de cDNAs longos, é necessário aumentar a temperatura de hibridização ou alterar a constante dielétrica do meio. Assim, conforme previamente proposto (Aaronson e Todaro, 1968), adicionou-se trealose e betaína à amostra. A betaína altera a constante dielétrica do meio ao passo que a trealose permite que a temperatura seja aumentada sem que a transcriptase reversa desnature ou perca sua atividade.

Após a síntese da primeira fita, a segunda fita foi sintetizada e o produto foi amplificado por uma reação de PCR. Dependendo da eficiência de síntese da primeira fita e da qualidade do mRNA, o número de ciclos necessários para se trabalhar na fase de amplificação exponencial do PCR costuma variar entre 15 e 24 ciclos. Portanto, foi realizada a padronização do número de ciclos necessários para amplificação do produto (Figura 13).

Figura 13 - Padronização da amplificação do cDNA de PLJ e MTWT

Imagem do Gel de agarose contendo as amostras utilizadas para a padronização do número de ciclos (15, 18, 21, 27, 24 e 31 ciclos). Foram amplificados cDNAs provenientes das linhagens PLJ (A) e MTWT (B).

Para ambas as amostras, selecionou-se 24 ciclos para a amplificação do cDNA. Ao contrário do que era esperado, houve um decréscimo do tamanho médio dos cDNAs após adição de Trealose e Betaína. Experimentos realizados anteriormente no laboratório mostram que a amplificação pelo kit *Smart PCR cDNA Synthesis Kit* pode gerar fragmentos de até 8Kb, enquanto que na abordagem utilizada foram gerados fragmentos de até 4Kb. Entretanto, como a corrida no gel mostrou um padrão de bandejamento, um dos indicativos de um cDNA de boa qualidade, os experimentos prosseguiram.

4.2.2. Busca dos genes diferencialmente expressos

Os cDNAs resultantes da subtração *Forward* e *Reverse* foram aplicados em uma membrana de nylon, utilizando-se um aplicador manual, e hibridizadas com sondas provenientes das linhagens PLJ e MTWT. Após hibridização, uma tela de *PhosphorImager* foi exposta as membranas e, decorridos 20 dias, lidas no escâner *Storm* (Figura 14).

Figura 14 – Macroarranjos de cDNAs proveniente do RDA

Foram aplicadas 768 amostras de cDNA em duplicata, sendo 382 amostras provenientes da biblioteca *Forward* e 368 da biblioteca *Reverse*. As membranas foram hibridizadas com sondas provenientes das linhagens MTWT (A) e PLJ (B). Para hibridização foi utilizada a mesma quantidade de sonda marcada em ambas as membranas,

Metade dos cDNAs imobilizados na membrana eram provenientes da subtração *Forward* e a outra metade da *Reverse*. Apesar de ter sido colocado a mesma quantidade de sonda marcada em ambas as membranas, houve um sinal bem mais expressivo na hibridização com sondas de PLJ (A) do que com MTWT (B), e ocorreram diversas manchas nas membranas. Ainda que a correção matemática tenha amenizado estes problemas, era maior a probabilidade de encontrarmos um falso positivo proveniente da análise computacional. A representação gráfica dos valores obtidos após a normalização pode ser observada na Figura 12.

Figura 15 – Gráfico de dispersão dos macroarranjos do RDA

Representação gráfica, em escala logarítmica, da dispersão dos sinais obtidos dos macroarranjos onde foram imobilizadas as bibliotecas subtraídas. Os macroarranjos foram hibridizados com cDNA proveniente das linhagens MTWT (ordenada) e PLJ (abscissa) marcados radioativamente. As duas linhas paralelas a diagonal principal representam um *cut-off*=2.

Apesar da maior parte dos pontos terem se concentrado próximos da diagonal, foi possível identificar um número considerável de candidatos a serem diferencialmente expressos. Após as análises, 67 candidatos foram selecionados. Como um dos problemas das técnicas de subtração é o de

gerar uma alta redundância, antes da confirmação da expressão diferencial dos genes candidatos foi realizado um seqüenciamento dos cDNAs selecionados.

4.2.3. Seqüenciamento e análise dos resultados

A reação de seqüenciamento foi montada a partir do kit *BigDye® Terminator Cycle Sequencing Kit* seguindo o protocolo do fabricante. As seqüências obtidas foram inicialmente comparadas com as seqüências depositadas no GeneBank, através do programa *BLAST* (<http://www.ncbi.nlm.nih.gov/BLAST/>). Após o seqüenciamento, foram identificados oito genes como candidatos a serem regulados por MT, sendo dois ativados e seis inibidos pela ação MT (Tabela 2).

Tabela 2 - Genes identificados pelo RDA

Lista dos genes identificados como candidatos a serem regulados por MT. Na primeira coluna está a descrição do gene, na segunda e na terceira coluna os *spots* dos macroarranjos hibridizados com sondas provenientes de PLJ e MTWT, respectivamente.

Gene	PLJ	MTWT
<i>Delta-like 1 homolog (Dlk1)</i>
<i>Septine 2 (Sept2)</i>
<i>Insulin-Induced Protein 1</i>
<i>NRAS-related protein (UNR)</i>
<i>Telomerase binding protein p23</i>
<i>Meningioma 1 (MN1)</i>
<i>AP-2 transcription factor, alpha (TFAP2A)</i>
<i>Mus musculus laminin, beta 2 (Lamb2)</i>

4.3. CodeLink bioarrays

4.3.1. Marcação das sondas, hibridização e lavagens

Neste projeto foi utilizado o *CodeLink UniSet Mouse I Bioarrays*, que é constituído por 10.000 oligonucleotídeos derivados de genes depositados no GenBank, RefSeq e UniGene até o final do ano de 2002. Cada oligo possui 20 nucleotídeos, tendo sido desenhado de forma a ser específico para o gene analisado. Foram hibridizados quatro *bioarrays*, sendo dois com sondas sintetizadas a partir de mRNA da linhagem PLJ e dois a partir do mRNA proveniente da linhagem MTWT. As sondas foram preparadas seguindo fielmente as especificações da empresa.

4.3.2. Aquisição e análise dos dados

As imagens foram adquiridas utilizando-se o escâner *GenePix 4000B* e a análise inicial foi realizada pelos programas *GenePix* e *CodeLink System Software*.

Após a obtenção das imagens dos *bioarrays* (Figura 16), a intensidade de cada spot foi quantificada pelo programa *GenePix Pro*. Como descrito na Introdução, os *bioarrays* são pré-analizados para identificação de *spots* que apresentam alguma irregularidade. Assim, os sinais referentes a esses *spots* foram previamente anulados através do programa *GenePix Pro*.

PLJ

MTWT

Figura 16 – Lâminas do *CodeLink Bioarrays*

As imagens foram obtidas pelo escâner *GenePix 4000B*. Acima temos a imagem de um *bioarray* que foi hibridizado com sondas provenientes da linhagem celular PLJ e abaixo um *bioarray* hibridizado com sondas provenientes da linhagem celular MTWT.

Posteriormente, o programa *CodeLink System Software* normaliza os dados em relação à mediana local, confere os controles positivos e negativos e permite a análise dos dados para réplicas da mesma sonda. Alguns exemplos dos dados que podem ser obtidos através do *CodeLink System Software* estão na (Figura 17) .

Figura 17 - Telas de saída do CodeLink System Software.

A: sinais obtidos na lâmina hibridizada com MTWT em função dos sinais das lâminas hibridizadas com PLJ. Em azul, estão os genes que apresentaram uma diferença de expressão entre as duas lâminas menor que dois e, em amarelo, aqueles cuja diferença de expressão entre as lâminas foi maior que dois. B: Em destaque, plotagem do coeficiente de variação (CV) em função da mediana. A direita superior, tem-se os mesmos dados divididos de acordo com a intensidade da fluorescência (pequena, média e grande) e plotados na forma de um gráfico de barras. Na faixa abaixo, os dados são mostrados numericamente.

Para a seleção dos genes diferencialmente expressos, inicialmente foi utilizada a normalização realizada através do *CodeLink System Software*. Assim, era calculada a razão do sinal entre as condições estudadas e esta comparada com um *cut-off* selecionado arbitrariamente. Posteriormente, ao invés de ser utilizado o *ratio* para a seleção dos genes diferencialmente expressos, foi utilizada uma análise estatística (e-Bayes) que identificava os genes que possuem uma diferença estatisticamente significativa entre os *bioarrays* estudados, ou seja, com $p < 0,05$.

Ainda que ambas as abordagens estejam corretas, a metodologia e-Bayes nos permitiu visualizar que o método de normalização que vinha sendo utilizado (mediana local) não era adequado para a análise das lâminas. Anteriormente, utilizando um *cut-off* de 1,6 (sugerido pelo fabricante como suficiente para a seleção dos genes diferencialmente expressos), foram selecionados 795 genes como sendo candidatos a serem diferencialmente

expressos, entretanto, quando os mesmo candidatos foram submetidos ao e-Bayes apenas 395 genes apresentaram um $p < 0,05$.

Essa diferença entre o número de genes identificados provavelmente se deu porque no método de normalização do *CodeLink* System Software cada lâmina é normalizada separadamente. Assim, quando as lâminas foram analisadas em conjunto pelo e-Bayes a falta de uma normalização global acarretou em um $p > 0,05$ de boa parte dos genes, o que nos levou a concluir que um novo método de normalização deveria ser utilizado.

Partindo dos dados não normalizados, foram testados dois métodos de normalização, o *Quantile* e o *LOWESS*, este último utilizando um índice de alisamento de 0,3. Os gráficos de dispersão resultantes dos três métodos de normalização (mediana, *Quantile* e *LOWESS*) estão representados na Figura 18.

A

B

C

Figura 18 – Gráficos de dispersão após diferentes métodos de normalização

Representação gráfica da intensidade dos *spots* dos *bioarrays* sendo PLJ1 os dados referentes à intensidade dos *spots* na primeira hibridização dos *bioarrays* com cRNA proveniente da linhagem PLJ, PLJ2 a segunda hibridização e, analogamente, para MT1 e MT2. Cada gráfico representa a intersecção dos dados segundo as indicações na diagonal. Assim, na primeira linha, segunda coluna, por exemplo, temos que na abscissa estão plotados os dados referentes à hibridização da lâmina PLJ2 e na ordenada os dados referentes à hibridização da lâmina PLJ1. Em A temos os dados normalizados pela mediana local, em B normalizados pela metodologia *Quantile* e em C pela metodologia *LOWESS*.

Na distribuição dos dados de um *array* é assumido que a quantidade de genes induzidos no modelo estudado tende a ser proporcional à quantidade de genes que são inibidos, e que a maior parte dos genes não sofre alteração; portanto, o gráfico que representa a distribuição das intensidades descreve uma curva com coeficiente angular 1 (Quackenbush, 2001). Levando em conta esse critério, na Figura 18-A pode-se observar que na normalização por mediana local, a menos de PLJ1xMT1 e PLJ2xMT2, as distribuições apresentam uma curvatura no sentido diagonal, mostrando a falta de normalização dos mesmos. Em contrapartida, tanto após a normalização por *Quantile* (Figura 18-B) quanto após a normalização por *LOWESS* (Figura 18-C) todos os gráficos descrevem uma distribuição homogênea ao longo da diagonal.

Após a nova normalização, foi aplicado o método e-Bayes para identificação dos genes diferencialmente expressos e, conforme pode ser observado na Tabela 3, ambos os novos métodos de normalização identificaram 1.129 genes como sendo candidatos a serem diferencialmente expressos ($p < 0,05$).

Tabela 3 – Número de genes selecionados em cada método de normalização

Método de normalização	No. de genes selecionados*
Mediana local	378
Quantile	1.129
<i>LOWESS</i>	1.129

* Genes que apresentaram uma diferença de expressão estatisticamente significativa ($p < 0,05$) entre os modelos estudados

Como não houve diferença significativa entre os genes identificados após a normalização por *Quantile* ou por *LOWESS*, por ser um método já consagrado na literatura para a normalização de dados de *arrays*, o *LOWESS* foi selecionado para ser utilizado nas análises futuras. Posteriormente, foi descrito que, assim como foi constatado neste trabalho, entre diversos

métodos o LOWESS consistia no melhor método para normalização dos dados provenientes do CodeLink (Wu, Dave *et al.*, 2005).

Ao invés de trabalharmos com gráficos de dispersão dos sinais de PLJ x MTWT, como estava sendo utilizado anteriormente, foram utilizados gráficos M x A (Dudoit e Fridlyand, 2002; Quackenbush, 2002), onde $M = \log_2(PLJ * MTWT) / 2$ e $A = \log_2(PLJ / MTWT)$ de cada elemento dos *arrays*, pois esse gráfico permite uma melhor visualização de artefatos da hibridização e da normalização (Figura 19).

Figura 19 - Gráfico MA do CodeLink

Gráfico MA das duas hibridizações estudadas. Acima do ponto $M=0$ temos os genes induzidos por MT e abaixo de $M=0$ os genes cuja expressão é reprimida por MT. Em vermelho está plotada a curva de regressão localmente ponderada para um índice de alisamento de 0,3. Em verde estão indicados os pontos selecionados pelo método e-Bayes em cada uma das normalizações, LOWESS (acima) e Mediana local (abaixo).

Conforme pode se observado pela curva de regressão indicada na Figura 19, enquanto na normalização *LOWESS* a curva descreve uma reta ao longo da ordenada $M=0$, na normalização pela mediana local a curva possui uma curvatura ao longo de $M=0$, indicando uma possível falha no método de normalização pela mediana.

4.3.3. Análise dos genes selecionados

Após a seleção dos genes cuja expressão diferencial apresentava um $p < 0,05$, foi utilizado o programa *PubGene* (<http://www.PubGene.org/>) para analisar as possíveis interações entre esses genes. O *PubGene* atribui notas (similares ao *e-value* utilizado pelo programa *Blast*) a cada uma das possíveis interações entre os genes analisados. Uma das vias identificadas pelo *PubGene* pode ser observada na Figura 20.

Figura 20 - Tela de saída do programa *PubGene*

Em tons de verde estão os genes induzidos por MT e em tons de vermelho os genes reprimidos por MT. Os números indicados representam a pontuação atribuída à possível interação existente entre os genes.

Das possíveis interações identificadas pelo *PubGene*, foram selecionados 29 genes para terem sua expressão diferencial analisada por *Real-Time PCR* (Tabela 4).

Para analisar a lista dos genes selecionados, também foi utilizado o programa *FatiGO* (Al-Shahrour, Diaz-Uriarte *et al.*, 2004) que, entre outras funções, classifica os genes de acordo com os critérios do *Gene Ontology Consortium*, ou seja, componente celular, processo biológico e função molecular. Parte da tela de saída do *FatiGO* é mostrada na Figura 21.

Figura 21 – Tela de saída do *FatiGO*

Estão representadas algumas das classificações fornecidas pelo *FatiGO*. As porcentagens/barras em vermelho são referentes aos genes selecionados como induzidos por MT e em verde os genes reprimidos pela atividade de MT.

A partir da classificação *FatiGO*, foram selecionados 34 genes relacionados à divisão celular e/ou ciclo celular para serem confirmados por *Real-Time PCR*. O *FatiGO* também forneceu uma lista de 69 genes que ainda não haviam sido classificados pelo *Gene Ontology*. Esses genes foram submetidos ao *Blast*, sendo que 14 destes genes ainda não haviam sido caracterizados, sendo selecionados também para futura análise por *Real-Time PCR*.

Tabela 4 - Genes selecionados a partir dos dados obtidos no CodeLink

Os genes listados foram selecionados a partir dos dados obtidos das lâminas do CodeLink com o auxílio dos programas PubGene, FatiGO e Blast. O $\log_2\text{ratio}$ é definido como o logaritmo na base 2 da razão da intensidade, obtida no CodeLink, da hibridização com MTWT e PLJ. Assim, os $\log_2\text{ratio}$ positivos representam os genes cuja expressão é induzida pela atividade de MT e os negativos correspondem aos genes reprimidos por MT.

Gene	Método	$\log_2\text{ratio}$
Adam11	PubGene	2.37
Adam12	PubGene	3.95
Adamts1	PubGene	2.52
Adamts5	PubGene	2.48
Ak013687	FatiGO / Blast	-1.61
AK017200	FatiGO / Blast	2.37
AK018939	FatiGO / Blast	1.53
AK019875	FatiGO / Blast	-1.15
AK029831	FatiGO / Blast	1.87
AK080810	FatiGO / Blast	-1.40
Basp1	FatiGO	7.59
Bmpr2	FatiGO	1.84
Brca1	PubGene	-1.15
Cdc20	FatiGO	-3.55
Cdc25c	FatiGO	-2.98
Cdc2a	PubGene	-2.00
Cdc6	FatiGO	-2.21
Cdca5	FatiGO	-4.81
Cdk1	FatiGO	-2.00
Chek2	FatiGO	-2.31
CiclinaA2	FatiGO	-3.96
CiclinaB2	FatiGO	-3.42
CiclinaG1	FatiGO	-2.11
Cks2	FatiGO	-3.54
Dkk1	PubGene	8.43
Dkk2	PubGene	2.46
Dkk3	PubGene	1.87
Fanca	PubGene	-3.36
Fgfr1	PubGene	1.11
Frat1	FatiGO	1.21
Fzd2	FatiGO	1.29
Gmnn	FatiGO	-3.08
Gspt2	FatiGO	2.39
Gtf2i	FatiGO	1.17
Id1	FatiGO	-1.57
Ier5l	FatiGO	1.15
Igf1	PubGene	3.19
Igf2	PubGene	6.70
Igf2r	PubGene	1.34

Gene	Método	$\log_2\text{ratio}$
Igfbp7	PubGene	2.48
Klf2	FatiGO	-1.39
MAD2l1	FatiGO	-3.44
Map3k5	FatiGO	2.18
Mcm3	FatiGO	-1.37
Mmp14	PubGene	1.37
Mmp3	PubGene	3.43
Mmp9	PubGene	2.34
Nek2	PubGene	-2.17
NM_007955	FatiGO / Blast	1.19
NM_020616	FatiGO / Blast	1.29
NM_025766	FatiGO / Blast	1.11
NM_026152	FatiGO / Blast	-2.04
Npdc1	PubGene	5.07
Oip5	FatiGO	-2.35
Pnpt1	FatiGO	-1.33
Pou5f1	FatiGO	-1.40
Ptn	PubGene	2.76
Ptp4a3	PubGene	-1.20
Pttg1	FatiGO	-4.21
Rad51	PubGene	-2.24
Rad51	FatiGO	-2.24
Rgs2	FatiGO	-6.93
Rhox5	FatiGO	-4.18
Rybp	PubGene	1.99
Sap30	FatiGO	-3.19
Sgce	PubGene	1.92
Smarca1	PubGene	3.76
Stat4	PubGene	-1.18
Tmem158	FatiGO	5.14
Tmem162	FatiGO	-1.04
Uhrf1	FatiGO	-1.79
W09175	FatiGO / Blast	-3.27
W70351	FatiGO / Blast	-1.56
Wdr70	FatiGO / Blast	-1.38
Wnt4	PubGene	2.61
XM_001004922	FatiGO / Blast	1.89
Zyx	PubGene	1.65

4.4. Análise da expressão dos genes candidatos por RT-PCR

4.4.1. Síntese da primeira fita do cDNA

Foi realizado um protocolo que já era utilizado pelo laboratório para a síntese de cDNAs no projeto *Transcription Finishing Initiative* (TFI), onde são utilizados tanto oligo-dT como *primers* randômicos para a síntese do cDNA.

O RNA total utilizado nas reações foi obtido através da purificação do lisado das linhagens celulares estudadas (A31, PLJ, MTWT, MT315, MT322, MT2F e 250phe) em colchão de Cloreto de Césio mediante ultra-centrifugação, conforme descrito no item 3.3.

4.4.2. Controle da qualidade do cDNA

Ainda seguindo os protocolos padronizados no projeto TFI, foram realizadas reações de PCR para confirmar a qualidade dos cDNAs. Assim, para averiguar a qualidade do cDNA sintetizado, foi realizado um PCR utilizando *primers* para um gene que, teoricamente, possui a mesma expressão em todas as linhagens celulares, ou seja, GAPDH (Figura 22-GAPDH). Para testar a presença de cDNAs referentes a genes de baixa expressão, foram utilizados *primers* para um gene de baixa expressão, p53 (Figura 22-p53)

O próximo teste consistiu em verificar se as amostras de RNA, que serviram como *template* para a síntese do cDNA, não estavam contaminadas com DNA genômico. Para tanto, foi realizado um PCR utilizando *primers* que se ligavam ao íntron do gene MLH (Figura 22-MLH). Este teste permite verificar se houve a síntese de cDNA de transcritos longos correspondentes à porção 5' do gene NOTCH-2, cujo transcrito completo possui aproximadamente 12Kb (Figura 22-NOTCH2).

Figura 22 - Gel de agarose para confirmação da qualidade dos cDNAs.

Fracionamento dos produtos das reações de PCR utilizando como *template* cDNAs provenientes das linhagens estudadas (A31, PLJ, MTWT, MT315, MT322, MT2F e 250phe) e um controle negativo sem *template* (Controle -). Foram utilizados *primers* para os genes p53, GAPDH, MLH1 e NOTCH2. O resultado das amplificações está indicado pelas setas.

4.4.3. Padronização das reações de RT-PCR

Para quantificar a expressão dos genes nas diferentes linhagens, era necessário que todas as reações de PCR contivessem a mesma quantidade de amostra. Assim, a expressão do gene GAPDH, um gene *housekeeping*, foi utilizada para padronizar a quantidade de amostra que seria utilizada em cada reação. Esta padronização foi realizada em duas etapas, uma para padronizar a quantidade de ciclos onde a amplificação de GAPDH ainda

estaria na sua fase exponencial. Uma vez definido esse número de ciclos (Figura 23-A), foi realizada uma segunda padronização da quantidade de cDNA utilizada como *template* de modo que todas as bandas observadas no gel apresentassem a mesma intensidade de sinal (Figura 23-B).

Figura 23 – Padronização da amplificação do cDNA de GAPDH para o RT-PCR

A: Padronização da fase exponencial de amplificação do gene GAPDH. Na reação de PCR foram utilizados como template cDNAs provenientes das linhagens PLJ e MTWT. Os números 10, 15, 20, 25 e 30 indicam os números de ciclos da reação de PCR. B: Padronização da quantidade de *template* que seria utilizada para os ensaios de RT-PCR. Foram utilizados, como *template*, cDNAs provenientes das linhagens A31, PLJ, MTWT, MT315, MT322, MT2F e 250phe

Na Figura 23-A tem-se a imagem do produto da reação de PCR realizada para determinar o número adequado de ciclos de PCR. Conforme há um aumento no número de ciclos, tem-se um aumento na quantidade de produto sendo que menor variação entre os ciclos 25 e 30, principalmente em MTWT, indica que a reação não está mais na sua fase exponencial de amplificação. Portanto, foi escolhido um número de ciclos que ainda estivesse na faixa exponencial de amplificação (entre 20 e 25 ciclos), ou seja, 23 ciclos. Uma vez determinado este número, foram realizadas reações para padronizar a intensidade do sinal, da amplificação de GAPDH, em todas as linhagens estudadas. Na Figura 23-B, apresenta-se o primeiro gel corrido antes que as quantidades de cDNA utilizadas como *template* fossem normalizadas.

Nessa etapa, a foto era escaneada e então as bandas da imagem eram quantificadas pelo programa *ImageQuant* (Molecular Dynamics). Baseado nessa quantificação, era realizada a alteração nas quantidades de *template* utilizado nas reações, e uma nova reação era realizada. Entretanto, havia uma grande discrepância entre o que se observava no gel e a intensidade das bandas que era fornecida pelo *ImageQuant*. Portanto, as imagens passaram a ser obtidas pelo aparelho *ImageMaster VDS* e, posteriormente, eram analisadas pelo *ImageQuant*.

Para padronização, uma das amostras foi tomada como padrão (PLJ), e as quantidades de *template* utilizadas nas demais reações eram alteradas de forma que o sinal ficasse igual ao de PLJ. Após, aproximadamente, 8 reações de PCR, para cada amostra, foi possível padronizar a quantidade de *template* a ser utilizada para que a amplificação de GAPDH gerasse a mesma intensidade quando o produto do PCR fosse analisado em um gel de agarose 1%, conforme pode ser observado na Figura 24.

Figura 24 - Resultado final da padronização de GAPDH

4.4.4. Confirmação da expressão diferencial

Uma vez definida a quantidade de *template* que seria utilizada para amplificar o cDNA, foi realizada a padronização do número de ciclos que seria utilizado para amplificar os genes MN1, Ferritina e UNR, seguindo a mesma abordagem utilizada para o gene GAPDH. Padronizado o número de ciclos, as amostras foram amplificadas utilizando a quantidade de *template* padronizada. O resultado dessas amplificações pode ser observado na Figura 25.

Figura 25 – RT-PCR dos genes selecionados a partir do RDA.

Imagem de géis resultantes das amplificações dos genes MN1, UNR, Ferritina e do gene utilizado como controle, GAPDH. Como *template*, foram utilizados cDNAs provenientes das linhagens celulares A31, PLJ, MTWT, MT315, MT322, MT2F e 250phe.

Para o gene MN1 foram padronizados 43 ciclos de amplificação, sendo que só foi observada sua amplificação nos cDNAs provenientes das linhagens A31 e PLJ. Posteriormente, foi realizada uma nova reação de PCR com 50 ciclos, onde também não foi observado sinal algum nessa reação nas linhagens que expressão MT selvagem e mutantes (figura não apresentada), o que mostrou que MN1 tem sua expressão completamente inibida por MT.

Os genes UNR (*upstream of NRAS*) e Ferritina apresentaram uma pequena variação nas diferentes das linhagens analisadas. Como essas variações eram pequenas, seria necessário que fossem realizadas, no mínimo, triplicatas de tais experimentos para afirmarmos, com segurança, a existência de alguma regulação desses genes por MT.

Entretanto, quando começamos a fazer as réplicas dos experimentos de confirmação, tanto dos genes descritos acima quanto de GAPDH, começamos a nos deparar com uma variação significativa entre os resultados

anteriores, sendo que nessa tentativa de obtermos dados coerentes foi realizado uma média de 15 reações por amostra.

Uma das suspeitas sobre esta variação estava sobre o aparelho utilizado para obter a imagem do gel (ImageMaster® VDS). Para testar tal hipótese, a imagem de um mesmo gel foi adquirida em três orientações diferentes, e depois as bandas foram densitometradas (Figura 26).

Figura 26 - Densitometria em três orientações diferentes.

No gel foram aplicadas seis amostras distintas (A-F). Em azul, temos o resultado da densitometria quando o gel foi escaneado na forma usual. Em roxo, o resultado quando o gel foi girado em 90° e em amarelo quando o gel foi girado em 180°.

Como pode ser observado na Figura 26, existe uma grande variação entre os resultados dos três escaneamentos, tanto na intensidade de cada banda quanto na relação entre essas intensidades, o que provavelmente ocorreu porque as lâmpadas realmente não estavam iluminando o gel homogeneamente.

Antes que o problema fosse solucionado, o Departamento de Bioquímica da USP adquiriu um termociclador capaz de fazer a quantificação das amostras em tempo real (*Real-Time PCR*). Assim, em vista dos

problemas que estavam ocorrendo e do fato *Real-Time PCR* ser uma metodologia quantitativa, ao contrário do RT-PCR que é semi-quantitativa, optou-se por fazer as confirmações por *Real-Time PCR*.

4.5. Confirmação dos genes por *Real-Time PCR*

4.5.1. Padronização da concentração dos *primers*

Para estabelecermos qual seria a concentração final dos *primers* nas reações de amplificação, foram realizadas reações com concentrações finais de *primers* iguais a 200 η M, 600 η M e 800 η M (Figura 27).

Figura 27 - Curvas de amplificação dos genes GAPDH, Dlk1 e UNR

Nas reações para amplificação dos genes GAPDH, Dlk1 e UNR, as concentrações finais dos *primers* eram 200 η M, 600 η M e 800 η M, em duplicata. Os controles negativos, sem *template*, não estão representados nesses gráficos.

Na Figura 27 estão representados os gráficos das amplificações dos genes GAPDH, DLK e UNR. Nessa primeira etapa, também foram padronizados *primers* para os genes Septina2, Ferritina, p23, Actina e Tubulina, esses dois últimos para serem utilizados, juntamente com GAPDH, como controles internos.

A seleção da concentração final dos *primers* foi baseada, principalmente, na variação do Ct apresentado entre as duplicatas. Na Tabela 5 tem-se, como exemplo, os Cts apresentados por GAPDH quando foi escolhido um *threshold* de 0,4. Os Cts que apresentaram menor variação entre as duplicatas foram aqueles cuja concentração do *primer* na reação era de 600 η M, portanto, essa concentração de *primer* foi escolhida para

amplificar todas as amostras de GAPDH. O mesmo procedimento foi realizado com os outros *primers*.

Tabela 5 – Ct do *primer* de GAPDH em função da sua concentração final

Em cada coluna tem-se, respectivamente, a posição da reação na placa de 96-wells, o *primer* utilizado, a concentração final do *primer* na reação e o Ct encontrado para um *threshold* de 0,4.

Poço	<i>Primer</i>	Concentração final do <i>primer</i> (nM)	Ct
A6	GAPDH	0	34,03
B6	GAPDH	0	34,36
C6	GAPDH	200	17,53
D6	GAPDH	200	17,66
E6	GAPDH	600	17,35
F6	GAPDH	600	17,31
G6	GAPDH	800	17,47
H6	GAPDH	800	17,19

4.5.2. Análise da eficiência dos *primers*

Antes de iniciar a análise dos genes selecionados, foi calculada a eficiência dos *primers* sintetizados. Paratanto, foram realizadas reações de *Real-Time PCR* utilizando como *template* diluições seriadas (1:40, 1:80, 1:160, 1:320, 1:640) de um *MIX* de cDNA das amostras que seriam analisadas. Após a reação, foi elaborado um gráfico do logaritmo da diluição utilizada em função do Ct obtido (Figura 28) sendo que a eficiência do *primer* foi calculada em função do coeficiente angular da curva obtida. Foram aceitos aqueles *primers* que apresentaram uma eficiência entre 80% e 110% e com um R^2 maior que 0,85. O mesmo raciocínio foi utilizado para o cálculo da eficiência de todos os *primers* que seriam utilizados para a análise dos dados provenientes do *CodeLink*.

Figura 28 – Diluição da amostra em função Ct para o gene Npdc1

Gráfico representativo do gene Npdc, no qual a distribuição dos pontos descreve a curva utilizada para o cálculo da eficiência dos *primers*.

4.5.3. Padronização do gene *housekeeping*

Para aumentar a confiabilidade dos resultados, foi amplificado um gene e este foi normalizado com as três possibilidades de genes *housekeeping*, sendo eles GAPDH, Actina e Tubulina. Quando os resultados foram analisados, percebeu-se que o perfil de expressão apresentado pelos genes variava de acordo com o gene *housekeeping* utilizado (Figura 29), de forma que a inibição observada na linhagem MT315 quando o controle interno é GAPDH, por exemplo, era trocada por uma indução quando se utiliza Actina ou Tubulina. Ainda que com intensidades distintas, os resultados apresentados pela Actina e pela Tubulina possuíam um perfil semelhante, o que é coerente visto que ambas são proteínas relacionadas ao citoesqueleto celular.

Figura 29 – Real-Time PCR utilizando diferentes controles internos

O gráfico mostra o perfil do sinal, em relação à PLJ, utilizando como controle interno GAPDH, Actina e Tubulina. Como *template* foi utilizada a linhagem controle (PLJ) e as linhagens expressando MT selvagem (MTWT) e mutantes (MTWT, MTW315, MT322, MT2F e 250phe).

Como não havia dados na literatura que indicassem os controles adequados para o sistema estudado, optou-se por se basear nos dados do *CodeLink* para esta seleção. De acordo com os dados do *CodeLink*, GAPDH apresentava uma indução 2,05 vezes, o que é uma variação mais do que suficiente para invalidar um resultado obtido no *Real-Time PCR*. A β -actina, para a qual foram desenhados os *primers*, não estava representada na lista de genes do *CodeLink*, entretanto, havia outras isoformas de actina no *bioarray*, e estas eram sempre inibidas por PLJ em uma razão de 5 a 10 vezes, o que dá indícios de que a actina também não consiste em um bom controle interno. A isoforma da tubulina utilizada foi a tubulina- $\alpha 6$, que no *bioarray* apresentava uma repressão por MT que variava cerca de 1,5 vezes.

Dentro das possibilidades, a tubulina se mostrou a mais viável. Entretanto, dentre os possíveis controles internos presentes no *bioarray*, a proteína ribossomal 36B4 (fosfoproteína ribossômica ácida P0) apresentava uma indução em MT de 1,1 vezes, que mostrou ser a opção mais coerente e, portanto, passou a ser o controle interno utilizado em todos os experimentos de *Real-Time PCR*.

4.5.4. Análise da expressão diferencial

Após cada reação de *Real-Time PCR*, definiu-se um *threshold* de 0,23 e, com base nesse *threshold*, o programa *ABI-Manager SDS* forneceu os Cts que foram utilizados para calcular os Δ CPs e os *ratios* em relação à linhagem PLJ. Como controle interno foi utilizado o nível de expressão do gene 36B4 e todas as reações foram realizadas em triplicata.

Para a análise por *Real-Time PCR*, além dos genes selecionados pelo *CodeLink* (Tabela 4), foram analisados os genes obtidos no RDA (Tabela 2), genes relacionados ao fator de transcrição AP-1, que é regulado por MT, e genes relacionados ao controle do ciclo celular, sendo eles: AKT1, Bcl2, CD44, CDK4, c-fos, ciclina D1, c-jun, E-caderina, MMP2, MMP7, MMP9, myc, p16, p21, p27, p53, PAI-2, PKC (todas isoformas), TGF- β 1, TIMP3, TNF α , UPA, UPAR.

Inicialmente, foi analisada a expressão dos genes selecionados somente na linhagem controle (PLJ) e na linhagem que expressa MT selvagem (MTWT). Após a reação de *Real-Time PCR*, foi realizado um teste de Tukey para identificar as diferenças de expressão estatisticamente significativas.

Assim, após os dois critérios de exclusão utilizados, o teste de Tukey e a análise da eficiência dos *primers* (item 4.5.2), 64 genes foram selecionados para serem também analisados nas linhagens que expressam MT mutado para a via de PI-3K (MT315), PLC- γ (MT322), PI-3K/PLC- γ (MT2F) e Ras/MAPK (250phe).

Para análise estatística dos dados, os *ratios* foram submetidos ao programa Prism onde, após uma análise de variância unidirecional seguida de um teste de Tukey, obteve-se para cada gene as diferenças de expressão entre as linhagens analisadas e a significância estatística de cada diferença (Figura 30).

Figura 30 - Resultado do *Real-Time PCR* dos genes analisados

Representação gráfica da análise estatística dos 64 genes estudados. No eixo da ordenada está representada a diferença de expressão em relação à amostra PLJ e, no eixo da abscissa, a linhagem cujo cDNA foi utilizado como *template* para reação.

Como pode ser observado na Figura 30, os genes estudados apresentaram diferentes respostas tanto à expressão de MT como à expressão dos mutantes de MT. Apesar do programa Prism ter identificado as diferenças de expressão estatisticamente significativas ao analisar cada gene isoladamente, esse dado não era suficiente para realizar uma análise global dos possíveis padrões de regulação existentes. Portanto, para analisar a relação entre os genes estudados foi realizada uma análise de correlação de Pearson.

A análise de correlação de Pearson relaciona os genes dois a dois e identifica se existe um padrão de expressão estatisticamente significativo entre as linhagens estudadas. Ao final da análise tem-se um coeficiente de relação (r) que varia de -1 a 1 representando uma correlação inversa (valores negativos) ou direta (valores positivos) entre os genes estudados. Ao coeficiente de correlação é atribuído um coeficiente de Pearson (p) que representa a significância estatística do coeficiente de correlação. O coeficiente de correlação cujo módulo se encontra entre 0,6 e 0,9 é considerado forte e com módulo acima de 0,9 é considerado muito forte.

Após o cálculo da correlação, o coeficiente p foi corrigido pelo método FDR (*False Discovery Rate*) (Benjamini e Hochberg, 1995) e os *outliers* foram minimizados através do método *leave-one-out*. Foram consideradas somente aquelas correlações com um $|r| > 0,7$ e $p < 0,05$. Para essa análise, foi desenvolvido um sistema na linguagem R de programação, no qual uma das

telas de saída pode ser observada na Figura 31. Ao todo foram analisadas 4096 correlações, sendo que 92 destas satisfizeram os critérios de seleção.

Figura 31 - Tela de saída do programa para análise das correlações

Representação gráfica dos níveis de expressão de dois genes (Brca1 e Nek2) que apresentaram um coeficiente de correlação positiva muito forte ($r=0,96$) com um p estatisticamente significativo ($p<0,05$). Na primeira linha estão os gráficos de barras referentes ao nível de expressão em relação ao gene PLJ, na esquerda inferior está o gráfico de dispersão e na direita inferior a representação, em forma de linhas, dos gráficos de barras sobrepostos.

Na Figura 31 estão representadas as análises de dois dos genes que foram identificados como apresentando uma correlação positiva entre si, o que pode ser observado, além do r e do p , pelo fato dos pontos estarem dispostos ao longo da diagonal no gráfico de correlação e pelos gráficos de

barra apresentarem um comportamento semelhante. Se o coeficiente de correlação fosse negativo, o padrão de expressão seria invertido, ou seja, o aumento da expressão de um gene em uma linhagem teria correlação com a diminuição do outro gene analisado na mesma linhagem. Em seguida, as correlações foram utilizadas para construção de mapas de correlação (Figura 32).

Figura 32 - Mapas de correlação dos genes analisados.

Nos círculos estão indicados os nomes dos genes e as linhas indicam que existe uma correlação entre os pares de genes. Os gráficos indicam os perfis de expressão dos genes presentes no respectivo mapa de correlação. Nos gráficos não estão representadas as correlações negativas.

No total, foram identificadas 92 correlações entre 39 dos genes analisados. Posteriormente, os mapas da Figura 32 foram utilizados para inferir possíveis relações entre os genes estudados, uma vez que uma correlação indica a possibilidade de que esses genes sejam ativados ou inibidos por mecanismo em comum, gene esse que possivelmente está sendo regulado pelas vias ativadas por MT e inibidas nos mutantes de MT.

4.6. Análise funcional do gene MN1

O gene MN1 foi encontrado nos experimentos de subtração como sendo completamente inibido pela ação de MT, resultado este que foi posteriormente confirmado por ensaios de RT-PCR e, recentemente, por ensaios de *Real-Time PCR* (Figura 33). Como o gene ainda não havia sido caracterizado, foi selecionado para os ensaios de análise funcional.

Foram escolhidas duas metodologias para análise funcional do gene MN1, que foram nocautear o gene utilizando a metodologia de RNAi, e estudar os efeitos dessa inibição na linhagem PLJ, e super-expressar MN1 na linhagem MTWT e nas linhagens que expressam MT mutante.

Figura 33 – Real Time PCR do gene MN1

No eixo da abscissa tem-se o cDNA cuja amostra serviu como *template* e no eixo das ordenadas a variação da fluorescência em relação ao sinal da fluorescência em PLJ.

4.6.1. Clonagem do cDNA completo do gene MN1

4.6.1.1. Amplificação através da metodologia de RACE

Os *primers* utilizados para amplificação do transcrito possuíam T_m (*Melting Temperature*) diferentes. Como não eram conhecidas as condições ideais para a amplificação de MN1, foram realizadas reações de PCR variando o número de ciclos de 38 a 45, e temperaturas de anelamento de 60°C a 70°C e os tempos de extensão de 45seg a 2min.

Figura 34 - Padronização da amplificação do gene MN1

Da esquerda para direita, reações do Race-3': 43 ciclos iniciando com a temperatura de extensão de 64°C e diminuindo 0,1°C da mesma por ciclo e 43 ciclos de 60°C. Em seguida, reações do Race-5': 44 ciclos de 58°C, 59,5°C, 64,7°C, 67°C, 67°C diminuindo 0,2°C por ciclo e 64°C diminuindo 0,1°C por ciclo, respectivamente.

Na Figura 34 está a imagem de um dos géis realizados para padronizar as condições para amplificação de MN1. O objetivo consistia em obter uma banda única e sem arraste, como quase foi obtido nas reações para o Race-5'. Por fim, para amplificação do Race-3' foi padronizada uma temperatura de anelamento 60°C/42 ciclos de amplificação. Para o Race-5'

foram encontradas duas possibilidades, sendo uma com a temperatura de anelamento de 64°C e outra com temperatura de anelamento de 65°C, ambas com 42 ciclos de amplificação(Figura 35).

Figura 35 – Condições selecionadas para amplificação de MN1

Da esquerda para direita, Race-3' 42 ciclos com temperatura de anelamento de 60°C, Race-5' com temperatura de anelamento de 64°C e Race-5' com temperatura de anelamento de 65°C, todos com 42 ciclos de amplificação.

Embora tenham sido realizados diversos testes, não conseguimos obter uma banda única no Race-5', sendo que as condições que apresentaram um menor número de bandas e menor arraste foram as duas representadas na Figura 35.

Após clonagem no vetor pUC18 e transformação de bactérias, estas foram então plaqueadas, as colônias transformantes isoladas e o resultado confirmado através de um PCR de colônia. Posteriormente, foi realizada uma reação de seqüenciamento do produto do PCR de colônia e os contigs referentes aos genes foram montados utilizando-se os programas

Phred/Phrap (<http://www.phrap.org/phredphrapconsed.html>) analisam que, respectivamente, a qualidade do seqüenciamento e montam os *contigs*.

Após os seqüenciamento foi constatado que todas as seqüências obtidas através do 5'RACE eram produto de um alinhamento inespecífico do *primer* MN1-Forward. Entretanto, foi possível amplificar uma região do gene MN1 através do 3'RACE.

A

B

Figura 36 - Tela de saída do programa BLAT

Na tela de saída do BLAT está mostrado o resultado do alinhamento com o genoma de camundongo da seqüência predita por ferramentas de Bioinformática para MN1 (XM_485667), a seqüência obtida pela metodologia de subtração RDA (Subtração) e a seqüência obtida no 3'RACE (RACE). Os retângulos pintados representam os éxons e as linhas que ligam os retângulos pintados representam íntrons. Em A tem-se a tela de saída mostrando toda a seqüência de XM_485667 e em B uma ampliação da área em vermelho de A onde pode ser observado o alinhamento das seqüências obtidas pela subtração e pelo RACE com o XM_485667.

Na Figura 36 pode ser observada a seqüência que foi clonada através do 3'RACE utilizando o *primer* MN1-Reverse. Para amplificar o restante do gene MN1, foram desenhados *primers* baseados nas predições da seqüência XM_485667 e nas ESTs descritas que alinham com essa região cromossômica de MN1 (Figura 36-B).

4.6.1.2. Amplificação do gene MN1 utilizando *primers* baseados em predições e ESTs conhecidas para o gene MN1

A segunda alternativa para amplificação do gene MN1 foi desenhar *primers* baseados na seqüência XM_485667 (*primers* F9, F1148, F2439, F3297, F4413, F6990, R1169, R2670, R3807, R4470, R5845, R6229 e R8049) e em algumas ESTs que alinham com a região cromossômica onde se encontra MN1 de camundongo, sendo elas BB232797 (*primers* F54, F320 e R281), BB658850 (*primers* F198, R468 e R583) e CA315561 (*primer* F87). Um esquema da disposição desses *primers* ao longo da seqüência XM_485667 pode ser observada na Figura 37.

Figura 37 - Esquema da disposição dos *primers* desenhados para amplificação de MN1

Os *primers* foram desenhados ao longo da sequência predita para MN1(XM_485667). BB658850, CA315561 representam ESTs que alinham com XM_485667. BB232797, apesar de não alinhar com XM_485667, está na mesma região cromossômica que o gene MN1 e os *primers* desenhados para sua amplificação podem ser observados na esquerda inferior da imagem. Em amarelo estão representados os *primers Forward* e em azul os *primers Reverse*. Em vermelho estão indicados os transcritos que foram amplificados por PCR utilizando os *primers Forward* e *Reverse* correspondentes

Figura 38 - Testes de amplificação de MN1

Nos géis estão representados algumas das condições que foram testadas para amplificação de MN1. Em A e B foram utilizadas diferentes temperaturas de anelamento do *primer*, 58°C e 56°C, e em C foi testado a adição de Betaína para facilitar a amplificação de transcritos ricos em GC.

Como não eram conhecidas as condições ideais para amplificação dos transcritos, foram realizados diversos testes sendo que alguns deles podem ser observados na Figura 38. Com esses *primers* foi possível amplificar oito transcritos utilizando as seguintes combinações diferentes de *primers* (Figura 37):

F6990 – R8049	F198 – R468
F4413 – R5845	F87 – R5845
F4413 – R6229	F87 – R6229
F1148 – R2670	F54 – R281

Apesar de uma porção significativa do gene MN1 ter sido amplificada, com essas combinações não foi possível obter o transcrito completo do gene. Para amplificar as regiões restantes optou-se por realizar uma nova amplificação por RACE utilizando, além do SMART-RACE PCR *primer*, os *primers* desenhados a partir das predições de ferramentas de Bioinformática.

Na Figura 37, em vermelho, estão indicadas as seqüências do gene MN1 que foram amplificadas e seqüenciadas. Para amplificar a porção 5' do gene MN1, obteve-se sucesso ao utilizarmos as combinações de *primers* 5' SMART CDS *primer* IIA-Sal3A1/ F3287 e 3' SMART CDS *primer* IIA-Sal3A1/ F4413 (Figura 39).

Figura 39 - Resultado da amplificação da porção 3' do gene MN1

Para a amplificação da porção 3' do gene MN1 foram utilizados os *primers* 3' SMART CDS *primer* IIA Sal3A1 e F3287 ou F4413.

As três bandas resultantes da amplificação com o *primer* F3287 foram isoladas e purificadas do gel de agarose utilizando o kit GFX e, em seguida, submetidas a reações de seqüenciamento para identificação das seqüências. Contudo, até o momento não houve sucesso em obter a seqüência dos três transcritos.

4.6.2. Inibição de MN1 através de RNA de interferência

As construções pMDL, pRSVRev, pMDLG, pEGFP e pMN1 foram transfectadas (transfecção mediada por lipossomas) nas células empacotadoras 293T. Como controle da eficiência de transfecção, foi utilizada a construção contendo EGFP (*Enhanced Green Fluorescet Protein*).

De acordo com o controle utilizado (EGFP), foi detectada uma eficiência de infecção de aproximadamente 100%, tanto para PLJ infectada com o lentivírus que codificava para proteína EGFP (Figura 40-A/B) como para o que codificava para o RNAi de MN1 e a proteína GFP (Figura 40-C/D). A morfologia celular das populações de PLJ infectadas não sofreram nenhuma alteração morfológica significativa.

4.6.2.1. Confirmação por *Real-Time PCR*

Foi utilizada a metodologia de *Real-Time PCR* para confirmar a degradação do mRNA de MN1 devido ao RNAi (Figura 41), sendo que não foi detectado a amplificação de produto algum no cDNA da linhagem que expressava o RNAi de MN1 (RNAi-MN1), o que indica que o RNAi levou ao *knock-down* de MN1.

4.6.2.2. Curva de crescimento

Uma vez confirmada a inibição da atividade de MN1 pelo RNAi, foi realizada uma curva de crescimento para verificar se a inibição deste gene estava exercendo algum efeito direto no controle da divisão celular. Todos os experimentos foram realizados em duplicata. Para evitar possíveis erros de leitura, foram realizadas três leituras de cada amostra no contador eletrônico (Figura 42).

Segundo os dados obtidos para a curva de crescimento, representados na Figura 42, as duplicatas se mostraram reprodutíveis e não foi encontrada nenhuma diferença estatisticamente significativa entre os tempos de dobramento das linhagens estudadas.

4.6.2.3. Ensaio de formação de colônia em meio semi-sólido

Para determinar se a inibição da expressão de MN1 provocaria alterações no potencial tumorigênico (agarose 0,3%) e metastático (agarose 0,9%) das células PLJ, foi realizado o ensaio de crescimento em suspensão, em duplicata (Li, Price *et al.*, 1989). As colônias que se desenvolveram foram classificadas de acordo com o tamanho em Grandes (de 300 a 600µm), Médias (de 150 a 300µm) e Pequenas (de 0 a 150µm).

Após a contagem das colônias de células (Figura 43), não foi observada uma alteração estatisticamente significativa no número e tamanho das colônias encontradas, quando foram comparadas diferentes concentrações de células em diferentes concentrações de agarose.

Figura 40 – Células PLJ infectadas com lentivírus

As células PLJ foram infectadas com lentivírus expressando EGFP (A,B) e RNA de interferência para MN1 e GFP (C,D), 48 horas após a infecção. Em A e C tem-se a imagem obtida em microscopia de luz visível e em B e C em microscopia de fluorescência.

Figura 41 – Real-Time PCR para confirmação do RNAi de MN1

No eixo da abscissa estão as linhagens cujo cDNA serviu como *template*, PLJ (PLJ), PLJ infectada com vetor expressando somente GFP (GFP), PLJ infectada com vetor expressando GFP e RNAi de MN1 (RNAi-MN1). No eixo das ordenadas tem-se a variação da fluorescência em relação ao sinal da fluorescência de PLJ

Figura 42 – Curva de Crescimento de células expressando o RNAi de MN1

Resultado da curva de crescimento realizada com a linhagem PLJ, linhagem PLJ infectada com vetor contendo GFP e linhagem PLJ infectada com vetor contendo GFP e RNAi-MN1. No eixo da abscissa tem-se o número de horas após o plaqueamento inicial de 10^4 células, período no qual foi realizada a coleta e, no eixo das ordenadas, tem-se representado, em escala logarítmica, o número de células contadas no momento da respectiva coleta.

Figura 43 - Crescimento em suspensão de agarose.

Foram plaqueadas 10^3 e 10^4 células das linhagens PLJ (PLJ), PLJ expressando somente GFP (GFP) e PLJ expressando GFP e RNAi-MN1 (MN1), em agarose 0,3% e agarose 0,9%. Todas as condições foram realizadas em triplicatas. Os gráficos são referentes ao número de colônias pequenas (0 a 150 μ m), médias (150 a 300 μ m) e grandes (300 a 600 μ m).

4.7. Clonagem do cDNA completo correspondente ao gene Dlk1

Dlk1 foi inicialmente identificado através das membranas comerciais (Tabela 1) e pela metodologia de subtração (Tabela 2) como sendo inibido por MT. Posteriormente teve sua inibição por MT selvagem e mutantes confirmada por *Real-Time PCR* (Figura 44) e foi identificada no *CodeLink* como tendo um nível de repressão por MT de 22 vezes ($p < 0,01$).

Como a sequência de Dlk1 para camundongos já era descrita (NM_010052), bastou desenhar *primers* que amplificassem a ORF do gene e que contivesse adaptadores para sua posterior clonagem em um vetor de expressão. Quando dois promotores fortes, provenientes de vetores de expressão, estão atuando em conjunto existe a possibilidade de que um interfira na expressão do outro. Como as células onde Dlk1 seria expresso (MTWT e mutantes) já possuem o vetor PLJ integrado com um promotor de SV40, a sequência de Dlk1 foi clonada em dois vetores diferentes, sendo eles pCXN2 e pcDNA3.

Figura 44 - Real Time PCR do genes Dlk1

No eixo da abscissa tem-se o cDNA cuja amostra serviu como *template* e no eixo das ordenadas a variação da fluorescência em relação ao sinal da fluorescência de PLJ.

4.7.1. Amplificação do transcrito completo de Dlk1

Após a amplificação do cDNA correspondente ao transcrito completo de Dlk1, além do transcrito esperado de aproximadamente 1,5Kb também foi amplificado um outro transcrito de 1,2Kb e, como o mesmo deveria se tratar de algum *splicing* alternativo de Dlk1, ele também foi clonado (Figura 45).

Os produtos da amplificação foram fracionados em um gel de agarose e, após as bandas serem recortadas, foram purificados utilizando o kit GFX. Em seguida, os vetores pcDNA3, pCXN2 e os dois transcritos de Dlk1 foram digeridos com a enzima de restrição *EcoRI* e, após uma nova purificação, foi realizada uma reação de ligação dos vetores com os transcritos de Dlk1. O produto da ligação foi plaqueado, as colônias foram isoladas e, para confirmação da clonagem, foi realizada uma reação de PCR de colônia utilizando *primers* para os vetores, ou seja, os *primers* T7 e SP6 para amplificar os transcritos clonados no vetor pcDNA3 e HNF e pLPCX-R para amplificar os transcritos clonados no vetor pCXN2.

Figura 45 - Amplificação de Dlk1

Para a amplificação de Dlk1 foram utilizadas diferentes temperaturas de anelamento, como está indicado acima do gráfico. Além da banda esperada de 1,5kb, também foi encontrada uma outra banda de ~1,4kb.

Dos seqüenciamentos do transcrito de 1,5Kb clonados em pcDNA3, após a montagem dos contigs utilizando os programas Phred/Phrap (Figura 46-A), foi possível identificar um cDNA clonado na orientação correta e que apresentava apenas uma mutação onde uma arginina foi alterada por uma histidina (Figura 46-B), que foi então selecionado para transfectar as linhagens MTWT e mutantes. Ao contrário do vetor pcDNA3, não foi obtido sucesso em clonar Dlk1 no vetor pCXN2.

A

>Contigl-Dlk1

cgcaaccagaagccagcgagccccgggAGCAGCCCCCTGCACCGCCTCCGCTCCCCGGACCGCGACCCAGGCCGCCCGG
 AGATGATCGCGACCGGAGCCCTCCTGCGCGTCTCTTGCTCCTGCTGGCTTTCGGCCACAGCACCTATGGGGCTGAATGC
 GACCCACCCTGTGACCCCCAGTATGGATTCTGCGAGGCTGACAATGTCTGCAGGTGCCATGTTGGCTGGGAGGGTCCCCCT
 CTGTGACAAGTGTGTAACAGCCCCCTGGCTGTGTCAATGGAGTCTGCAAGGaaCCATGGCAGTGCATCTGCAAGGATGGCT
 GGGACGGGAaATTCTGCGAAATAGACGTTCTGGGCTTGCACTCAACCCCCCTGCGccAACAATGGAACTTGCGTGGACCTG
 GAGAAAGGCCAGTACGAATGCTCCTGCACACCTGGATTCTCTGGAAAGGACTGCCAGCACAAAGGCTGGGCCCTGCGTAAT
 CAATGGTTCTCCCTGCCAGCACGGAGGCGCTGCGTGGATGATGAGGGCCAGGCcTCGCATGCTTCTGCTGTGCCCCC
 CTGGCTTCTCAGGCAACTTCTGTGAGATCGTAGCCGCAACCAACAGCTGTACCCCTAACCCATGCGAGAACGATGGCGTC
 TGCACCGACATCGGGGGTGACTTCCGTTGCCGCTGCCCAGCTGGATTCTGTCGACAAGACCTGCAGCCGCCCGGTGAGCAA
 CTGCGCCAGTGGCCCCGTGCCAGAACGGGGGCACCTGCCAGCACACCCAGGCCATCTGCTTACCATCCTGGGCGTGC
 TCACCAGCTGGTGGTGTGGGCACCGTGGCCATCGTCTTTCTCAACAAGTGCAGAACCTGGGTGTCCAACCTGCGCTAC
 AACCACATGCTTCGCAAGAAGAAGAACCTCCTGTTGCAGTATAACAGCGCGAGGAGCTGGCGGTCAATATCATCTTCCC
 CGAGAAGATTGACATGACCACTTTCAacaaggaggctggtgatgagggc

B

Query	92	MIATGAXXXXXXXXXXAFGR	STYGAECDPPCDPQYGFCEADNVCRCHVGWEGPLCDKCVTA	271
		MIATGALLRVLLLLLLAFGR	STYGAECDPPCDPQYGFCEADNVCRCHVGWEGPLCDKCVTA	
Sbjct	1	MIATGALLRVLLLLLLAFGR	STYGAECDPPCDPQYGFCEADNVCRCHVGWEGPLCDKCVTA	60
Query	272	PGCVNGVCKEPWQCICKDGDGDKFCEIDVRACTSTPCANNGTCVDLEKGQYECSTPGFS		451
		PGCVNGVCKEPWQCICKDGDGDKFCEIDVRACTSTPCANNGTCVDLEKGQYECSTPGFS		
Sbjct	61	PGCVNGVCKEPWQCICKDGDGDKFCEIDVRACTSTPCANNGTCVDLEKGQYECSTPGFS		120
Query	452	GKDCQHKAGPCVINGSPCQHGGACVDDEGQASHASCLCPPGFSGNFCEIVAATNSCTPNP		631
		GKDCQHKAGPCVINGSPCQHGGACVDDEGQASHASCLCPPGFSGNFCEIVAATNSCTPNP		
Sbjct	121	GKDCQHKAGPCVINGSPCQHGGACVDDEGQASHASCLCPPGFSGNFCEIVAATNSCTPNP		180
Query	632	CENDGVCTDIGGDFRCRCPAGFVDKTCRSPVSNCASGPCQNGGTCLQHTQVSFECLCKPP		811
		CENDGVCTDIGGDFRCRCPAGFVDKTCRSPVSNCASGPCQNGGTCLQHTQVSFECLCKPP		
Sbjct	181	CENDGVCTDIGGDFRCRCPAGFVDKTCRSPVSNCASGPCQNGGTCLQHTQVSFECLCKPP		240
Query	812	FMGPTCAKKRGASPVQVTHLPSGYGLTYRLTPGVHELVPVQPEQHILKVSMEKLNKSTPL		991
		FMGPTCAKKRGASPVQVTHLPSGYGLTYRLTPGVHELVPVQPEQHILKVSMEKLNKSTPL		
Sbjct	241	FMGPTCAKKRGASPVQVTHLPSGYGLTYRLTPGVHELVPVQPEQHILKVSMEKLNKSTPL		300
Query	992	LTEGQAICFTIXXXXXXXXXXXXXAIVFLNKCETWVSNLRYNHMLRKKKNLLQLYNSGEE		1171
		LTEGQAICFTILGVLTSLVVLGTVAIVFLNKCETWVSNLRYNHMLRKKKNLLQLYNSGEE		
Sbjct	301	LTEGQAICFTILGVLTSLVVLGTVAIVFLNKCETWVSNLRYNHMLRKKKNLLQLYNSGEE		360
Query	1172	LAVNIIFPEKIDMTTFNKEAGDE	1240	
		LAVNIIFPEKIDMTTFNKEAGDE		
Sbjct	361	LAVNIIFPEKIDMTTFNKEAGDE	383	

Figura 46 – Análise da seqüência clonada do gene Dlk1

A: Saída do programa Phred/Phrap indicando o contig montado com todas as seqüências do transcrito analisado. Em azul escuro estão as bases cujo resultado do seqüenciamento possui uma boa qualidade, em azul claro uma qualidade média e em vermelho uma qualidade baixa. B: Saída do programa *BLAST* mostrando o alinhamento da seqüência proteína Dlk1 com a tradução do resultado obtido com o seqüenciamento. Para cada bloco de três linhas, na primeira linha (Query) está a seqüência de Dlk1 e na terceira linha (Sbjct) a seqüência obtida com seqüenciamento após sua tradução. Em vermelho está indicada a única alteração existente no qual uma Arginina foi alterada por uma Histidina.

Após extração do RNA das linhagens transfectadas e síntese do cDNA, foi realizado um experimento de *Real-Time PCR* para confirmação da expressão de Dlk1 na linhagem MTWT transfectada com o vetor vazio, MTWT expressando a forma canônica de Dlk1 (1,5Kb) e MTWT expressando a possível isoforma de Dlk1(1,2Kb) (Figura 47).

Na Figura 47-A pode ser observado que as duas populações que expressam Dlk1 (A03 e A08) possuem um nível de expressão cerca de 5 vezes menor do que PLJ enquanto que a linhagem MTWT e a linhagem MTWT transfectada com o vetor pcDNA3 vazio não expressam o gene Dlk1.

Uma vez constatado que a construção pcDNA3+Dlk1 estava levando à expressão do gene Dlk1, a segunda etapa consistiu em, através da utilização de *primers* para o gene MT, averiguar se o promotor do vetor pcDNA3 estava exercendo alguma influência na expressão do gene MT (Figura 47-B), sendo que foi observado uma diminuição de cerca de três vezes no nível de expressão de MT quando comparada com a linhagem MTWT.

Para verificar se a proteína Dlk1 sintetizada era funcional, foi realizada uma reação de *Real-Time PCR* utilizando *primers* para genes possivelmente eram regulados por Dlk1, como o gene Brca1, que havia sido identificado pelo software *PubGene* como sendo regulado por Dlk1. Embora tenha ocorrido a ativação de Brca1 pela ação da forma canônica de Dlk1 e uma ativação menor pela ação da isoforma de 1,2Kb, o vetor pcDNA3 exerceu um efeito semelhante em Brca1 (Figura 47-C).

Em resumo, foi detectado que os níveis de Dlk1 expressos pela construção pcDNA3+Dlk1 foram consideravelmente menores que aqueles induzidos pela ação de MT (

Figura 47-A) o que pode ter derivado da interferência entre os promotores do vetor PLJ e MTWT (Figura 47-B), e que o vetor pcDNA3, por si só, era capaz de exercer um efeito na célula nos mesmos níveis que a construção pcDNA3+Dlk1 (Figura 47-C). Como essa interferência do vetor, por si só, já invalidaria os resultados obtidos posteriormente, optou-se por clonar Dlk1 em um novo vetor.

Figura 47 – Real-Time PCR da superexpressão de Dlk1

Nos gráficos representados foram utilizados *primers* para Dlk1 (A), MT (B) e Brca1(C). No eixo da abscissa temos as linhagens a partir das quais se sintetizou os cDNAs que serviram como template, sendo elas: PLJ (PLJ), MTWT (MTWT), MTWT transfectada o vetor pcDNA3 vazio (pcDNA3), MTWT transfectada com o vetor pcDNA3 contendo a forma canônica de Dlk1 (A03) e MTWT transfectada com o vetor pcDNA3 com a isoforma menor de Dlk1 (A08). No eixo das ordenadas temos a variação da fluorescência em relação ao sinal da fluorescência de PLJ.

Ensaio anteriores do nosso laboratório, em outros sistemas, mostraram que o efeito produzido pelo vetor plasmideal pcDNA3 não é observado com o vetor retroviral pLPCX, portanto, este vetor foi selecionado para uma nova clonagem de Dlk1. Assim, foram desenhados novos *primers* para a amplificação de Dlk1 e as duas isoformas de Dlk1 foram amplificadas novamente, o produto foi digerido com *EcoRI* posteriormente ligados ao vetor pLPCX. Após o seqüenciamento inserto não foi identificada nenhuma alteração na seqüência do gene clonado.

4.7.1.1. Transfecção e produção das partículas virais

Após a confirmação da clonagem por PCR de colônia das bactérias transformadas e seqüenciamento dos clones obtidos, dois clones foram selecionados para serem transfectados em células empacotadoras ϕ nx-Ampho, sendo eles os clones E11 e F11, onde estão clonadas, respectivamente, a isoforma canônica e o *splicing* alternativo de Dlk1.

As construções contendo pLPCX+Dlk1 foram transfectadas (transfecção mediada por lipossomas) nas células empacotadoras ϕ nx-Ampho e os sobrenadantes virais coletados a cada 12h durante 72h. A partir dos sobrenadantes virais, foi realizado o protocolo de infecção para gerar populações de células MTWT

4.7.1.2. Curva de crescimento

Uma vez confirmada a superexpressão de Dlk1, foram realizados ensaios de curva de crescimento para se verificar se Dlk1 estava exercendo algum efeito no controle da divisão celular. Todos os experimentos foram realizados em duplicata, e para evitar possíveis erros de leitura no contador eletrônico, foram realizadas duas leituras de cada amostra.

Figura 48 – Curva de crescimento – Dlk1

Foi realizado o ensaio de curva de crescimento para as linhagens MTWT (MTWT), MTWT infectada com o vetor pLPCX vazio (pLPCX), MTWT expressando a isoforma canônica de Dlk1 (E11) e linhagem expressando o *splicing* alternativo de Dlk1 (F11). No eixo da abscissa está representado o tempo em que foi realizada a coleta, após o plaqueamento, e no eixo das ordenadas o número de células contadas na respectiva coleta.

Com base dos dados representados na 0, foi realizado o cálculo do tempo de dobramento das linhagens estudadas utilizando o programa Prism (Tabela 6 - Tempo de dobramento – Dlk1).

Tabela 6 - Tempo de dobramento – Dlk1

Foram analisadas as linhagens MTWT, MTWT infectada com vetor pLPCX vazio (pLPCX), MTWT infectada com vetor pLPCX expressando a isoforma canônica de Dlk1 (E11) e MTWT expressando a isoforma com *splicing* alternativo de Dlk1 (F11).

Linhagem	Tempo de dobramento (horas)	R ²
MTWT	14,54	0,9716
pLPCX	16,23	0,9878
E11	16,82	0,9974
F11	17,20	0,9688

Foi realizada um teste ANOVA entre os tempos da curva de crescimento e foi constatada uma diferença estatisticamente significativa ($p < 0,001$) entre o tempo de dobramento das linhagens E11 e MTWT, E11 e pLPCX, F11 e pLPCX, E11 e pLPCX, e não encontrada uma diferença significativa entre o tempo de dobramento das linhagens E11 e F11.

5. DISCUSSÃO

Para a busca de novos genes regulados pela ação do antígeno MT foram utilizadas três metodologias, as membranas comerciais, a construção de bibliotecas subtraídas de cDNA e o *CodeLink bioarray*. Apesar destas três metodologias permitirem a identificação de genes diferencialmente expressos, as mesmas foram realizadas com objetivos diferentes. Como as membranas comerciais e o RDA fornecem um número reduzido de genes, a sua finalidade consistia em identificar alguns genes que seriam caracterizados no modelo estudado. O *CodeLink*, como trabalha com uma lista maior de genes, foi utilizado para identificar novas vias de regulação ainda não relacionadas a MT.

5.1. Busca de genes regulados por MT através das membranas comerciais e do RDA

A primeira metodologia utilizada foi as membranas comerciais, e tinha como principal objetivo identificar genes cuja expressão ainda não havia sido caracterizada como sendo regulada pelas vias induzidas por MT. Dentre os genes selecionados, o que apresentou a maior diferença de expressão entre a linhagem controle (PLJ) e que expressa MT (MTWT) foi o gene *Dlk1 (delta-like 1 homolog)*, um gene homeótico membro da família EGF-like (*Epidermal Growth Factor*) que codifica para uma proteína que pode ser encontrada tanto na forma solúvel quanto associado à membrana, cujo produto havia sido caracterizado como um fator importante no controle da diferenciação de diversos tipos de células durante o desenvolvimento (Lee, Helman *et al.*, 1995; Ruiz-Hidalgo, Gubina *et al.*, 2002).

Utilizando a técnica RDA de subtração de cDNAs, foram construídas bibliotecas de cDNA enriquecidas para genes diferencialmente expressos nas linhagens PLJ e MTWT. Após a geração das bibliotecas e busca dos genes através de macroarranjos de cDNA (Figura 14), foram selecionados 76 cDNAs como candidatos a serem regulados pela ação de MT, sendo 22 da biblioteca *Forward* e 54 da biblioteca *Reverse*. Entretanto, após o seqüenciamento foi encontrada uma redundância de ~90%, ou seja, os 22 cDNAs selecionados na biblioteca *Forward* eram correspondentes a apenas 2 genes, enquanto que

somente seis genes distintos estavam representados na biblioteca *Reverse* (Tabela 2).

Devido à possibilidade de existirem falsos positivos entre os genes selecionados, antes de iniciar a seleção dos genes para uma futura análise funcional foram realizados experimentos de RT PCR e, como estes apresentaram diversos problemas, posteriormente foi utilizada a metodologia de *Real-Time PCR* para confirmar a expressão diferencial.

Após a análise por *Real-Time PCR* foi confirmada a expressão diferencial dos genes *Insig1 (Insulin Induced Gene 1)*, *AP-2*, *Lamb2 (laminin, beta 2)*, *UNR (Upstream of NRAS)*, *MN1 (Meningiome 1)* e *Dlk1 (Delta-Like homolog 1)*, dois quais *Lamb2*, um importante membro da matriz extracelular (Noda, Kitayama *et al.*, 1989; Wewer, Thornell *et al.*, 1997), e *Insig1*, ligada ao metabolismo de lipídeos (Li, Takaishi *et al.*, 2003; Attie, 2004), ainda não possuíam relação conhecida com o controle do ciclo celular.

Dentre os genes encontrados através do RDA e que possuem relação com o ciclo celular, *AP-2*, apesar de estar bem caracterizado no processo de diferenciação de diferentes tipos celulares, regula diversos genes relacionados ao controle do ciclo celular como *p21* (Zeng, Somasundaram *et al.*, 1997) e *c-myc* (Gaubatz, Imhof *et al.*, 1995), sendo que sua é comumente relacionada a parada do ciclo celular (Bar-Eli, 2001; Pfisterer, Ehlermann *et al.*, 2002). Curiosamente, *AP-2* foi encontrado como sendo induzido pela ação de MT o que, apesar de provavelmente não possuir relação com o processo de transformação celular mediada por MT, deve estar relacionado com a ativação do *enhancer* de MT, uma vez que *AP-2*, através de um mecanismo de *feedback*, ativa o *enhancer* do genoma de um outro poliomavírus, o vírus SV40 (Mitchell, Wang *et al.*, 1987).

O gene *UNR* foi descrito inicialmente como um transcrito que se localizava imediatamente *upstream* ao gene *N-ras* do genoma de diversos mamíferos, indicando a existência de um mecanismo de regulação comum aos dois genes (Jeffers, Paciucci *et al.*, 1990; Boussadia, Amiot *et al.*, 1997). *UNR* possui a capacidade de interagir com RNAs, dando indícios de que esteja envolvido com o metabolismo de mRNA (Triqueneaux, Velten *et al.*, 1999). Assim, a indução de *UNR* por MT pode estar ligada à transcrição de *N-ras* ou, em alguma instância, com o metabolismo de mRNA.

O gene MN1 de camundongo possui 93% de identidade com o seu homólogo humano, que consiste em uma proteína nuclear, expressa ubiquamente, que funciona como um co-ativador da transcrição e, até onde foi descrito, relacionada a processos de desenvolvimento (Meester-Smoor, Vermeij *et al.*, 2005). Curiosamente, a translocação do gene Tel (ETV6) pode gerar uma proteína fusionada com MN1, que faz com que MN1 mude de um repressor para um fator de transcrição com características oncogênicas, sendo que meningiomas (Buijs, Sherr *et al.*, 1995) e leucemias mielóides (Buijs, Van Rompaey *et al.*, 2000) já foram atribuídas a essa translocação.

Dentre os genes candidatos para uma análise funcional, os que se mostraram mais interessantes foram os genes MN1 e Dlk1. O gene MN1 porque foi completamente inibido pela ação de MT (Figura 33), ainda não foi caracterizado em camundongos, o seu homólogo humano é pouco estudado e possui a característica de poder agir como repressor e indutor do processo tumoral. Portanto, a identificação dos fatores que levaram MT a inibir MN1 levaria não só a uma melhor compreensão dos mecanismos controlados por MT, mas também iria gerar dados fundamentais para a compreensão do papel fisiológico de MN1 no contexto celular.

Assim como MN1, o gene Dlk1 também foi completamente inibido pela atividade de MT (Figura 44) e, o fato dele ter sido o gene mais encontrado no RDA e o mais expresso tanto nas membranas comerciais quanto no *CodeLink*, despertou interesse acerca do seu papel na transformação mediada por MT.

O primeiro passo idealizado para os ensaios funcionais constituiu-se em inibir a atividade dos genes MN1 na linhagem PLJ e superexpressar Dlk1 e MN1 na linhagem MTWT e nas linhagens que expressavam mutantes do gene MT.

A inibição do gene MN1 foi realizada através da metodologia de interferência de RNA (RNAi), sendo que os vetores lentivirais contendo as seqüências do RNAi foram construídas a partir de uma colaboração com o Dr. Ricardo Corrêa, que na época encontrava-se no *Salk Institute for Biological Studies*. Após a produção das partículas virais e infecção da linhagem PLJ foi possível identificar uma eficiência de infecção de aproximadamente 100% através do EGFP presente no vetor lentiviral (Figura 40). A confirmação da inibição de MN1 foi realizada através de ensaios de *Real-Time PCR*, onde foi constatada a completa inibição de MN1 na linhagem PLJ (Figura 41).

Para estudar se a perda de MN1 acarretaria em alguma alteração que pudesse transformar a linhagem PLJ, foi realizada uma curva de crescimento e um plaqueamento em substrato semi-sólido, que mimetiza o crescimento da célula em um animal (Figura 42 e Figura 43). Em ambos os ensaios, não foi observado nenhuma alteração ao se comparar a linhagem PLJ com a linhagem PLJ expressando o RNAi de MN1. O fato de não ter sido observado alteração nas características fenotípicas de PLJ não descarta a possibilidade de MN1 ter algum papel no controle do ciclo celular ou na atividade de MT, sendo que MN1 pode fazer parte de um processo no qual a sua ausência seria encoberta por genes diferentes com atividades semelhantes.

O próximo passo consistiu em testar o feito da superexpressão de MN1 na linhagem MTWT e, posteriormente, nos mutantes para MT. Como MN1 ainda não havia sido caracterizado, a primeira metodologia utilizada foi o RACE (*Rapid amplification of cDNA ends*) que, apesar de ter amplificado um pequeno fragmento da porção 3' do gene, não amplificou a porção 5' do gene. Foram desenhados *primers* baseados nas predições do gene MN1 e em ESTs, de camundongo, depositadas nos bancos, os quais, associados ao RACE, permitiram a amplificação de um fragmento de ~4Kb (Figura 37). O fragmento da porção 5' já foi amplificado, mas ainda não se obteve sucesso em sequenciar esse fragmentos (Figura 39). Assim que o fragmento for seqüenciado, futuros projetos poderão desenhar *primers* para amplificação do transcrito completo do gene MN1, deposição do mesmo no *GeneBank* e superexpressão de MN1 em linhagens que expressam MT selvagem e mutante.

Como o gene Dlk1 já havia sido caracterizado, foram desenhados *primers* para amplificar sua ORF, sendo que a amplificação de Dlk1 gerou duas bandas no gel de agarose, o que acreditava-se ser um *splicing* alternativo (Figura 45). Em concordância com a nossa hipótese, recentemente foi publicado que a isoforma encontrada se tratava de um *splicing* alternativo de Dlk1 (Deiuliis, Li *et al.*, 2006).

Posteriormente, Dlk1 canônica e sua isoforma foram clonados nos vetores pcDNA3 e pCXN2, sendo que só foi obtido sucesso na clonagem utilizando-se o vetor pcDNA3. Após as células PLJ terem sido transfectadas, a expressão de Dlk1 foi confirmada através de experimentos de *Real-Time PCR* onde foi constatado que Dlk1 estava sendo expresso na linhagem MTWT

(Figura 47-A). Para assegurar que não havia interferência entre os promotores do vetor PLJ e pcDNA3, foi realizado um *Real-Time PCR* utilizando *primers* para MT, onde foi constatada uma inibição de mais de 50% na atividade do vetor PLJ (Figura 47-B). Posteriormente, para testar se o vetor pcDNA3 estava exercendo alguma influência na expressão dos genes regulados por Dlk1, foi realizado um experimento de *Real-Time PCR* com *primers* para um gene identificado pelo *PubGene* como sendo regulado por Dlk1, o gene *Brca1* (*breast cancer 1*), sendo constatado que o efeito do vetor era semelhante ao efeito de Dlk1 clonado (Figura 47-C). Esses resultados foram suficientes para concluirmos que um novo vetor deveria ser utilizado para a superexpressão de Dlk1 na linhagem MTWT.

O vetor escolhido para a nova clonagem de Dlk1 foi o vetor retroviral pLPCX, sendo que após a clonagem, produção de partículas virais e seleção por puromicina, foram realizados experimentos de *Real-Time PCR* com *primers* para Dlk1, MT e *Brca1* sendo que não foi observado efeito significativo do vetor no nível de expressão desses genes.

Para analisar os efeitos de Dlk1 na linhagem PLJ foi realizada uma curva de crescimento para a linhagem MTWT, MTWT infectada com vetor vazio e MTWT expressando cada uma das duas isoformas de Dlk1 (Figura 48). De acordo com a curva, ainda que o vetor tenha exercido um certo efeito na taxa de divisão celular, houve um aumento estatisticamente significativa no tempo de dobramento das linhagens que expressavam as duas isoformas de Dlk1, quando comparadas com a linhagem expressando o vetor vazio (Tabela 6). Este aumento na taxa de divisão celular, apesar de estar em contradição com o fato de já ter sido demonstrado que a expressão de Dlk1 leva à uma inibição da taxa de proliferação celular em diversos sistemas (Li, Forman *et al.*, 2005), incluindo células Balb-3T3 (Baladron, Ruiz-Hidalgo *et al.*, 2002), de onde a linhagem A31 é derivada, está de acordo com outros trabalhos que mostram um aumento da taxa de proliferação induzida por Dlk1 (Raza, Alvi *et al.*, 1997; Yin, Xie *et al.*, 2006). Este resultado sugere que a expressão de MT levou à regulação de vias de transdução de sinal onde o sinal desencadeado por Dlk1 passou a ter um efeito de indução da proliferação. Uma possibilidade é que Dlk1 esteja interagindo com a via de Ras/MAPK, uma vez que Dlk1 é capaz de ativar vias de MAPK em resposta ao estímulo de IGF-1 (*insulin-like growth factor*

1)/Insulina (Ruiz-Hidalgo, Gubina *et al.*, 2002). Futuros projetos poderão investigar a interação de Dlk1 com as vias de transdução de sinal mediadas por MT e, dessa maneira, descrever novos mecanismos que relacionam a atividade de Dlk1 com o controle da divisão celular.

5.2. Análise dos genes regulados por MT através do CodeLink Bioarrays

Após a busca inicial de genes através de membranas comerciais e da metodologia de RDA, passou-se a utilizar a plataforma CodeLink para realizar, além da identificação de alguns genes relacionados a MT, uma busca de novas vias de sinalização reguladas por MT (Figura 16). Durante todo o projeto, a análise dos dados do CodeLink passou por diversas etapas que incluíram a escolha do método de normalização (*LOWESS*)(Figura 18), seleção do melhor método para seleção dos genes diferencialmente expressos (*e-Bayes*) e, por fim, o teste de diversos programas que se propunham a identificar possíveis vias de sinalização entre os genes selecionados, sendo que o programa que apresentou o resultado mais satisfatório foi o *PubGene*, que fornecia uma pontuação para as interações entre cada par de genes analisados (Figura 20).

Em seguida, foram desenhados *primers* para os genes selecionados e, posteriormente, foi calculada a eficiência destes *primers*. Alguns autores afirmam que *primers* com tamanho em torno de 20pb e que amplificam fragmentos menores que 150pb, como aqueles que foram utilizados neste projeto, possuem eficiência próxima a 100% e, sendo assim, não precisariam ter sua eficiência testada (Livak e Schmittgen, 2001). Entretanto, atualmente tem-se exigido que mesmo os *primers* que obedecem estes requisitos tenham sua eficiência testada. Assim, utilizando um protocolo já descrito (Pfaffl, 2001), antes de iniciarmos os ensaios de *Real-Time PCR* foi analisada a eficiência dos *primers*, sendo que cerca de 90% dos *primers* estavam dentro dos padrões definidos, ou seja, eficiência entre 80% e 110% e um coeficiente de correlação (R^2) maior que 0,85. Após o teste de eficiência, os *primers* que satisfizeram os padrões pré-determinados tiveram o nível de expressão de seus respectivos genes analisado através de experimentos de *Real-Time PCR*, na linhagem controle (PLJ), na linhagem que expressa MT

selvagem (MTWT) e nos mutantes de MT para a via de PI-3K(MT315), PLC- γ (MT322), PI-3K/PLC- γ (MT2F) e Ras/MAPK (250phe).

Após a análise dos *primers*, iniciou-se os experimentos de *Real-Time PCR*. Estes experimentos tinham como objetivo não só caracterizar novas vias reguladas por MT, mas também utilizar o modelo MT selvagem/mutantes para identificar possíveis vias de regulação de genes ainda não caracterizados. Para caracterizar estas vias, a abordagem utilizada foi a construção de Redes de Relevância (*Relevance Networks*), que constituem redes através das quais se infere associações entre os genes de acordo com o seu perfil de expressão nas condições estudadas. Para medir a suposta dependência entre os genes pode ser utilizada a Informação Mútua (IM) ou Correlação de Pearson (CP). Enquanto a IM analisa os perfis de expressão como um todo, a PC analisa as correlações pra cada par de genes. Ainda que a IM apresente algumas vantagens em relação a PC como uma melhor quantificação da dependência e o fato de não ser influenciada por *outliers*, optou-se por utilizar a PC, pois essa permite que sejam identificadas correlações negativas entre os genes analisados, além de já ter sido relatado que as correlações obtidas pelos dois métodos não apresentam diferenças significativas (Lindlof e Lubovac, 2005). Para minimizar a presença de *outliers*, foi utilizado o algoritmo *leave-one-out*, no qual os dados de cada uma das condições são retirados alternadamente e o conjunto de dados restante é novamente testado para as condições de exclusão ($|r| > 0,7$ e $p < 0,05$).

Antes de ser realizada a análise com todos os genes selecionados e ser desenvolvido um programa para este fim, foi realizada a análise de um número reduzido de genes, com o auxílio do programa *Prism*, para validar a abordagem utilizada. O primeiro diagrama identificado está representado na Figura 49.

Figura 49 – Primeiro mapa de correlação identificado

Os números indicam os coeficientes correlação (r) entre cada par de gene analisado.

Neste diagrama, estão representados os genes Nek2 (*NIMA-related kinase 2*), uma serina/treonina quinase relacionada com a separação do centrômero durante a mitose, controle da condensação da cromatina e controle do ciclo celular (Di Agostino, Rossi *et al.*, 2002; Hayward e Fry, 2006); Cdc2a (*cell division cycle 2 homolog A*), uma outra serina/treonina quinase relacionada com o controle do ciclo celular, principalmente na transição G2/M, e que se torna ativa quando complexada com Ciclina B (Kim, Buchholz *et al.*, 1994; Yu, Wang *et al.*, 2005); Fanca (*Fanconi anemia, complementation group A*), uma proteína relacionada com a resposta a dano celular e cuja alteração está relacionada com a anemia de Falconi (Noll, Battaile *et al.*, 2002; Yang, Herceg *et al.*, 2005) e Brca1, conhecida por estar relacionada ao câncer de mama familiar e por interagir e regular diversas proteínas incluindo aqueles codificados por oncogenes, genes supressores de tumor e outros (Deng, 2006). Foi descrito que Brca1 regula a expressão de Nek2 (Wang, Wang *et al.*, 2004) e Cdc2a (Bae, Rih *et al.*, 2005; Chabalier, Lamare *et al.*, 2006) e, apesar de não ser conhecido o mecanismo de regulação em comum, já foi relatada a interação de Brca1 com Fanca (Folias, Matkovic *et al.*, 2002). Brca1 é conhecido por ser ativado por PI-3K via Akt (Altioik, Batt *et al.*, 1999), e nas vias indicadas neste mapa há uma diminuição da expressão dos genes nas linhagens MT315 e MT322, que possuem a via de PI-3K inativada (Figura 30). Portanto, esta via é regulada por Brca1 que, conseqüentemente, regula Nek2, Cdc2a e Fanca. O fato das vias representadas na Figura 49 estarem descritas na literatura e a

modulação da expressão dos respectivos genes ser explicada pelas vias de sinalização contidas nos mutantes deu fortes indícios de que a metodologia utilizada seria uma ferramenta poderosa na identificação de novas vias de sinalização reguladas por MT.

Quando foram analisados todos os mapas de correlação, as correlações Brca1, Nek2, Cdc2a e Fanca estavam presentes no mapa de correlação da Figura 32-A, sendo que os genes nesse mapa têm sua atividade inibida na ausência da via de PI-3K (MT315 e MT322) e não sofrem alteração na inibição da via de Ras/MAPK (250phe). Os demais genes estão relacionados com o controle do ciclo celular e, em alguma instância, todos têm sua atividade regulada pelas vias de sinalização ativada por PI-3K. Neste mapa de correlação, somente dois genes não estão descritos como relacionados com as demais correlações, o gene Oip5 (*Opa interacting protein 5*) e W09175.

Oip5 é uma proteína de células epiteliais que interage com outras proteínas da membrana celular, participando do processo de invasão. Os mecanismos relacionados à regulação e ativação de Oip5 ainda não são conhecidos, entretanto, um estudo já relatou a expressão irregular de Oip5 em 50% dos tumores gástricos analisados, indicando uma possível relação com Oip5 e a progressão no ciclo celular. No mapa, Oip5 apresenta uma correlação direta com Mad2l1 (*MAD2 mitotic arrest deficient-like 1*), Chek2 (*CHK2 checkpoint homolog*) e Rad51 (*BRCA1/BRCA2-containing complex*), que estão relacionados à parada do ciclo celular induzida por dano no DNA ou reparo do DNA (Miyazaki, Bressan *et al.*, 2004; Nevanlinna e Bartek, 2006; Yuan, Xu *et al.*, 2006); portanto, existe a possibilidade de Oip5 estar ligado ao processo de parada do ciclo celular através de um mecanismo induzido por PI-3K.

W09175 faz parte da lista dos genes ainda não caracterizados que estão presentes no CodeLink, sendo que ainda não foi atribuída função alguma a este gene. Foi realizada uma análise da sequência deste gene através de ferramentas de Bioinformática (Pfam, CDD, SAGE) que mostraram indícios de que W09175 esteja relacionado com o processo de replicação, o que poderia explicar sua correlação com Oip5.

Na Figura 32-B, está representado uma mapa de correlação contendo seis genes, sendo que destes apenas dois já estão caracterizados, Id1

(*inhibitor of DNA binding 1*) e Cdk4 (*cyclin-dependent kinase 4*). Dentre eles, foi demonstrado que PKC, ativado na via de PLC- γ , leva à expressão de Id1 (Kebebew, Treseler *et al.*, 2003) e Cdk4 (Colomer, Lopez-Girona *et al.*, 1994), o que está de acordo com a inibição destes genes na linhagem MT322. Tanto Cdk4 quanto Id1 são ativados pela via de PI-3K (Gao, Zhang *et al.*, 2003; Birkenkamp, Essafi *et al.*, 2007), entretanto, o mapa de correlação mostrou um efeito inverso, ou seja, a ativação de Cdk4 e Id1 quando houve a inibição de PI-3K, tanto na linhagem MT315 quanto na MT2F. Como já foi descrito que Id1 e Cdk4 também podem ser ativadas pela via de MAPK (Langenfeld e Langenfeld, 2004; Yang e Roy, 2004), e na linhagem 250phe houve um aumento na expressão dos genes, ao que parece está havendo um *feedback* negativo entre as vias de PI-3K e Ras/MAPK, de modo que a inibição de uma das duas vias acarreta no aumento do efeito da outra.

Em relação aos genes não caracterizados, NM_020616, Wdr70, W70351, NM_026152 e Cdk4, a análise de suas respectivas seqüências utilizando ferramentas de Bioinformática não identificou características como homologia com outras proteínas ou domínios conservados, sendo que as únicas características que podemos inferir estão relacionadas a Id1 e Cdk4. Id1 é uma proteína que forma heterodímeros com fatores de transcrição, podendo inibir a atividade destes fatores, que possui papel em processos como crescimento, divisão, senescência e diferenciação celular (Alani, Young *et al.*, 2001; Ling, Lau *et al.*, 2005; Hui, Cheung *et al.*, 2006; Jang, Han *et al.*, 2006), enquanto que Cdk4, assim como a Cdk2, descrita anteriormente, consiste em uma quinase dependente de ciclinas com papel fundamental no controle do ciclo celular (Kim, Yang *et al.*, 2003). Ambos os genes possuem papel no controle do ciclo celular via ativação de fatores de transcrição, sendo Cdk4 de uma forma indireta e Id1 de uma forma direta. Assim, como os genes descritos possuem o mesmo perfil de regulação que Id1 e Cdk4, possivelmente eles possuem algum papel no controle da proliferação celular, e sua regulação está relacionada com as três principais vias ativadas por MT, ou seja, PI-3K, PLC- γ e Ras/MAPK.

Adam12 (*ADAM metalloproteinase domain 12*) é um membro da família de metaloproteínases, que se encontram ancoradas à membrana e apresentam uma estrutura complexa contendo domínios de desintegrina, rico em cistina, trans-membrânico e citoplasmático (Schlondorff e Blobel, 1999;

Le Pabic, L'helgoualc'h *et al.*, 2005), sendo que a atividade de Adam12 tem sido relacionada a diversos tumores como de mama (Kveiborg, Frohlich *et al.*, 2005), próstata (Peduto, Reuter *et al.*, 2006) e bexiga (Frohlich, Albrechtsen *et al.*, 2006). Foi descrito que tanto a via de PI-3K quanto a de Ras/MAPK induzem Adam12. Na rede da Figura 32-C observa-se um aumento da expressão de Adam12 nos mutantes para PI-3K e Ras/MAPK. Portanto, assim como no mapa da Figura 32-B, deve estar ocorrendo um *feedback* negativo entre as vias de PI-3K e Ras/MAPK onde a inibição de uma das vias permite o funcionamento da outra via.

O gene Reck (*reversion-inducing-cysteine-rich protein with kazal motifs*) codifica para uma glicoproteína que funciona como um regulador negativo da metaloproteinase-9, uma enzima da matriz extracelular relacionada à invasão e metástase. Recentemente, diversos trabalhos têm sido realizados no sentido de analisar a expressão do gene RECK em tumores humanos como hepatocarcinoma (Furumoto, Aii *et al.*, 2001) e carcinoma de mama (Span, Sweep *et al.*, 2003), sendo que níveis elevados RECK nestes tumores correlacionaram com um prognóstico favorável para os pacientes (Furumoto, Aii *et al.*, 2001; Masui, Doi *et al.*, 2003; Van Der Jagt, Sweep *et al.*, 2006). Apesar de Adam12 e Reck pertecerem à via de degradação da matriz extracelular, a correlação encontrada consiste na primeira descrição de um perfil de regulação semelhante e, conseqüentemente, um possível mecanismos de regulação em comum entre Reck e Adam12.

Tmem158 (*transmembrane protein 158*) foi identificado, recentemente, como sendo induzido na senescência mediada por Ras (Barradas, Gonos *et al.*, 2002). Apesar de haver indícios de que Tmem158 esteja agindo como um gene supressor de tumor, seu mecanismo de ação ainda não foi identificado (Silva, Silva *et al.*, 2006).

O gene fos é um proto-oncogene que codifica para a proteína Fos que, ao formar dímeros com a proteína Jun, forma o fator de transcrição AP-1, que está implicado na regulação da proliferação celular, diferenciação e transformação (Shaulian e Karin, 2001). Apesar da via de PI-3K e Ras/MAPK serem conhecidas por induzir a transcrição de Fos (Raivich e Behrens, 2006; Neto, Coelho *et al.*, 2007), o resultado obtido na Figura 32-C está de acordo com dados anteriores do nosso laboratório que mostraram que a ativação simultânea das vias de PI-3K e Ras/MAPK induzida por MT não

leva a expressão de Fos (Rameh e Armelin, 1991). Como não foi ouve a expressão de Fos nas linhagens MT322 e MT2F, e Fos foi ativado nas linhagens MT315 e 250phe, assim como em Adam12, provavelmente existe *crosstalking* entre as vias PI-3K e Ras/MAPK de modo que, a ativação de ambas as vias, leve a uma inibição da expressão de Fos.

Os demais genes da via, *Ier5l* (*immediate early response 5-like*) e *Fzd2* (*frizzled homolog 2*), ainda não foram caracterizados, sendo que sua possível relação tanto com os sinais mediados por MT como com os genes *Temem158*, *Adam12*, *Fos* e *Reck* foi primeiramente descrita no presente trabalho.

A interação indicada na Figura 32-D está claramente relacionada com a atividade de PLC- γ , uma vez que o sinal correspondente aos genes está ausente nas linhagens PLJ, MT322 e MT2F. *Uhrf1* (*ubiquitin-like, containing PHD and RING finger domains, 1*) está ligado ao controle da expressão gênica através de alterações epigenéticas, estando envolvido no controle da proliferação celular. Além disso, sua expressão irregular já foi descrita em tumores (Mousli, Hopfner *et al.*, 2003; Jeanblanc, Mousli *et al.*, 2005). A inibição de *Uhrf1* nas linhagens MT322 e MT2F é facilmente explicada com base em relatos de que o seqüestro de cálcio leva à inibição de *Uhrf1* (Abbady, Bronner *et al.*, 2005).

A proteína MCM3 (*minichromosome maintenance deficient 3*) possui um papel no processo de replicação e sua superexpressão já foi relacionada com o processo neoplásico (Ha, Shin *et al.*, 2004; Shechter e Gautier, 2004). Entretanto, até o momento não foi descrito um mecanismo de regulação comum entre MCM3 e PLC- γ ou *Uhrf1*.

A última via identificada (Figura 32-E) funcionou como um mecanismo para validar o método, uma vez que a relação entre Ciclina D1 e Myc já foi descrita em diversos sistemas, bem como sua dependência das vias ativadas por MT (Gera, Mellinghoff *et al.*, 2004; Yu, Ciemerych *et al.*, 2005; Marderosian, Sharma *et al.*, 2006).

Em resumo, os experimentos de *Real-Time PCR* com os resultados provenientes do CodeLink permitiram identificar possíveis mecanismos de regulação desconhecidos até então, bem como inferir possíveis regulações de genes ainda não caracterizados, sendo que essas predições foram embasadas na identificação de outras vias de regulação já descritas na literatura.

6. CONCLUSÕES

- . Através das membranas comerciais, foi possível isolar 22 genes como sendo regulados por MT, dos quais 17 eram reprimidos e 5 induzidos por MT.
- . Após o rastreamento de 750 clones, obtidos através das bibliotecas subtraídas de cDNA, foram identificados 6 genes reprimidos e 2 induzidos por MT.
- . O RDA permitiu a identificação do gene MN1 (*meningioma 1*), que até então não havia sido caracterizado.
- . A completa inibição de MN1 foi obtida através da construção e expressão de um RNA de interferência (RNAi), específico para MN1, na linhagem PLJ.
- . Não foram observadas alterações fenotípicas (morfologia e taxa de crescimento) após *knock-down* de MN1 pela metodologia RNAi.
- . Para amplificar a sequência completa de MN1 foi utilizada a metodologia RACE associada com *primers* desenhados ao longo da sequência de uma predição do gene MN1. Apesar da sequência completa do gene aparentemente ter sido amplificada, ainda não foi possível confirmar esse resultado através de sequenciamento.
- . A superexpressão de Dlk1, mediada pelo vetor retroviral pLPCX, levou a um aumento da taxa de divisão celular na linhagem MTWT.
- . O programa desenvolvido para a análise do CodeLink *bioarrays* permitiu a normalização e seleção de 1.129 genes que apresentavam uma diferença estatisticamente significativa ($p < 0,05$) nas condições analisadas.
- . O programa *Pubgene* possibilitou a identificação de possíveis vias de regulação após a análise dos dados do CodeLink, permitindo selecionar 29 genes para terem sua expressão diferencial avaliada por experimentos de *Real-Time PCR*.
- . Os programa *FatiGO* classificou a lista dos genes de acordo com o processo biológico em que estão envolvidos, sendo que foram selecionados 34 genes relacionados ao controle do ciclo celular para análise por *Real-Time PCR*.
- . Dos genes que não possuíam classificação de acordo com o *FatiGO*, com o auxílio do programa *Blast*, foram selecionados 14 genes novos para serem analisados por *Real-Time PCR*.

- . Somando os genes selecionados pelo *FatiGO*, *PubGene*, *Blast*, RDA e diversos genes relacionados ao ciclo celular, foram sintetizados *primers* para 86 genes, dos quais 64 satisfizeram as condições de exclusão
- . Após a análise das 4.096 interações, foram identificadas 92 correlações possíveis, envolvendo 35 dos genes selecionados.
- . Das interações encontradas, 13 delas envolviam genes não caracterizados, sendo que essas interações permitiram inferir possíveis funções e mecanismos de regulação destes genes.
- . O fato da maioria das correlações identificadas estarem descritas na literatura, validou a metodologia utilizada e comprovou que o sistema MT selvagem/mutantes consiste em uma ferramenta poderosa na caracterização de novos genes e eventos relacionados ao controle do ciclo celular e da transformação maligna.

7. BIBLIOGRAFIA

Aaronson, S. A. e G. J. Todaro. Development of 3T3-like lines from Balb-c mouse embryo cultures: transformation susceptibility to SV40. J Cell Physiol, v.72, n.2, p.141-8. 1968.

Abbadly, A. Q., C. Bronner, *et al.* TCR pathway involves ICBP90 gene down-regulation via E2F binding sites. Biochem Pharmacol, v.70, n.4, Aug 15, p.570-9. 2005.

Abram, C. L. e S. A. Courtneidge. Src family tyrosine kinases and growth factor signaling. Exp Cell Res, v.254, n.1, Jan 10, p.1-13. 2000.

Al-Shahrour, F., R. Diaz-Uriarte, *et al.* FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. Bioinformatics, v.20, n.4, Mar 1, p.578-80. 2004.

Alani, R. M., A. Z. Young, *et al.* Id1 regulation of cellular senescence through transcriptional repression of p16/Ink4a. Proc Natl Acad Sci U S A, v.98, n.14, Jul 3, p.7812-6. 2001.

Altiock, S., D. Batt, *et al.* Heregulin induces phosphorylation of BRCA1 through phosphatidylinositol 3-Kinase/AKT in breast cancer cells. J Biol Chem, v.274, n.45, Nov 5, p.32274-8. 1999.

Amalfitano, A., L. G. Martin, *et al.* Different roles for two enhancer domains in the organ- and age-specific pattern of polyomavirus replication in the mouse. Mol Cell Biol, v.12, n.8, Aug, p.3628-35. 1992.

Arena, S., S. Benvenuti, *et al.* Genetic analysis of the kinome and phosphatome in cancer. Cell Mol Life Sci, v.62, n.18, Sep, p.2092-9. 2005.

Attie, A. D. Insig: a significant integrator of nutrient and hormonal signals. J Clin Invest, v.113, n.8, Apr, p.1112-4. 2004.

Ausubel, F. M., R. Brent, *et al.* Current Protocols in Molecular Biology: Wiley. 2001

Bae, I., J. K. Rih, *et al.* BRCA1 regulates gene expression for orderly mitotic progression. Cell Cycle, v.4, n.11, Nov, p.1641-66. 2005.

Baladron, V., M. J. Ruiz-Hidalgo, *et al.* The EGF-like homeotic protein dlk affects cell growth and interacts with growth-modulating molecules in the yeast two-hybrid system. Biochem Biophys Res Commun, v.291, n.2, Feb 22, p.193-204. 2002.

Bar-Eli, M. Gene regulation in melanoma progression by the AP-2 transcription factor. Pigment Cell Res, v.14, n.2, Apr, p.78-85. 2001.

Barradas, M., E. S. Gonos, *et al.* Identification of a candidate tumor-suppressor gene specifically activated during Ras-induced senescence. Exp Cell Res, v.273, n.2, Feb 15, p.127-37. 2002.

Bashir, T., R. Horlein, *et al.* Cyclin A activates the DNA polymerase delta - dependent elongation machinery in vitro: A parvovirus DNA replication model. Proc Natl Acad Sci U S A, v.97, n.10, May 9, p.5522-7. 2000.

Beijerinck, M. W. Over een Contagium vivum fluidum alsoorzaak van de Vlekziekte der tabaksbladen. Versl. Gew. Verg. Wis- en Natuurk. Afd, v.VII, p.229-935 (English translation published in 1942: Concerning the mosaic disease of tobacco. *Phytopathol. Classics*). 1898.

Benjamin, T. L. Polyoma virus: old findings and new challenges. Virology, v.289, n.2, Oct 25, p.167-73. 2001.

Benjamini, Y. e Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J.R.Statist.Soc.B., v. 57(1), p.289-300. 1995.

Berger, J. R., R. M. Levy, *et al.* Predictive factors for prolonged survival in acquired immunodeficiency syndrome-associated progressive multifocal leukoencephalopathy. Ann Neurol, v.44, n.3, Sep, p.341-9. 1998.

Berger, J. R., L. Pall, *et al.* Progressive multifocal leukoencephalopathy in patients with HIV infection. J Neurovirol, v.4, n.1, Feb, p.59-68. 1998.

Berridge, M. J. Inositol trisphosphate and diacylglycerol as second messengers. Biochem J, v.220, n.2, Jun 1, p.345-60. 1984.

Berridge, M. J. e R. F. Irvine. Inositol trisphosphate, a novel second messenger in cellular signal transduction. Nature, v.312, n.5992, Nov 22-28, p.315-21. 1984.

Birkenkamp, K. U., A. Essafi, *et al.* FOXO3a induces differentiation of Bcr-Abl-transformed cells through transcriptional down-regulation of Id1. J Biol Chem, v.282, n.4, Jan 26, p.2211-20. 2007.

Blaikie, P. A., E. Fournier, *et al.* The role of the Shc phosphotyrosine interaction/phosphotyrosine binding domain and tyrosine phosphorylation sites in polyoma middle T antigen-mediated cell transformation. J Biol Chem, v.272, n.33, Aug 15, p.20671-7. 1997.

Bolstad, B. M., R. A. Irizarry, *et al.* A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics, v.19, n.2, Jan 22, p.185-93. 2003.

Bonnet, F. e P. Morlat. [Cancer and HIV infection: any association?]. Rev Med Interne, v.27, n.3, Mar, p.227-35. 2006.

Boussadia, O., F. Amiot, *et al.* Transcription of unr (upstream of N-ras) down-modulates N-ras expression in vivo. FEBS Lett, v.420, n.1, Dec 22, p.20-4. 1997.

Boyle, W. J., T. Smeal, *et al.* Activation of protein kinase C decreases phosphorylation of c-Jun at sites that negatively regulate its DNA-binding activity. Cell, v.64, n.3, Feb 8, p.573-84. 1991.

Braun, B. S., R. Frieden, *et al.* Identification of target genes for the Ewing's sarcoma EWS/FLI fusion protein by representational difference analysis. Mol Cell Biol, v.15, n.8, p.4623-30. 1995.

Brewster, C. E., H. R. Glover, *et al.* pp60c-src binding to polyomavirus middle T-antigen (MT) requires residues 185 to 210 of the MT sequence. J Virol, v.71, n.7, Jul, p.5512-20. 1997.

Brown, M. T. e J. A. Cooper. Regulation, substrates and functions of src. Biochim Biophys Acta, v.1287, n.2-3, Jun 7, p.121-49. 1996.

Buijs, A., S. Sherr, *et al.* Translocation (12;22) (p13;q11) in myeloproliferative disorders results in fusion of the ETS-like TEL gene on 12p13 to the MN1 gene on 22q11. Oncogene, v.10, n.8, Apr 20, p.1511-9. 1995.

Buijs, A., L. Van Rompaey, *et al.* The MN1-TEL fusion protein, encoded by the translocation (12;22)(p13;q11) in myeloid leukemia, is a transcription factor with transforming activity. Mol Cell Biol, v.20, n.24, Dec, p.9281-93. 2000.

Campbell, B. A. e L. P. Villarreal. Functional analysis of the individual enhancer core sequences of polyomavirus: cell-specific uncoupling of DNA replication from transcription. Mol Cell Biol, v.8, n.5, May, p.1993-2004. 1988.

Campbell, K. S., K. R. Auger, *et al.* Identification of regions in polyomavirus middle T and small t antigens important for association with protein phosphatase 2A. J Virol, v.69, n.6, Jun, p.3721-8. 1995.

Campbell, S. L., R. Khosravi-Far, *et al.* Increasing complexity of Ras signaling. Oncogene, v.17, n.11 Reviews, Sep 17, p.1395-413. 1998.

Carmichael, G. G., B. S. Schaffhausen, *et al.* Carboxy terminus of polyoma middle-sized tumor antigen is required for attachment to membranes, associated protein kinase activities, and cell transformation. Proc Natl Acad Sci U S A, v.79, n.11, Jun, p.3579-83. 1982.

Carpenter, C. L., B. C. Duckworth, *et al.* Purification and characterization of phosphoinositide 3-kinase from rat liver. J Biol Chem, v.265, n.32, Nov 15, p.19704-11. 1990.

Cartwright, C. A., P. L. Kaplan, *et al.* Altered sites of tyrosine phosphorylation in pp60c-src associated with polyomavirus middle tumor antigen. Mol Cell Biol, v.6, n.5, May, p.1562-70. 1986.

Chabalier, C., C. Lamare, *et al.* BRCA1 downregulation leads to premature inactivation of spindle checkpoint and confers paclitaxel resistance. Cell Cycle, v.5, n.9, May, p.1001-7. 2006.

Chang, C., D. T. Simmons, *et al.* Identification and partial characterization of new antigens from simian virus 40-transformed mouse cells. J Virol, v.31, n.2, Aug, p.463-71. 1979.

Chen, M. H. e T. Benjamin. Roles of N-glycans with alpha2,6 as well as alpha2,3 linked sialic acid in infection by polyoma virus. Virology, v.233, n.2, Jul 7, p.440-2. 1997.

Cheng, S. H., R. Harvey, *et al.* Peptide antibodies to the human c-fyn gene product demonstrate pp59c-fyn is capable of complex formation with the middle-T antigen of polyomavirus. Embo J, v.7, n.12, Dec 1, p.3845-55. 1988.

Chirgwin, J. M., A. E. Przybyla, *et al.* Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry, v.18, n.24, Nov 27, p.5294-9. 1979.

Cleveland, W. S. Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Association, v.74, n.368, p.829-836. 1979.

Clever, J. L., D. Mirandar, Jr., *et al.* RNA structure and packaging signals in the 5' leader region of the human immunodeficiency virus type 1 genome. J Virol, v.76, n.23, Dec, p.12381-7. 2002.

Coffin, J. M., S. H. Hughes, *et al.* Retroviruses - Reverse transcriptase and the generation of retroviral DNA. Plainview, N.Y.: Cold Spring Harbor Laboratory Press. 1997. xv, 843 p. p.

Colomer, J., A. Lopez-Girona, *et al.* Calmodulin regulates the expression of cdks, cyclins and replicative enzymes during proliferative activation of human T lymphocytes. Biochem Biophys Res Commun, v.200, n.1, Apr 15, p.306-12. 1994.

Cooper, G. M. Oncogenes. Boston, USA: Jones and Barllet Publishers Inc. 1995

Courtneidge, S. A. e A. Heber. An 81 kd protein complexed with middle T antigen and pp60c-src: a possible phosphatidylinositol kinase. Cell, v.50, n.7, Sep 25, p.1031-7. 1987.

Courtneidge, S. A. e A. E. Smith. Polyoma virus transforming protein associates with the product of the c-src cellular gene. Nature, v.303, n.5916, Jun 2-8, p.435-9. 1983.

Creager, A. N., K. B. Scholthof, *et al.* Tobacco mosaic virus. Pioneering research for a century. Plant Cell, v.11, n.3, Mar, p.301-8. 1999.

Dawe, C. J., R. Freund, *et al.* Variations in polyoma virus genotype in relation to tumor induction in mice. Characterization of wild type strains with widely differing tumor profiles. Am J Pathol, v.127, n.2, May, p.243-61. 1987.

Deiuliis, J. A., B. Li, *et al.* Alternative splicing of delta-like 1 homolog (DLK1) in the pig and human. Comp Biochem Physiol B Biochem Mol Biol, v.145, n.1, Sep, p.50-9. 2006.

Deleu, L., A. Pujol, *et al.* Activation of promoter P4 of the autonomous parvovirus minute virus of mice at early S phase is required for productive infection. J Virol, v.73, n.5, May, p.3877-85. 1999.

Deng, C. X. BRCA1: cell cycle checkpoint, genetic instability, DNA damage response and cancer evolution. Nucleic Acids Res, v.34, n.5, p.1416-26. 2006.

Di Agostino, S., P. Rossi, *et al.* The MAPK pathway triggers activation of Nek2 during chromosome condensation in mouse spermatocytes. Development, v.129, n.7, Apr, p.1715-27. 2002.

Diehl, J. A., M. Cheng, *et al.* Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. Genes Dev, v.12, n.22, Nov 15, p.3499-511. 1998.

Dilworth, S. M. Cell alterations induced by the large T-antigens of SV40 and polyoma virus. Semin Cancer Biol, v.1, n.6, Dec, p.407-14. 1990.

Polyoma virus middle T antigen and its role in identifying cancer-related molecules. Nat Rev Cancer, v.2, n.12, Dec, p.951-6. 2002.

Doorbar, J. The papillomavirus life cycle. J Clin Virol, v.32 Suppl 1, Mar, p.S7-15. 2005.

Dorris, D. R., A. Nguyen, *et al.* Oligodeoxyribonucleotide probe accessibility on a three-dimensional DNA microarray surface and the effect of hybridization time on the accuracy of expression ratios. BMC Biotechnol, v.3, Jun 11, p.6. 2003.

Druker, B. J. e T. M. Roberts. Generation of a large library of point mutations in polyoma middle T antigen. Nucleic Acids Res, v.19, n.24, Dec 25, p.6855-61. 1991.

Dubridge, R. B., P. Tang, *et al.* Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol Cell Biol, v.7, n.1, Jan, p.379-87. 1987.

Dudoit, S. e J. Fridlyand. A prediction-based resampling method for estimating the number of clusters in a dataset. Genome Biol, v.3, n.7, Jun 25, p.RESEARCH0036. 2002.

- Dunant, N. M., M. Senften, *et al.* Polyomavirus middle-T antigen associates with the kinase domain of Src-related tyrosine kinases. J Virol, v.70, n.3, Mar, p.1323-30. 1996.
- Eddy, B. E., S. E. Stewart, *et al.* Cytopathogenicity in tissue culture by a tumor virus from mice. Proc Soc Exp Biol Med, v.98, n.4, Aug-Sep, p.848-51. 1958.
- El-Deiry, W. S., J. W. Harper, *et al.* WAF1/CIP1 is induced in p53-mediated G1 arrest and apoptosis. Cancer Res, v.54, n.5, Mar 1, p.1169-74. 1994.
- Elis, W., E. Lessmann, *et al.* Mutations in the inter-SH2 domain of the regulatory subunit of phosphoinositide 3-kinase: effects on catalytic subunit binding and holoenzyme function. Biol Chem, v.387, n.12, Dec, p.1567-73. 2006.
- Ellerman, C. e O. Bang. O Experimentelle Leukemie bei Huhnern. Centralbl. Bakteriol., v.46, p.595-609. 1908.
- Epstein, M. A., B. G. Achong, *et al.* Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma. Lancet, v.15, Mar 28, p.702-3. 1964.
- Farmerie, W. G. e W. R. Folk. Regulation of polyomavirus transcription by large tumor antigen. Proc Natl Acad Sci U S A, v.81, n.22, Nov, p.6919-23. 1984.
- Feinberg, A. P. e B. Vogelstein. "A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity". Addendum. Anal Biochem, v.137, n.1, Feb, p.266-7. 1984.
- Fields, B. N., D. M. Knipe, *et al.* Fields' virology. Philadelphia: Lippincott Williams & Wilkins. 2001. 2 v. (xix, 3087, 72 p.) p.
- Flint, S. J., L. W. Enquist, *et al.* Principles of virology : molecular biology, pathogenesis, and control of animal viruses. Washington, D.C.: ASM Press. 2000. xxiv, 804 p. p.
- Folias, A., M. Matkovic, *et al.* BRCA1 interacts directly with the Fanconi anemia protein FANCA. Hum Mol Genet, v.11, n.21, Oct 1, p.2591-7. 2002.
- Fraenkel-Conrat, H., B. Singer, *et al.* Infectivity of viral nucleic acid. Biochim Biophys Acta, v.25, n.1, Jul, p.87-96. 1957.
- Franco, D., H. B. Pathak, *et al.* Stimulation of poliovirus synthesis in a HeLa cell-free in vitro translation-RNA replication system by viral protein 3CDpro. J Virol, v.79, n.10, May, p.6358-67. 2005.
- Franke, T. F., C. P. Hornik, *et al.* PI3K/Akt and apoptosis: size matters. Oncogene, v.22, n.56, Dec 8, p.8983-98. 2003.
- Freed, E. O. HIV-1 replication. Somat Cell Mol Genet, v.26, n.1-6, Nov, p.13-33. 2001.

Freund, R., R. T. Bronson, *et al.* Separation of immortalization from tumor induction with polyoma large T mutants that fail to bind the retinoblastoma gene product. Oncogene, v.7, n.10, Oct, p.1979-87. 1992.

Frohlich, C., R. Albrechtsen, *et al.* Molecular profiling of ADAM12 in human bladder cancer. Clin Cancer Res, v.12, n.24, Dec 15, p.7359-68. 2006.

Frolova, E., I. Frolov, *et al.* Packaging signals in alphaviruses. J Virol, v.71, n.1, Jan, p.248-58. 1997.

Furumoto, K., S. Aii, *et al.* RECK gene expression in hepatocellular carcinoma: correlation with invasion-related clinicopathological factors and its clinical significance. Reverse-inducing--cysteine-rich protein with Kazal motifs. Hepatology, v.33, n.1, Jan, p.189-95. 2001.

Gao, N., Z. Zhang, *et al.* Role of PI3K/AKT/mTOR signaling in the cell cycle progression of human prostate cancer. Biochem Biophys Res Commun, v.310, n.4, Oct 31, p.1124-32. 2003.

Gaubatz, S., A. Imhof, *et al.* Transcriptional activation by Myc is under negative control by the transcription factor AP-2. Embo J, v.14, n.7, Apr 3, p.1508-19. 1995.

Georgopoulou, U., P. Tsitoura, *et al.* The protein phosphatase 2A represents a novel cellular target for hepatitis C virus NS5A protein. Biochimie, v.88, n.6, Jun, p.651-62. 2006.

Gera, J. F., I. K. Mellinghoff, *et al.* AKT activity determines sensitivity to mammalian target of rapamycin (mTOR) inhibitors by regulating cyclin D1 and c-myc expression. J Biol Chem, v.279, n.4, Jan 23, p.2737-46. 2004.

Gordon-Shaag, A., O. Ben-Nun-Shaul, *et al.* Cellular transcription factor Sp1 recruits simian virus 40 capsid proteins to the viral packaging signal, *ses.* J Virol, v.76, n.12, Jun, p.5915-24. 2002.

Gottlieb, K. A. e L. P. Villarreal. Natural biology of polyomavirus middle T antigen. Microbiol Mol Biol Rev, v.65, n.2, Jun, p.288-318 ; second and third pages, table of contents. 2001.

Gross, L. A filterable agent, recovered from Ak leukemic extracts, causing salivary gland carcinomas in C3H mice. Proc Soc Exp Biol Med, v.83, n.2, Jun, p.414-21. 1953.

Oncogenic viruses. Oxford ; New York: Pergamon Press. 1983. 2 v. (xxxii, 1203 p.) p.

Ha, S. A., S. M. Shin, *et al.* Cancer-associated expression of minichromosome maintenance 3 gene in several human cancers and its involvement in tumorigenesis. Clin Cancer Res, v.10, n.24, Dec 15, p.8386-95. 2004.

- Hanks, S. K., A. M. Quinn, *et al.* The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science, v.241, n.4861, Jul 1, p.42-52. 1988.
- Hannon, G. J. RNA interference. Nature, v.418, n.6894, Jul 11, p.244-51. 2002.
- Harper, J. W., G. R. Adami, *et al.* The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell, v.75, n.4, Nov 19, p.805-16. 1993.
- Harrich, D. e B. Hooker. Mechanistic aspects of HIV-1 reverse transcription initiation. Rev Med Virol, v.12, n.1, Jan-Feb, p.31-45. 2002.
- Harris, C. C. p53 tumor suppressor gene: from the basic research laboratory to the clinic--an abridged historical perspective. Carcinogenesis, v.17, n.6, Jun, p.1187-98. 1996.
- Hawkins, P. T., A. Eguinoa, *et al.* PDGF stimulates an increase in GTP-Rac via activation of phosphoinositide 3-kinase. Curr Biol, v.5, n.4, Apr 1, p.393-403. 1995.
- Hayward, D. G. e A. M. Fry. Nek2 kinase in chromosome instability and cancer. Cancer Lett, v.237, n.2, Jun 18, p.155-66. 2006.
- Hibbert, C. S., J. Mirro, *et al.* mRNA molecules containing murine leukemia virus packaging signals are encapsidated as dimers. J Virol, v.78, n.20, Oct, p.10927-38. 2004.
- Hibi, M., A. Lin, *et al.* Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain. Genes Dev, v.7, n.11, Nov, p.2135-48. 1993.
- Hinds, P. W., S. Mitnacht, *et al.* Regulation of retinoblastoma protein functions by ectopic expression of human cyclins. Cell, v.70, n.6, Sep 18, p.993-1006. 1992.
- Hong, Y. K., A. Mikami, *et al.* A new class of mutations reveals a novel function for the original phosphatidylinositol 3-kinase binding site. Proc Natl Acad Sci U S A, v.100, n.16, Aug 5, p.9434-9. 2003.
- Hu, Q., A. Klippel, *et al.* Ras-dependent induction of cellular responses by constitutively active phosphatidylinositol-3 kinase. Science, v.268, n.5207, Apr 7, p.100-2. 1995.
- Hubank, M. e D. G. Schatz. Identifying differences in mRNA expression by representational difference analysis of cDNA. Nucleic Acids Res, v.22, n.25, p.5640-8. 1994.
- Hui, C. M., P. Y. Cheung, *et al.* Id-1 promotes proliferation of p53-deficient esophageal cancer cells. Int J Cancer, v.119, n.3, Aug 1, p.508-14. 2006.

Hunter, T. Cooperation between oncogenes. Cell, v.64, n.2, Jan 25, p.249-70. 1991.

Ichaso, N. e S. M. Dilworth. Cell transformation by the middle T-antigen of polyoma virus. Oncogene, v.20, n.54, Nov 26, p.7908-16. 2001.

Ivanovsky, D. Ueber die Mosaikkrankheit der Tabakspflanze. St Petersburg. Acad. Imp. Sci. Bull, v.35, 1892, p.67-70 (English translation published in 1942: Concerning the mosaic disease of tobacco. Phytopathol. Classics). 1892.

Jackson, P. K., S. Chevalier, *et al.* Early events in DNA replication require cyclin E and are blocked by p21CIP1. J Cell Biol, v.130, n.4, Aug, p.755-69. 1995.

Jang, K. S., H. X. Han, *et al.* Id-1 overexpression in invasive ductal carcinoma cells is significantly associated with intratumoral microvessel density in ER-negative/node-positive breast cancer. Cancer Lett, v.244, n.2, Dec 8, p.203-10. 2006.

Janssens, V., J. Goris, *et al.* PP2A: the expected tumor suppressor. Curr Opin Genet Dev, v.15, n.1, Feb, p.34-41. 2005.

Jayakar, H. R., E. Jeetendra, *et al.* Rhabdovirus assembly and budding. Virus Res, v.106, n.2, Dec, p.117-32. 2004.

Jeanblanc, M., M. Mousli, *et al.* The retinoblastoma gene and its product are targeted by ICBP90: a key mechanism in the G1/S transition during the cell cycle. Oncogene, v.24, n.49, Nov 10, p.7337-45. 2005.

Jeffers, M., R. Paciucci, *et al.* Characterization of unr; a gene closely linked to N-ras. Nucleic Acids Res, v.18, n.16, Aug 25, p.4891-9. 1990.

Kaiser, A. D. Mutations in a temperate bacteriophage affecting its ability to lysogenize Escherichia coli. Virology, v.3, n.1, Feb, p.42-61. 1957.

Katan, M. Families of phosphoinositide-specific phospholipase C: structure and function. Biochim Biophys Acta, v.1436, n.1-2, Dec 8, p.5-17. 1998.

Kebebew, E., P. A. Treseler, *et al.* The helix-loop-helix protein, Id-1, is overexpressed and regulates growth in papillary thyroid cancer. Surgery, v.134, n.2, Aug, p.235-41. 2003.

Kern, F. G., S. Pellegrini, *et al.* Regulation of polyomavirus late promoter activity by viral early proteins. J Virol, v.60, n.1, Oct, p.275-85. 1986.

Kim, D. M., K. Yang, *et al.* Biochemical characterizations reveal different properties between CDK4/cyclin D1 and CDK2/cyclin A. Exp Mol Med, v.35, n.5, Oct 31, p.421-30. 2003.

- Kim, Y. H., M. A. Buchholz, *et al.* Up-regulation of c-myc induces the gene expression of the murine homologues of p34cdc2 and cyclin-dependent kinase-2 in T lymphocytes. J Immunol, v.152, n.9, May 1, p.4328-35. 1994.
- Kornbluth, S., M. Sudol, *et al.* Association of the polyomavirus middle-T antigen with c-yes protein. Nature, v.325, n.7000, Jan 8-14, p.171-3. 1987.
- Kress, M., E. May, *et al.* Simian virus 40-transformed cells express new species of proteins precipitable by anti-simian virus 40 tumor serum. J Virol, v.31, n.2, Aug, p.472-83. 1979.
- Kung, H. J., C. Boerkoel, *et al.* Retroviral mutagenesis of cellular oncogenes: a review with insights into the mechanisms of insertional activation. Curr Top Microbiol Immunol, v.171, p.1-25. 1991.
- Kveiborg, M., C. Frohlich, *et al.* A role for ADAM12 in breast tumor progression and stromal cell apoptosis. Cancer Res, v.65, n.11, Jun 1, p.4754-61. 2005.
- Lane, D. P. e L. V. Crawford. T antigen is bound to a host protein in SV40-transformed cells. Nature, v.278, n.5701, Mar 15, p.261-3. 1979.
- Langenfeld, E. M. e J. Langenfeld. Bone morphogenetic protein-2 stimulates angiogenesis in developing tumors. Mol Cancer Res, v.2, n.3, Mar, p.141-9. 2004.
- Le Pabic, H., A. L'helgoualc'h, *et al.* Involvement of the serine/threonine p70S6 kinase in TGF-beta1-induced ADAM12 expression in cultured human hepatic stellate cells. J Hepatol, v.43, n.6, Dec, p.1038-44. 2005.
- Lee, Y. L., L. Helman, *et al.* dlk, pG2 and Pref-1 mRNAs encode similar proteins belonging to the EGF-like superfamily. Identification of polymorphic variants of this RNA. Biochim Biophys Acta, v.1261, n.2, Apr 4, p.223-32. 1995.
- Lekanne Deprez, R. H., A. C. Fijnvandraat, *et al.* Sensitivity and accuracy of quantitative real-time polymerase chain reaction using SYBR green I depends on cDNA synthesis conditions. Anal Biochem, v.307, n.1, Aug 1, p.63-9. 2002.
- Li, J., K. Takaishi, *et al.* Insig-1 "brakes" lipogenesis in adipocytes and inhibits differentiation of preadipocytes. Proc Natl Acad Sci U S A, v.100, n.16, Aug 5, p.9476-81. 2003.
- Li, L., S. J. Forman, *et al.* Expression of DLK1 in hematopoietic cells results in inhibition of differentiation and proliferation. Oncogene, v.24, n.27, Jun 23, p.4472-6. 2005.
- Li, L., J. E. Price, *et al.* Correlation of growth capacity of human tumor cells in hard agarose with their in vivo proliferative capacity at specific metastatic sites. J Natl Cancer Inst, v.81, n.18, Sep 20, p.1406-12. 1989.

- Li, M., M. K. Lyon, *et al.* In vitro phosphorylation of the polyomavirus major capsid protein VP1 on serine 66 by casein kinase II. J Biol Chem, v.270, n.43, Oct 27, p.26006-11. 1995.
- Li, N., A. Batzer, *et al.* Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature, v.363, n.6424, May 6, p.85-8. 1993.
- Liang, J. e J. M. Slingerland. Multiple roles of the PI3K/PKB (Akt) pathway in cell cycle progression. Cell Cycle, v.2, n.4, Jul-Aug, p.339-45. 2003.
- Liang, T. J., G. G. Carmichael, *et al.* A polyoma mutant that encodes small T antigen but not middle T antigen demonstrates uncoupling of cell surface and cytoskeletal changes associated with cell transformation. Mol Cell Biol, v.4, n.12, Dec, p.2774-83. 1984.
- Lindlof, A. e Z. Lubovac. Simulations of simple artificial genetic networks reveal features in the use of Relevance Networks. In Silico Biol, v.5, n.3, p.239-49. 2005.
- Ling, M. T., T. C. Lau, *et al.* Overexpression of Id-1 in prostate cancer cells promotes angiogenesis through the activation of vascular endothelial growth factor (VEGF). Carcinogenesis, v.26, n.10, Oct, p.1668-76. 2005.
- Liu, Z. e G. G. Carmichael. Polyoma virus early-late switch: regulation of late RNA accumulation by DNA replication. Proc Natl Acad Sci U S A, v.90, n.18, Sep 15, p.8494-8. 1993.
- Livak, K. J. e T. D. Schmittgen. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods, v.25, n.4, Dec, p.402-8. 2001.
- Lodish, H. F. Molecular cell biology. New York: W.H. Freeman. 2004. xxxiii, 973 p. p.
- Lomax, M. e M. Fried. Polyoma virus disrupts ARF signaling to p53. Oncogene, v.20, n.36, Aug 16, p.4951-60. 2001.
- Lonnstedt, I. e T. Speed. Replicated microarray data. Statistica Sinica, v.12, n.1, JAN, p.31-46. 2002.
- Macleod, K. Tumor suppressor genes. Curr Opin Genet Dev, v.10, n.1, Feb, p.81-93. 2000.
- Marderosian, M., A. Sharma, *et al.* Tristetraprolin regulates Cyclin D1 and c-Myc mRNA stability in response to rapamycin in an Akt-dependent manner via p38 MAPK signaling. Oncogene, v.25, n.47, Oct 12, p.6277-90. 2006.
- Marsh, M. e A. Helenius. Virus entry: open sesame. Cell, v.124, n.4, Feb 24, p.729-40. 2006.

- Masui, T., R. Doi, *et al.* RECK expression in pancreatic cancer: its correlation with lower invasiveness and better prognosis. Clin Cancer Res, v.9, n.5, May, p.1779-84. 2003.
- Matthews, J. T. e T. L. Benjamin. 12-O-tetradecanoylphorbol-13-acetate stimulates phosphorylation of the 58,000-Mr form of polyomavirus middle T antigen in vivo: implications for a possible role of protein kinase C in middle T function. J Virol, v.58, n.2, May, p.239-46. 1986.
- Meester-Smoor, M. A., M. Vermeij, *et al.* Targeted disruption of the Mn1 oncogene results in severe defects in development of membranous bones of the cranial skeleton. Mol Cell Biol, v.25, n.10, May, p.4229-36. 2005.
- Meijer, L. Cyclin-dependent kinases inhibitors as potential anticancer, antineurodegenerative, antiviral and antiparasitic agents. Drug Resist Updat, v.3, n.2, Apr, p.83-88. 2000.
- Minden, A., A. Lin, *et al.* Differential activation of ERK and JNK mitogen-activated protein kinases by Raf-1 and MEKK. Science, v.266, n.5191, Dec 9, p.1719-23. 1994.
- Mitchell, P. J., C. Wang, *et al.* Positive and negative regulation of transcription in vitro: enhancer-binding protein AP-2 is inhibited by SV40 T antigen. Cell, v.50, n.6, Sep 11, p.847-61. 1987.
- Miyazaki, T., D. A. Bressan, *et al.* In vivo assembly and disassembly of Rad51 and Rad52 complexes during double-strand break repair. Embo J, v.23, n.4, Feb 25, p.939-49. 2004.
- Morgan, W. C., D. R. Kaplan, *et al.* Recombinant retroviruses that transduce middle T antigen cDNAs derived from polyomavirus mutants: separation of focus formation and soft-agar growth in transformation assays and correlations with kinase activities in vitro. J Virol, v.62, n.9, Sep, p.3407-14. 1988.
- Morrison, D. K. e R. E. Cutler. The complexity of Raf-1 regulation. Curr Opin Cell Biol, v.9, n.2, Apr, p.174-9. 1997.
- Moule, M. G., C. H. Collins, *et al.* Role for PP2A in ARF signaling to p53. Proc Natl Acad Sci U S A, v.101, n.39, Sep 28, p.14063-6. 2004.
- Mousli, M., R. Hopfner, *et al.* ICBP90 belongs to a new family of proteins with an expression that is deregulated in cancer cells. Br J Cancer, v.89, n.1, Jul 7, p.120-7. 2003.
- Murray, A. W. Recycling the cell cycle: cyclins revisited. Cell, v.116, n.2, Jan 23, p.221-34. 2004.
- Nemethova, M., M. Smutny, *et al.* Transactivation of E2F-regulated genes by polyomavirus large T antigen: evidence for a two-step mechanism. Mol Cell Biol, v.24, n.24, Dec, p.10986-94. 2004.

- Neto, E. H., A. L. Coelho, *et al.* Activation of human T lymphocytes via integrin signaling induced by RGD-disintegrins. Biochim Biophys Acta, v.1773, n.2, Feb, p.176-84. 2007.
- Nevanlinna, H. e J. Bartek. The CHEK2 gene and inherited breast cancer susceptibility. Oncogene, v.25, n.43, Sep 25, p.5912-9. 2006.
- Niwa, H., K. Yamamura, *et al.* Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene, v.108, n.2, Dec 15, p.193-9. 1991.
- Noda, M., H. Kitayama, *et al.* Detection of genes with a potential for suppressing the transformed phenotype associated with activated ras genes. Proc Natl Acad Sci U S A, v.86, n.1, Jan, p.162-6. 1989.
- Noll, M., K. P. Battaile, *et al.* Fanconi anemia group A and C double-mutant mice: functional evidence for a multi-protein Fanconi anemia complex. Exp Hematol, v.30, n.7, Jul, p.679-88. 2002.
- Novak, U. e B. E. Griffin. Requirement for the C-terminal region of middle T-antigen in cellular transformation by polyoma virus. Nucleic Acids Res, v.9, n.9, May 11, p.2055-73. 1981.
- Nunbhakdi-Craig, V., L. Craig, *et al.* Simian virus 40 small tumor antigen induces deregulation of the actin cytoskeleton and tight junctions in kidney epithelial cells. J Virol, v.77, n.5, Mar, p.2807-18. 2003.
- Ogris, E., I. Mudrak, *et al.* Catalytically inactive protein phosphatase 2A can bind to polyomavirus middle tumor antigen and support complex formation with pp60(c-src). J Virol, v.73, n.9, Sep, p.7390-8. 1999.
- Ogris, E., H. Rotheneder, *et al.* A binding site for transcription factor E2F is a target for trans activation of murine thymidine kinase by polyomavirus large T antigen and plays an important role in growth regulation of the gene. J Virol, v.67, n.4, Apr, p.1765-71. 1993.
- Oliveira, M. L., S. M. Brochado, *et al.* Mechanisms of cell transformation induced by polyomavirus. Braz J Med Biol Res, v.32, n.7, Jul, p.861-5. 1999.
- Oliveira, M. L., T. M. Roberts, *et al.* Mapping of polyomavirus middle T domain that is responsible for AP-1 activation. Oncogene, v.16, n.23, Jun 11, p.2975-82. 1998.
- Osaki, M., M. Oshimura, *et al.* PI3K-Akt pathway: its functions and alterations in human cancer. Apoptosis, v.9, n.6, Nov, p.667-76. 2004.
- Pallas, D. C., V. Cherington, *et al.* Cellular proteins that associate with the middle and small T antigens of polyomavirus. J Virol, v.62, n.11, Nov, p.3934-40. 1988.

Pallas, D. C., L. K. Shahrik, *et al.* Polyoma small and middle T antigens and SV40 small t antigen form stable complexes with protein phosphatase 2A. Cell, v.60, n.1, Jan 12, p.167-76. 1990.

Parkin, D. M. The global health burden of infection-associated cancers in the year 2002. Int J Cancer, v.118, n.12, Jun 15, p.3030-44. 2006.

Peduto, L., V. E. Reuter, *et al.* ADAM12 is highly expressed in carcinoma-associated stroma and is required for mouse prostate tumor progression. Oncogene, v.25, n.39, Aug 31, p.5462-6. 2006.

Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res, v.29, n.9, May 1, p.e45. 2001.

Pfisterer, P., J. Ehlermann, *et al.* A subtractive gene expression screen suggests a role of transcription factor AP-2 alpha in control of proliferation and differentiation. J Biol Chem, v.277, n.8, Feb 22, p.6637-44. 2002.

Poulin, D. L. e J. A. Decaprio. Is there a role for SV40 in human cancer? J Clin Oncol, v.24, n.26, Sep 10, p.4356-65. 2006.

Quackenbush, J. Computational analysis of microarray data. Nat Rev Genet, v.2, n.6, p.418-27. 2001.

_____. Microarray data normalization and transformation. Nat Genet, v.32 Suppl, Dec, p.496-501. 2002.

Raabe, T. e U. R. Rapp. Ras signaling: PP2A puts Ksr and Raf in the right place. Curr Biol, v.13, n.16, Aug 19, p.R635-7. 2003.

Raivich, G. e A. Behrens. Role of the AP-1 transcription factor c-Jun in developing, adult and injured brain. Prog Neurobiol, v.78, n.6, Apr, p.347-63. 2006.

Ramakrishnan, R., D. Dorris, *et al.* An assessment of Motorola CodeLink microarray performance for gene expression profiling applications. Nucleic Acids Res, v.30, n.7, Apr 1, p.e30. 2002.

Rameh, L. E. e M. C. Armelin. T antigens' role in polyomavirus transformation: c-myc but not c-fos or c-jun expression is a target for middle T. Oncogene, v.6, n.6, Jun, p.1049-56. 1991.

Rangarajan, A., S. J. Hong, *et al.* Species- and cell type-specific requirements for cellular transformation. Cancer Cell, v.6, n.2, Aug, p.171-83. 2004.

Rassoulzadegan, M., A. Cowie, *et al.* The roles of individual polyoma virus early proteins in oncogenic transformation. Nature, v.300, n.5894, Dec 23, p.713-8. 1982.

Raza, A., S. Alvi, *et al.* Excessive proliferation matched by excessive apoptosis in myelodysplastic syndromes: the cause-effect relationship. Leuk Lymphoma, v.27, n.1-2, Sep, p.111-8. 1997.

Reed, W., J. Carroll, *et al.* The etiology of yellow fever. An additional note. JAMA, v.36, p.431-440. 1901.

Rochford, R., B. A. Campbell, *et al.* Genetic analysis of the enhancer requirements for polyomavirus DNA replication in mice. J Virol, v.64, n.2, Feb, p.476-85. 1990.

Rochford, R., J. P. Moreno, *et al.* Enhancer dependence of polyomavirus persistence in mouse kidneys. J Virol, v.66, n.6, Jun, p.3287-97. 1992.

Rodriguez-Viciana, P., C. Collins, *et al.* Polyoma and SV40 proteins differentially regulate PP2A to activate distinct cellular signaling pathways involved in growth control. Proc Natl Acad Sci U S A, v.103, n.51, Dec 19, p.19290-5. 2006.

Rodriguez-Viciana, P., P. H. Warne, *et al.* Phosphatidylinositol-3-OH kinase as a direct target of Ras. Nature, v.370, n.6490, Aug 18, p.527-32. 1994.

Roner, M. R. e W. K. Joklik. Reovirus reverse genetics: Incorporation of the CAT gene into the reovirus genome. Proc Natl Acad Sci U S A, v.98, n.14, Jul 3, p.8036-41. 2001.

Roskoski, R., Jr. Src protein-tyrosine kinase structure and regulation. Biochem Biophys Res Commun, v.324, n.4, Nov 26, p.1155-64. 2004.

Rous, P. Landmark article (JAMA 1911;56:198). Transmission of a malignant new growth by means of a cell-free filtrate. By Peyton Rous. Jama, v.250, n.11, Sep 16, p.1445-9. 1983.

Ruiz-Hidalgo, M. J., E. Gubina, *et al.* dlk modulates mitogen-activated protein kinase signaling to allow or prevent differentiation. Exp Cell Res, v.274, n.2, Apr 1, p.178-88. 2002.

Sambrook, J. e D. W. Russell. Molecular cloning : a laboratory manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. 2001. 3 v. (various pagings) p.

Samuels, Y. e K. Ericson. Oncogenic PI3K and its role in cancer. Curr Opin Oncol, v.18, n.1, Jan, p.77-82. 2006.

Schlegel, R. e T. L. Benjamin. Cellular alterations dependent upon the polyoma virus Hr-t function: separation of mitogenic from transforming capacities. Cell, v.14, n.3, Jul, p.587-99. 1978.

Schlondorff, J. e C. P. Blobel. Metalloprotease-disintegrins: modular proteins capable of promoting cell-cell interactions and triggering signals by protein-ectodomain shedding. J Cell Sci, v.112 (Pt 21), Nov, p.3603-17. 1999.

Schuchner, S., M. Nemethova, *et al.* Transactivation of murine cyclin A by polyomavirus large and small T antigens. J Virol, v.75, n.14, Jul, p.6498-507. 2001.

Schummer, M., W. Ng, *et al.* Inexpensive handheld device for the construction of high-density nucleic acid arrays. Biotechniques, v.23, n.6, Dec, p.1087-92. 1997.

Serunian, L. A., K. R. Auger, *et al.* Production of novel polyphosphoinositides in vivo is linked to cell transformation by polyomavirus middle T antigen. J Virol, v.64, n.10, Oct, p.4718-25. 1990.

Shah, K. V. SV40 and human cancer: a review of recent data. Int J Cancer, v.120, n.2, Jan 15, p.215-23. 2007.

Shaulian, E. e M. Karin. AP-1 in cell proliferation and survival. Oncogene, v.20, n.19, Apr 30, p.2390-400. 2001.

Shechter, D. e J. Gautier. MCM proteins and checkpoint kinases get together at the fork. Proc Natl Acad Sci U S A, v.101, n.30, Jul 27, p.10845-6. 2004.

Shin, S. Y., S. C. Yoon, *et al.* Phosphorylation of glycogen synthase kinase-3beta at serine-9 by phospholipase Cgamma1 through protein kinase C in rat 3Y1 fibroblasts. Exp Mol Med, v.34, n.6, Dec 31, p.444-50. 2002.

Silva, J., J. M. Silva, *et al.* Analysis of the candidate tumor suppressor Ris-1 in primary human breast carcinomas. Mutat Res, v.594, n.1-2, Feb 22, p.78-85. 2006.

Smith, A. E. e A. Helenius. How viruses enter animal cells. Science, v.304, n.5668, Apr 9, p.237-42. 2004.

Sogayar, M. C., A. A. Camargo, *et al.* A transcript finishing initiative for closing gaps in the human transcriptome. Genome Res, v.14, n.7, Jul, p.1413-23. 2004.

Span, P. N., C. G. Sweep, *et al.* Matrix metalloproteinase inhibitor reversion-inducing cysteine-rich protein with Kazal motifs: a prognostic marker for good clinical outcome in human breast carcinoma. Cancer, v.97, n.11, Jun 1, p.2710-5. 2003.

Stanley, W. M. Isolation of a crystalline protein possessing the properties of tobacco mosaic virus. Science, v.81, n.2113, 06/1935, p.644-645. 1935.

Stehelin, D., H. E. Varmus, *et al.* DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature, v.260, n.5547, Mar 11, p.170-3. 1976.

Su, W., W. Liu, *et al.* Association of Polyomavirus middle tumor antigen with phospholipase C-gamma 1. J Biol Chem, v.270, n.21, May 26, p.12331-4. 1995.

Talbot, S. J. e D. H. Crawford. Viruses and tumours--an update. Eur J Cancer, v.40, n.13, Sep, p.1998-2005. 2004.

- Tessema, M., U. Lehmann, *et al.* Cell cycle and no end. Virchows Arch, v.444, n.4, Apr, p.313-23. 2004.
- Tiscornia, G., O. Singer, *et al.* A general method for gene knockdown in mice by using lentiviral vectors expressing small interfering RNA. Proc Natl Acad Sci U S A, v.100, n.4, Feb 18, p.1844-8. 2003.
- Toker, A., M. Meyer, *et al.* Activation of protein kinase C family members by the novel polyphosphoinositides PtdIns-3,4-P2 and PtdIns-3,4,5-P3. J Biol Chem, v.269, n.51, Dec 23, p.32358-67. 1994.
- Treisman, R., A. Cowie, *et al.* The structures of the spliced mRNAs encoding polyoma virus early region proteins. J Mol Appl Genet, v.1, n.2, p.83-92. 1981.
- Treisman, R., U. Novak, *et al.* Transformation of rat cells by an altered polyoma virus genome expressing only the middle-T protein. Nature, v.292, n.5824, p.595-600. 1981.
- Triqueneaux, G., M. Velten, *et al.* RNA binding specificity of Unr, a protein with five cold shock domains. Nucleic Acids Res, v.27, n.8, Apr 15, p.1926-34. 1999.
- Uren, A. G., J. Kool, *et al.* Retroviral insertional mutagenesis: past, present and future. Oncogene, v.24, n.52, Nov 21, p.7656-72. 2005.
- Van Der Jagt, M. F., F. C. Sweep, *et al.* Correlation of reversion-inducing cysteine-rich protein with kazal motifs (RECK) and extracellular matrix metalloproteinase inducer (EMMPRIN), with MMP-2, MMP-9, and survival in colorectal cancer. Cancer Lett, v.237, n.2, Jun 18, p.289-97. 2006.
- Van Hoof, C. e J. Goris. PP2A fulfills its promises as tumor suppressor: which subunits are important? Cancer Cell, v.5, n.2, Feb, p.105-6. 2004.
- Vermeulen, K., D. R. Van Bockstaele, *et al.* The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Prolif, v.36, n.3, Jun, p.131-49. 2003.
- Vivanco, I. e C. L. Sawyers. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nat Rev Cancer, v.2, n.7, Jul, p.489-501. 2002.
- Walter, G., R. Ruediger, *et al.* Association of protein phosphatase 2A with polyoma virus medium tumor antigen. Proc Natl Acad Sci U S A, v.87, n.7, Apr, p.2521-5. 1990.
- Wang, X., R. H. Wang, *et al.* Genetic interactions between Brca1 and Gadd45a in centrosome duplication, genetic stability, and neural tube closure. J Biol Chem, v.279, n.28, Jul 9, p.29606-14. 2004.
- Wang, Z. e D. D. Brown. A gene expression screen. Proc Natl Acad Sci U S A, v.88, n.24, p.11505-9. 1991.

Watson, J. D. e F. H. Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature, v.171, n.4356, Apr 25, p.737-8. 1953.

Weinberg, R. A. Oncogenes and tumor suppressor genes. CA Cancer J Clin, v.44, n.3, May-Jun, p.160-70. 1994.

Welford, S. M., J. Gregg, *et al.* Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. Nucleic Acids Res, v.26, n.12, p.3059-65. 1998.

Wewer, U. M., L. E. Thornell, *et al.* Extrasynaptic location of laminin beta 2 chain in developing and adult human skeletal muscle. Am J Pathol, v.151, n.2, Aug, p.621-31. 1997.

White, D. O. e F. Fenner. Medical virology. San Diego: Academic Press. 1994. xvii, 603 p. p.

White, M. K., J. Gordon, *et al.* Human polyomaviruses and brain tumors. Brain Res Brain Res Rev, v.50, n.1, Dec 1, p.69-85. 2005.

Whitman, M., C. P. Downes, *et al.* Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate. Nature, v.332, n.6165, Apr 14, p.644-6. 1988.

Winnischofer, S. M., M. L. De Oliveira, *et al.* Suppression of AP-1 constitutive activity interferes with polyomavirus MT antigen transformation ability. J Cell Biochem, v.90, n.2, Oct 1, p.253-66. 2003.

Wu, W., N. Dave, *et al.* Comparison of normalization methods for CodeLink Bioarray data. BMC Bioinformatics, v.6, p.309. 2005.

Yamamoto, T., B. De Crombrughe, *et al.* Identification of a functional promoter in the long terminal repeat of Rous sarcoma virus. Cell, v.22, n.3, Dec, p.787-97. 1980.

Yang, P. e S. K. Roy. Follicle stimulating hormone-induced DNA synthesis in the granulosa cells of hamster preantral follicles involves activation of cyclin-dependent kinase-4 rather than cyclin d2 synthesis. Biol Reprod, v.70, n.2, Feb, p.509-17. 2004.

Yang, S. I., R. L. Lickteig, *et al.* Control of protein phosphatase 2A by simian virus 40 small-t antigen. Mol Cell Biol, v.11, n.4, Apr, p.1988-95. 1991.

Yang, Y. G., Z. Herceg, *et al.* The Fanconi anemia group A protein modulates homologous repair of DNA double-strand breaks in mammalian cells. Carcinogenesis, v.26, n.10, Oct, p.1731-40. 2005.

Yin, D., D. Xie, *et al.* DLK1: increased expression in gliomas and associated with oncogenic activities. Oncogene, v.25, n.13, Mar 23, p.1852-61. 2006.

Yu, B., Y. Wang, *et al.* Protein kinase A regulates cell cycle progression of mouse fertilized eggs by means of MPF. Dev Dyn, v.232, n.1, Jan, p.98-105. 2005.

Yu, Q., M. A. Ciemerych, *et al.* Ras and Myc can drive oncogenic cell proliferation through individual D-cyclins. Oncogene, v.24, n.47, Oct 27, p.7114-9. 2005.

Yuan, B., Y. Xu, *et al.* Increased expression of mitotic checkpoint genes in breast cancer cells with chromosomal instability. Clin Cancer Res, v.12, n.2, Jan 15, p.405-10. 2006.

Zeng, Y. X., K. Somasundaram, *et al.* AP2 inhibits cancer cell growth and activates p21WAF1/CIP1 expression. Nat Genet, v.15, n.1, Jan, p.78-82. 1997.

Zhang, X. F., J. Settleman, *et al.* Normal and oncogenic p21ras proteins bind to the amino-terminal regulatory domain of c-Raf-1. Nature, v.364, n.6435, Jul 22, p.308-13. 1993.

Zhu, W., A. Eicher, *et al.* At the onset of transformation polyomavirus middle-T recruits shc and src to a perinuclear compartment coincident with condensation of endosomes. Oncogene, v.17, n.5, p.565-76. 1998.

8. APÊNDICES

Apêndice A - Primers

Os primers para PCR foram desenhados com o auxílio do programa *Primer 3* (<http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi>) (Rozen e Skaletsky, 2000), e *primers* para os experimentos de *Real-Time PCR* foram desenhados com o auxílio do programa *Primer Express v2.0* (Applied Biosystems). As seqüências estão representadas na direção 5'→3'.

Construção das bibliotecas subtraídas (RDA)

Síntese da primeira fita do cDNA

Smart IIA - Sau3A1 - AAGCAGTGGTATCAACGCAGAGATCGCGGG
CDS IIA - Sau3A1 - AAGCAGTGGTATCAACGCAGAGATCT (29) VN

Síntese da segunda fita e amplificação do cDNA

PCR SmartIIA-Sau3A1 - AAGCAGTGGTATCAACGCAGAGATCGCGGG

Controle da qualidade do cDNA

Smart IIA - Sau3A1 - AAGCAGTGGTATCAACGCAGAGATCGCGGG
CDS IIA - Sau3A1 - AAGCAGTGGTATCAACGCAGAGATCT (29) VN
GAPDHF - CTGCACCACCAACTGCTTA
GAPDHR - CATGACGGCAGGTCAGGTC
MLH1F - TGGTGTCTCTAGTTCTGG
MLH1R - CATTGTTGTAGTAGCTCTGC
NotchF - ACTGTGGCCAACCAGTTCTC
NotchR - CTCTCACAGGTGCTCCCTTC

Geração das Representações e da população Driver

R12 - GATCTGCGGTGA
R24 - AGCACTCTCCAGCCTCTCACCGCA

Geração da população Tester I

J12 - GATCTGTTTCATG
J24 - ACCGACGTCGACTATCCATGAACA

Geração da população Tester II

N12 - GATCTTCCCTCG
N24 - AGGCAACTGTGCTATCCGAGGGAA

PCR de colônia

PBlueF	- GTTTTCCCAGTCACGACGTT
PBlueR	- GGCTCGTATGTTGTGTGGAA

Clonagem de Dlk1 em pLPCX

mDLK1-EcoRI-R	- CCGGAATTCCGGGTGGTCCGCAACCAGAAGC
mDLK1-EcoRI-R	- CCGGAATTCCGGCCTCATCACCAGCCTCCTTGT

Amplificação de MN1

BB632797-F-320	- AGCTCTGCTTAGCTGAAGATGGA
BB632797-F-54	- TACCCGCTGAGCCATCTTAC
BB632797-R-281	- GGAAGTTGAATCAAGCCACAGAG
BB658850-F-198	- CCGCGTCTACTGCTGTGAATC
BB658850-R-468	- CTAGGAGCTCTCTCCGCTGCT
BB658850-R-583	- GCACAGGCGGTGTGAACTC
CA315561-F-87	- GCCATGAGTACTATCGACCTGGA
XM_485667-F-1148	- CACAGGCGACACTTAGCGACTT
XM_485667-F-2439	- TTATGAGCACATGGCGGAGA
XM_485667-F-3297	- TTACATGAACGTGGCCAAGAG
XM_485667-F-4413	- CTCCTACTGCTTGCCAGAACA
XM_485667-F-6990	- AGCAGAGATTGCTGAGCTGATGT
XM_485667-F-9	- CATCTTCACCGAGCCTTCTGT
XM_485667-R-1169	- AAGTCGCTAAGTGTGCGCCTGT
XM_485667-R-2670	- ATCTACGGCTCCTTGGTTCGT
XM_485667-R-3807	- CTGCTGCTCAAAGCCACTCA
XM_485667-R-4470	- AGTCACGTTGAGGTTAGGAGCA
XM_485667-R-5845	- GGTGGAGATATCAGAGTGCAGAG
XM_485667-R-6229	- CTTCTAGAACACACGCCGAGTACA
XM_485667-R-8049	- CAGTATGCGTGAATACGGAAGACT

Real-Time PCR – Controles internos

36B4-F	- TTCTCTTAAAACTCCGGTCTGGAT
36B4-R	- AGAAACAAAGATGCCAACACAAAG
GAPDH-F	- CATGGCCTTCCGTGTTTCCTA
GAPDH-R	- CCTGCTTCACCACCTTCTTGA

Actina-F	- ACGGCCAGGTCATCACTATTG
Actina-R	- AGTTTCATGGATGCCACAGGAT
Tubulina-F	- AGTTTTTCGCGGACCACTTCA
Tubulina-R	- TAGAGCTCCCAGCAGGCATT

Real-Time PCR - CodeLink bioarrays

Adam11-F	- GTGGAGGTGTGAACGAGTATGG
Adam11-R	- GCTCCGGTGCTTATTCCACAT
Adam12-F	- ATCAGTGTCTTCGGCGTTCAC
Adam12-R	- GGCAATTCTTCCTGTTGTTACATACC
Adamts1-F	- CCTTCGGAATTTCTGCAACTG
Adamts1-R	- AATCCTGTCTGGTGAACAGAATTG
Adamts5-F	- TCTTCAATCCTTACCAGCATCGAT
Adamts5-R	- AGCAAACAATTACCATGACCATCA
Brcal-F	- CCAAACGCTGACTCCCTTAGTG
Brcal-R	- TCGGTGGTAGCTCCAGAATTCT
Cdc2a-F	- TATGATCCTGCCAAACGAATCTC
Cdc3a-R	- TCCAGAGGGCTACATCTTCTTAATC
Dkk1-F	- GCTGCCCCGGGAACACTACT
Dkk1-R	- CAAGGTTTTCAATGATGCTTTCCT
Dkk2-F	- CCACACTCCAAGATGCCTCATAT
Dkk2-R	- CCAGAAGTGGCGAGCACAA
Dkk3-F	- TGCCCTGGACATCACCTTAAA
Dkk3-R	- TCCCTCAGGTGGCTTCTTGT
Dlk1-F	- GCTGGGACGGGAAATTCTG
Dlk1-R	- GGAGCATTCGTACTGGCCTTT
Fanca-F	- TGCTGGCTTTTTTCACTTGATACC
Fanca-R	- CACGTATGTCCGGAGTAGATAGGA
Ferritina-F	- CGTGGATCTGTGTACTTGCTTCA
Ferritina-R	- GCTGCGACTGGAGAGACTTGT
Fgfr1-F	- GCTGGTTGAAAAATGGCAAAG
Fgfr1-R	- GGCACCACAGAATCCATTATGA
GAPDH-F	- CATGGCCTTCCGTGTTTCTA
GAPDH-R	- CCTGCTTCACCACCTTCTTGA
Igf1-F	- ACCCGGACCTACCAAAATGAC
Igf1-R	- CTGGTGAAGGTGAGCAAGCA
Igf2-F	- GCCCTTCTCCTCCGATCCT
Igf2-R	- CAAAGAGATGAGAAGCACCAACA
Igf2r-F	- CCAGGAGCCCCCAAATTC
Igf2r-R	- GAGTACAAGCCAATGCAGTTTCA
Igfbp7-R	- CCATCACCCAGGTCAGCAA
Igfbp7-R	- ACCTTGGCACCAGTGACGTT
Insig1-F	- GCGGCTGTTGTGCGTTTACT
Insig1-R	- CACGGCAATACAGCGCATAA
Ak013687-F	- CCAAGCCACACCTGATCACA
Ak013687-R	- AGGCAGCAACCTGGGTATGT
AK017200-F	- TAGGTGGCTGATTCATTCACTCA
AK017200-R	- AGGCATTTTGCAAGACCAGGTA

AK018939-F	- CCAATGGCTGGAGGAAAGG
AK018939-F	- CCAATGGCTGGAGGAAAGG
AK018939-R	- TCCCCACCGGTCTTTTCTC
AK019875-F	- TGAGGCTTGGATTAAGGATCCA
AK019875-R	- GTCAACCCCCTCAGCATCTG
AK029831-F	- ACGTGAAGTTACATGTCCGTAGGA
AK029831-R	- AGTTGGCCTCATCCCTTCACT
AK029831-R	- AGTTGGCCTCATCCCTTCACT
AK080810-F	- GAACTGGAGAGGTACAGCCTGAA
AK080810-R	- TTGGCCAGATCTTTTTTACATCAGA
Basp1-F	- ACCCAAAGCCGAACCTCAA
Basp1-R	- CGCCTTCGGCCTTCTTGT
Bmpr2-F	- TTTCCACCCCCTGACACAA
Bmpr2-R	- GGCGACTATCAAAACAGCTAACAC
Cdc20-F	- AGCGCCTTGATGTTGAGTGA
Cdc20-F	- GGCAAGTGGTGGCAATGATA
Cdc25c-F	- AGGATTTTCAAGAAGACCCAATGG
Cdc25c-R	- AATGGCGTTCATGTCACAGAGT
Cdc6-F	- TCCGTAAAGCGCTGGATGTT
Cdc6-R	- CGACTCGCTGGGTGATTTACA
Cdca5-F	- TGAAGAAGATCGTAGCTCATGCA
Cdca5-F	- TGAAGAAGATCGTAGCTCATGCA
Cdca5-R	- TACCTTGAGCGGAGGGTTGTT
Cdk1-F	- AGAACACCTTTCCCAAGTGGAA
Cdk1-R	- TTGGCAGGATCATAGACTAGCATT
Chek2-F	- CTGGCTCCTGAGGTTCTTGTCT
Chek2-R	- TGGATAACCCACTTAGGCAGATGA
CiclinaA2-F	- CCACTGACACCTCTTGACTATCCA
CiclinaA2-R	- TCATTAACGTTCACTGGCTTGTCT
CiclinaB2-F	- CGACCTTTGCCACTTCACTTC
CiclinaB2-R	- CGTCAGCTCCATCAGGTACTTG
CiclinaG1-F	- TCACGGTTAGGGAAACTCCTGAT
CiclinaG2-R	- GTCTTTAGCCTTGATTTGAATGCTT
Cks2-F	- CAAAACCTCATCTGATGTCCGAAGA
Cks2-R	- AGAATATGCGGTTCTGGCTCAT
Frat1-F	- CAACACGGAGACAACCGAATG
Frat1-R	- CCCACCTCCTCTGGGATGT
Fzd2-F	- ACATCGCCTACAACCAGACCAT
Fzd2-R	- TCACCAGCGGGTAGAACTGAT
Gmn-F	- TGTCTTCCTCCACGGATGCTA
Gmn-R	- AGCACTTGACCCGTGTAGTTAGTTT
Gspt2-F	- CCCTAGAAGAATCTTGGGAGCTT
Gspt2-R	- CCTCCATCACTTCCTTGACACTT
Gtf2i-F	- ATTGCTGGCGAAGGAAAGG
Gtf2i-R	- CACTCCTGAGGCTGGTTTATCAA
Id1-F	- CGCACTGAGGGACCAGATG
Id1-R	- GCGGTAGTGTCTTTCCCAGAGA
Ier51-F	- GACGTTCCGCTCTCCTGAGT
Ier51-R	- ACCGAAACGTCCCTTTTGG

Klf2-F	- TACACCAAGAGCTCGCACCTAA
Klf2-R	- GTCAGAGCGCGGAACTT
MAD211-F	- ATGTAAACACAGCAGAAGGGATCA
MAD211-R	- GTCCTCCCTGCCTCTCTCACT
Map3k5-F	- TGAGCTCATGCAGCCTAACTTC
Map3k5-R	- CTCGCTTGCGCCACCTT
Mcm3-F	- CGGGAAGCGCAGAGAGACTA
Mcm3-R	- CCGATACTGGTTGTCACTGATCA
Mmp14-F	- GACAGGCAAGGCTGATTTGG
Mmp14-R	- CGAACATTGGCCTTGATCTCA
Mmp3-F	- TGCACCTGGACCAGGGATTA
Mmp3-R	- TGAGCAGCAACCAGGAATAGG
Mmp9-R	- CCAGACGTGGGTCGATTCC
Mmp9-R	- GTCTCGCGGCAAGTCTTCAG
NM_007955-F	- GGAGAGCCCTGCTGATAACATG
NM_007955-R	- ACACCATTTCAGAGTTCCCTTT
NM_020616-F	- AGTCCACGCAAGCCACAGTAG
NM_020616-R	- TGCGGCCACGATAGTAACAA
NM_025766-F	- CACGGTGCTCACAGACTTTGA
NM_025766-R	- TGAACATGCGGATGCTGAAT
NM_026152-F	- AGTCCACGCAAGCCACAGTAG
NM_026152-R	- TGCGGCCACGATAGTAACAA
Oip5-F	- AGGTTTTCTCAAGTCCAGCACATA
Oip5-R	- CCTCTCAAGGCAGCCATAGC
Pnpt1-F	- ACATCCCACTGCCCTAGGACTA
Pnpt1-R	- CGAGAAAGCCTCATTCTTCCA
Pou5f1-F	- CTCACCCTGGGCGTTCTCT
Pou5f1-R	- GCCGCAGCTTACACATGTTCT
Pttg1-F	- TGGCCAGCCTTTGGTGAT
Pttg1-R	- TGGAGTGCTCATCCAGTCTCAGT
Rad51-F	- TGACACCGAGGGCACCTT
Rad51-R	- GAACCCTCGCGCATATGC
Rgs2-F	- GGGCAGAAGCATTGATGAAC
Rgs2-R	- GCCAACCAGAATTCAATGTTTTTC
Rhox5-F	- GAGGAAGCACAAAGAATCATCCA
Rhox5-R	- TGTTCTTCCGAGTCTTCCTTGAC
Sap30-F	- AGGCAACGCCAGCTTCAG
Sap30-R	- CACAAATGTAAAGATGCCTTGCA
Tmem158-F	- CGCAAGCCCATCGAGTCTAC
Tmem158-R	- ACGCTCCACACCACGATGA
Tmem162-F	- AGACCAACCAGCCGTACTTTGT
Tmem162-R	- TAGATGGAAGCCAAGGGACACT
Uhrf1-F	- CACTCTCAGAGCAGCAGGCTAA
Uhrf1-R	- GATCTGGTACGGCCCATCCT
W09715-F	- CAAGCGCACCTTAAAGCAGAA
W09715-R	- TCTGCCAATCGCTGAATAAAAA
W70351-F	- CCTTGGGTTTGGCCTCCTA
W70351-R	- AGTGGTGTTTGGTGGCCTTGTC
Wdr70-F	- CATTTGACTTGGAGGCCATGT

Wdr70-R	- TGGCCCAGCCTTCCAAATA
XM_001004922-F	- AGGAAGACTGGGTGCTGCAA
XM_001004922-R	- TGCTGATGCAAGTTTGGTAATACA
Nek2-F	- GGTTGCATGTCAGAGGTGTTGT
Nek2-R	- GGAATGGGAATCCTGGTGAGT
Npdc1-F	- AGGAAAATGAAGATGGCGACTTC
Npdc1-R	- TGACAGCGTGGAATGGTCAA
p23-R	- CAAAGGAAAGGGCAAAGCTTAA
p23-R	- TCAGAGAAACGGTCAAAATTCTGA
Ptn-f	- CCTGTGGCAAGCTCACCAA
Ptn-R	- GCGTCTTTTAATCCAGCATCTTC
Ptp4a3-F	- AGGTGAGCTACCGGCACATG
Ptp4a3-R	- TGGTAGCCCCGTACTTCTTCA
Rad51-F	- GGTGAAGGGAAGGCCATGTAC
Rad51-R	- AGGACATCGCTGCCAGAGA
Rybp-F	- TGTCCACAAAGGCAGTGCTATT
Rybp-R	- CAGCTTGCATCCCAGACACA
Sept2-F	- GATGAGCAGTTTGAACGCTACCT
Sept2-R	- AATGCAACATCTAAGGGCTTCAG
Sgce-F	- ATACCTCCCATGCACACAGACA
Sgce-R	- CTGCGGTTGAGGGATCTGA
Smarcal1-F	- GGCAAGTGCTGATGTGGTCAT
Smarcal1-R	- TACTGGCTTCTTCTGACCAATGC
Sod1-F	- GCGGTGAACCAGTTGTGTTG
Sod1-R	- AGCCTTGTGTATTGTCCCCATACT
Sod2-F	- TTAACGCGCAGATCATGCA
Sod2-R	- CCAGAGCCTCGTGGTACTTCTC
Sod3-F	- AGAGCCTGACAGGTGCAGAGA
Sod3-R	- TCCCCTGGATTTGACATGGT
Stat4-F	- TGAAGTACCTCTACCCTGACATTCC
Stat4-R	- CCCGTTTCGGTTGGTCTTG
UNR-F	- GATGCCAGTGCTCCACGTCTA
UNR-R	- ATGACACCAGCTTGACGGATCT
Wnt4-F	- GCGTAGCCTTCTCACAGTCCTT
Wnt5-R	- CTGCCAGCCTCGTTGTTGT
Zyx-F	- GGACGACATGACCAAGAACGA
Zyx-R	- GTTTGGTACTAGGCTTGGGAACA
GAPDH-F	- CATGGCCTTCCGTGTTTCTTA
GAPDH-R	- CCTGCTTCACCACCTTCTTGA
Actina-F	- ACGGCCAGGTCATCACTATTG
Actina-R	- AGTTTCATGGATGCCACAGGAT
Tubulina-F	- AGTTTTTCGCGGACCACTTCA
Tubulina-R	- TAGAGCTCCCAGCAGGCATT