IDENTIFICAÇÃO DE GENES DIFERENCIALMENTE METILADOS EM LINHAGENS TUMORAIS DE MAMA ATRAVÉS DE TRATAMENTO COM AGENTE DESMETILANTE E CONSTRUÇÃO DE BIBLIOTECAS SUBTRATIVAS DE cDNA

MURILO VIEIRA GERALDO

Dissertação apresentada à Fundação Antônio Prudente para obtenção do título de Mestre em Ciências

Área de concentração: Oncologia

Orientadora: Dra. Anamaria Aranha Camargo

São Paulo
2006

FUNDAÇÃO ANTONIO
PRUDENTE

FICHA CATALOGRÁFICA

Preparada pela Biblioteca do Centro de Tratamento e Pesquisa

Hospital do Câncer A.C. Camargo

Geraldo, Murilo Vieira

Identificação de genes diferencialmente metilados em linhagens tumorais de mama através de tratamento com agente desmetilante e construção de bibliotecas subtrativas de cDNA / Murilo Vieira Geraldo -- São Paulo, 2006.

123p.

Dissertação(mestrado)-Fundação Antônio Prudente.

Curso de Pós-Graduação em Ciências-Área de concentração: Oncologia.

Orientadora: Anamaria Aranha Camargo

Descritores: 1. CÂNCER DA MAMA/genética. 2. MARCADORES BIOLÓGICOS DE TUMOR. 3. METILAÇÃO DE DNA. 4. BIBLIOTECA DE GENES. 5. AZACITIDINA. 6. LINHAGEM CELULAR TUMORAL.

"A vida é aquilo que
acontece enquanto fazemos planos"
(John Lennon)

DEDICATÓRIA

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenador de Ensino

Aos meus alicerces:

meu pai, Ivaldir, exemplo de justiça e perseverança.

minha mãe, Elvira, exemplo de ternura e dedicação.

minha irmã, Nola, pelo carinho incondicional e amizade.

minha Rita, pelo carinho e amor sempre tão intensos.

FUNDAÇÃO ANTONIO
PRUDENTE

AGRADECIMENTOS

À Dra. Anamaria A. Camargo, pela oportunidade de fazer parte de um excelente grupo de pesquisa e pelo exemplo profissional. Muito obrigado pela leitura criteriosa desta dissertação e pela disponibilidade e paciência infinitas.

Ao MSc Ricardo P. Moura, pelo convívio diário durante a realização do trabalho, pelas críticas sempre construtivas e dicas sempre valiosas na realização dos experimentos de PCR em Tempo Real. A disponibilidade e as muitas longas conversas contribuíram muito para a realização deste trabalho.

Ao Dr. Newton V. Verbisck, pelas críticas sempre construtivas, pela conversa, pela leitura cuidadosa deste trabalho e principalmente pela amizade e descontração, sempre importantes.

À Maria Cristina Costa pela padronização das reações de seqüenciamento, etapa muito importante deste trabalho.

À Dra. Ana Paula M. Silva, pela paciência nas instruções sobre o geNorm.

À MSc Fabiana Bettoni, pela paciência, receptividade e atenção dedicados a mim durante minha chegada ao laboratório.

Ao Dr. Raphael B. Parmigiani, pelas dicas e pelo tratamento das linhagens celulares.

Aos colegas Fabrício F. Costa, Fernando Camargo, Maria Cristina F. Costa, Daniela F. Ierardi, Érico T. Costa, Daniel O. Vidal e Ricardo P. Moura

pela convivência, pelas ótimas conversas e pelas atividades "extracurriculares" desenvolvidas ao longo do trabalho.

Aos outros colegas de laboratório Felícia P. Cavalher, Valéria Paixão, Fabiana Bettoni, Raphael Parmigiani, Lílian Pires, Lílian Inoue, Ana Paula Silva, Anna Christina Salim e Elisa N. Ferreira pelo ambiente de trabalho e pelos momentos de descontração.

À Dra. Luísa L Villa (Laboratório de Virologia – I.L.P.C.) e sua equipe, em especial ao José Carlos M. de Prado pela disponibilidade e ajuda na fixação do cDNA nas membranas de nylon.

À Dra. Dirce M. Carraro (Laboratório de Análise da Expressão Gênica – I.L.P.C.), pela disponibilidade e receptividade em seu laboratório e, em especial, aos seus alunos Maria Cristina Rangel, pela atenção, paciência e ajuda valiosa na realização das hibridizações do Northern Reverso e Paulo Pineda na amplificação do mRNA. A todo o laboratório, Nádia, Mariana, Tiago, Elisa, Ellen e Louise, obrigado pelos momentos de descontração.

À Dra. Vilma R. Martins (Laboratório de Biologia Celular – I.L.P.C.) e ao Dr. Edgar Graner (Faculdade de Odontologia da UNICAMP - Campus de Piracicaba) pelo acompanhamento do meu trabalho desde o início, pelas críticas e sugestões.

Ao Dr. Wilson Araújo Silva Jr. (Faculdade de Medicina da USP – Campus de Ribeirão Preto) pela receptividade e, em especial ao Dr. Paulo Peitl Jr. pelos ensinamentos na realização do RaSH, pela amizade e pela ambientação à Ribeirão Preto, à Anemari D. R. Santos, pela extrema dedicação e ajuda no RaSH e ao Daniel G. Ribeiro, pela atenção e esforços dispensados com a *webpage*. A todo o laboratório pelo carinho especial nesta etapa do trabalho.

À Dra. Mari Cleide Sogayar (Instituto de Química da USP), pela receptividade em seu laboratório e, em especial à Rita de Cássia S. Figueira pelo fornecimento do cDNA das amostras tumorais de mama, pelas dicas valiosas e pela leitura cuidadosa desta dissertação e ao Dr. Christian Collin pelas valiosas dicas para a elaboração, realização e interpretação dos resultados dos experimentos de Northern Reverso.

À Dra. Silvia R. B. Uliana (Instituto de Ciências Biomédicas da USP), minha primeira orientadora e quem me apresentou à Biologia Molecular. Muito obrigado pelos puxões de orelha e pela confiança. Muito obrigado a todos os meus amigões do Laboratório de Leishmanioses: Camila, Renata, Denise, Henrique, Jenicer, Danilo e Fernando, pelo fantástico ambiente de trabalho, pelo carinho e pelos infinitos momentos de descontração.

A todos os funcionários do I.L.P.C., pelo suporte técnico e administrativo.

A todos os colegas do I.L.P.C., pela criação de um ambiente de trabalho favorável à obtenção de resultados positivos.

À Diretora do I.L.P.C., a Ilma Profa. Dra. Luísa L. Villa, que, pela eficiente administração do Instituto, possibilitou excelentes condições para a realização do presente trabalho.

À Suely Francisco e aos demais funcionários da biblioteca da Fundação Antônio Prudente, pela atenção e boa vontade na organização e revisão desta dissertação.

À Ana Maria Kuninari e Luciana Pitombeira, da secretaria de Pós-graduação, pela competência e pela atenção dispensadas aos alunos.

Aos grandes amigos: Rodrigo A. Prado e Thiago M. Marques, pelos mais de vinte anos de amizade e Julio M. Ricarte, pela amizade dentro e fora da Biologia. Cada um, do seu jeito, fez com que cada resultado ruim fosse amenizado e rapidamente superado nesta trajetória. As "intimações" do fim de tarde sempre foram bem-vindas.

Aos meus amigos Alexandre Takaki, Danilo F. Rodil, Marcelo W. Leite, Thierry Faria Lima, Ernesto S. Nakayasu e Almir P. Figueiredo, que fizeram parte da minha formação escolar e universitária.

Aos Profissionais (Mineiro, Blanes, Sinas, Paulinho, Lili, Camila, Rita, Julio, Pipoca, Mari, Mona, Guilherme, PCR, Pererê entre outros) pelos muitos momentos de descontração juntos.

À FAPESP e CAPES pelo suporte financeiro durante a realização deste trabalho.

A todos os meus familiares que, com alegria, me empurraram "ladeira acima". Sempre.

Ao meu tio Armando A. H. Vieira, que, pelo sucesso na carreira acadêmica, foi o incentivo para a escolha da minha profissão. Aos outros biólogos da família (Ana Maria e Aline) que compartilham o gosto pela Biologia.

À minha querida família, Ivaldir, Elvira e Nola pelo amor e amparo incondicionais, presentes em TODOS os momentos, pelo incentivo constante e pelas críticas construtivas. Vocês têm participação importantíssima em cada um dos meus passos até aqui.

À minha Ritoca, meu amor, pela compreensão e pelos inesquecíveis "simples" momentos juntos. Seu carinho e dedicação foram imprescindíveis para a superação de todos os obstáculos neste trabalho. Muito obrigado pelo apoio intelectual e principalmente emocional da lindona nestes anos de graduação e pós-graduação. Obrigado à sua família, Jorge, Maria e Cézar, pela receptividade e carinho.

A todos aqueles direta ou indiretamente envolvidos nesta jornada. Muito obrigado!

RESUMO

Geraldo MV. **Identificação de genes diferencialmente metilados em linhagens tumorais de mama através do tratamento com agente desmetilante e a construção de bibliotecas subtrativas de cDNA.** São Paulo; 2006. [Dissertação de Mestrado – Fundação Antônio Prudente]

Uma das áreas mais promissoras na pesquisa sobre o câncer de mama é a identificação de marcadores moleculares que possam ser utilizados na detecção precoce da doença, na predição do prognóstico do paciente ou no desenvolvimento de abordagens terapêuticas alternativas, reduzindo, assim, as taxas de mortalidade e morbidade associadas à doença. Alterações epigenéticas têm sido associadas à tumorigênese, e a metilação do DNA é um dos mecanismos epigenéticos mais estudados. A metilação do DNA está associada ao silenciamento gênico e alterações no padrão de metilação têm sido utilizadas com freqüência na caracterização de genes relacionados ao câncer, assim como marcador tumoral para o diagnóstico e prognóstico da doença. Neste trabalho, nós identificamos genes possivelmente silenciados por metilação do DNA em linhagens celulares tumorais de mama, como uma forma de aumentar o número de marcadores moleculares disponíveis para o câncer de mama. Para tanto, 3 linhagens celulares tumorais de mama (MCF-7, MDA-MB-231, MDA-MB-435) foram tratadas com o agente desmetilante 5-aza-2'-deoxicitidina, capaz de induzir a expressão de genes silenciados por metilação do DNA. RNA total extraído de culturas celulares submetidas ou não ao tratamento com o agente desmetilante foi utilizado para a construção de bibliotecas subtrativas de cDNA utilizando a metodologia de RaSH (*Rapid Subtraction Hybridization*). Duas bibliotecas diferentes foram construídas para cada linhagem celular: uma utilizando como *tester* cDNA sintetizado a partir de RNA extraído da linhagem celular tratada com o agente desmetilante, e contendo genes induzidos pelo tratamento; e a outra,

utilizando como tester cDNA da linhagem celular não submetida ao tratamento, e que foi utilizada como controle da subtração. Um total de 2367 seqüências foram geradas a partir de 3 bibliotecas subtraídas utilizando o cDNA das linhagens MCF-7, MDA-MB-435 e MDA-MB-231 tratadas como tester. A comparação entre cada biblioteca subtrativa e seu experimento controle possibilitou a exclusão de um total de 50 genes candidatos falso-positivos. A comparação das seqüências oriundas dessas bibliotecas com seqüências depositadas em bancos de dados públicos, permitiu a identificação de 533 genes induzidos pelo tratamento e possivelmente regulados por metilação. Para a identificação de genes induzidos na linhagem MDA-MB-231, foi realizado um experimento de Northern Blot Reverso, e 4 genes que apresentaram indução superior a 3 vezes foram escolhidos para validação por PCR em Tempo Real. Ao todo, 41 candidatos foram selecionados para validação experimental por PCR em Tempo Real utilizando o cDNA sintetizado a partir do mesmo RNA utilizado para a construção das bibliotecas de RaSH. Dos 37 genes para os quais foi possível padronizar a reação de PCR em Tempo Real, 5 genes mostraram aumento da expressão de pelo menos 2 vezes após o tratamento (*Sequestosome 1, SQSTM1; Superoxide dismutase 2, SOD2; Connective Tissue Growth Factor, CTGF; Small Ubiquitin-like Modifier 2, SUMO2* e *Fibronectin 1, FN1).* Uma análise preliminar do padrão de expressão destes genes em 18 amostras de tumores de mama revelou a redução da expressão do gene *SUMO2* em 10 amostras tumorais de mama quando comparadas a um pool de amostras de mama normal.

SUMMARY

Geraldo MV. **Identification of differentially methylated genes in breast tumor cell lines through treatment with demethylating agent and generation of cDNA subtracted libraries.** São Paulo; 2006. [Dissertação de Mestrado-Fundação Antônio Prudente]

One of the most promising areas in breast cancer research is the identification of molecular markers that can be used for early detection of the disease, for predicting patient prognosis and for the development of alternative therapeutic approaches, reducing, in this way, the mortality and morbidity associated with the disease. Epigenetic alterations have been frequently associated with tumorigenesis, and DNA methylation is one of the most studied epigenetic mechanism. DNA methylation is associated with gene silencing and changes in the DNA methylation pattern have been frequently used in the characterization of cancer-related genes, as well as a tumor marker for disease diagnosis and prognosis. In this work, we have identified genes putatively silenced by DNA methylation in breast tumor cell lines, as a way to enlarge the number of molecular markers available for breast cancer. For this purpose, 3 breast tumor cell lines (MCF-7, MDA-MB-435 and MDA-MB-231) were treated with the demethylating agent 5-aza-2'-deoxycytidine (5aza-dC), which can induce the expression of genes silenced by DNA methylation. Total RNA extracted from cell cultures submitted or not to the treatment with the demethylating agent was used to construct subtractive cDNA libraries using the RaSH (*Rapid Subtraction Hybridization*) methodology. Two different libraries were constructed for each cell line: one using as tester cDNA synthesized from RNA extracted from the cell line treated with the demethylating agent and containg the genes induced by the treatment; and the other using as tester cDNA from the cell line not submitted to the treatment and that was used as control for the subtraction. A total of

2367 sequences were generated from 3 cDNA libraries subtracted using cDNA from the treated MCF-7, MDA-MB-435 and MDA-MB-231 cell lines as tester. Comparison between each subtractive library and its control allowed the exclusion of a total of 50 false-positives candidate genes. Comparison of the sequences generated from these libraries with sequences submitted to public databases allowed the identification of 533 genes induced by the treatment and possibly regulated by DNA methylation. For the identification of induced genes in the MDA-MB-231 cell line, a reverse Northern blot experiment was carried out, and 4 genes with a more than three-fold induction were selected for validation by Real Time PCR. In all, 41 candidates were selected for experimental validation by Real Time PCR, using cDNA synthesized from the same RNA source used fro the construction of the RaSH libraries. From the 37 genes for which it was possible to carry on Real Time PCR analysis, 5 genes showed more than 2-fold induction upon treatment (*Sequestosome 1, SQSTM1; Superoxide dismutase 2, SOD2; Connective Tissue Growth Factor, CTGF; Small Ubiquitin-like Modifier 2, SUMO2 e Fibronectin 1, FN1*). A preliminary analysis of the expression pattern of these genes in 18 breast tumor samples revealed reduced expression of the *SUMO2* gene, in 10 breast tumor samples when compared to a pool of normal breast samples.

LISTA DE FIGURAS

Figura 1	Esquema representando a anatomia da mama feminina e as principais estruturas afetadas pelo câncer de mama.	3
Figura 2	Taxas de sobrevida em 5 anos de acordo com o estadiamento ao diagnóstico.	8
Figura 3	Metilação de DNA na Ilha de CpG presente na região promotora de um gene.	12
Figura 4	Silenciamento epigenético no câncer.	14
Figura 5	Estruturas das moléculas de citidina e seu análogo 5-azacitidina.	24
Figura 6	Esquema representando a técnica de RaSH.	33
Figura 7	Estratégia adotada para a identificação de genes silenciados por metilação e reativados pelo tratamento com 5-aza-dC.	46
Figura 8	Exemplo de padronização das reações de PCR quantitativo em tempo real.	61
Figura 9	Fórmula proposta por PFAFFL para a quantificação relativa por PCR em Tempo Real.	63
Figura 10	Construção das bibliotecas subtrativas.	71
Figura 11	*Webpage* utilizada para análise *in silico.*	76
Figura 12	Análise global das seqüências geradas pelos experimentos 5AZA-*tester.*	79
Figura 13	Verificação da presença de Ilha de CpG nos genes identificados pela técnica de RaSH.	82
Figura 14	Presença de Ilhas de CpG na região promotora nos genes re-induzidos por 5-aza-dC.	83
Figura 15	Análise em larga escala por *Northern* Reverso.	87
Figura 16	Reativação da expressão gênica após tratamento com o agente desmetilante 5aza-dC.	93
Figura 17	Análise preliminar da expressão gênica dos cinco genes validados em amostras tumorais de mama.	104

FUNDAÇÃO ANTONIO PRUDENTE BIBLIOTECA

LISTA DE TABELAS

Tabela 1	Estimativas das taxas brutas de incidência por 100.000 e de número de casos novos por câncer, em mulheres, para o ano 2006.	2
Tabela 2	Metilação e câncer de mama.	18
Tabela 3	O tratamento de linhagens celulares tumorais com 5aza-2'-deoxicitidina para a identificação de genes diferencialmente metilados.	27
Tabela 4	Distribuição dos dados clínico-patológicos das 18 amostras de tumores primários de mama	39
Tabela 5	Análise seqüencial realizada pelo *pipeline* desenvolvido pelo Dr. Wilson Araújo da Silva Jr.	53
Tabela 6	Seqüências dos *primers* desenhados para a validação experimental.	57
Tabela 7	Seqüências dos *primers* dos genes normalizadores utilizados.	64
Tabela 8	Características gerais das linhagens celulares utilizadas.	66
Tabela 9	Bibliotecas subtrativas construídas para as linhagens MCF-7, MDA-MB-435 e MDA-MB-231 e seqüências geradas.	70
Tabela 10	Genes selecionados para validação experimental.	86
Tabela 11	Reativação da expressão gênica após tratamento com o agente desmetilante 5-aza-dC.	92
Tabela 12	Aberrações estruturais encontradas regiões cromossômicas dos cinco genes re-induzidos pelo tratamento com 5-aza-dC.	100

LISTA DE ABREVIATURAS

5-aza-dC	5-Aza 2'-deoxicitidina
µg	Micrograma
µL	Microlitro
µM	Micromolar
Abs	Absorbância
ACS	*American Cancer Society*
BCRP	do inglês *Breast Cancer Resistance Protein*
BLAT	do ingles *BLAST-like Alignment Tool*
BLAST	do ingles *Basic Local Alignment Sequence Tool*
BSA	Albumina de soro bovino
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
cDNA	DNA complementar
CpG	do inglês *cytosine phosphate guanine*
Ct	do inglês *Cycle Threshold*
DEPC	Di-etil-pirocarbonato
DMEM	do inglês *Dulbecco's Modified Eagle's Medium*
DMSO	Dimetilsulfóxido
DNMT	DNA metil-transferase
DNA	Ácido desoxirribonucléico
DNase	Desoxirribonuclease
dNTP	Desoxinucleotídeo
DTT	Ditiotreitol
EDTA	Ácido etilenodiaminotetracético
EST	do inglês *Expressed Sequence Tags*
EUA	Estados Unidos da América
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
GAPDH	Gliceraldeído 3-fosfato desidrogenase

HMEC	do inglês *Human Mammary Epithelial Cells*
HDAC	Histona Deacetilase
HPRT	Hipoxantina fosforribosil transferase
HCl	Ácido Clorídrico
INCA	Instituto Nacional de Câncer
Kb	Kilobase
kDa	Kilodalton
LOH	do inglês *Loss Of Heterozigosity*
LRES	do inglês *Long-Range Epigenetic Silencing*
M	Molar
Mb	Megabase
MBP	do inglês *Methylcytosine Binding Protein*
$MgCl_2$	Cloreto de Magnésio
MOPS	Ácido 3-N-Morfolino propano-sulfônico
mRNA	RNA mensageiro
mM	Milimolar
NaCl	Cloreto de Sódio
NaOH	Hidróxido de Sódio
NCBI	do inglês *National Center for Biotechnology Information*
NH_4OAc	Acetato de amônio
ng	nanograma
nM	Nanomolar
nm	Nanômetro
pb	Pares de base
PCR	do inglês: *Polimerase Chain Reaction*
pH	Potencial hidrogeniônico
PM	Peso molecular
RaSH	do inglês *Rapid Subtraction Hybridization*
RE	Receptor de estrógeno
RNA	Ácido ribonucléico
RNase	Ribonuclease
RP	Receptor de progesterona

RPM	Rotações por minuto
RPMI	do inglês *Roswell Park Memorial Institute*
rRNA	RNA ribossomal
RT-PCR	Reverse Transcription PCR
SAGE	do inglês *Serial Analysis of Gene Expression*
SAME	Serviço de Arquivo Médico e Estatística
SDS	Dodecil sulfato de sódio
SFB	Soro fetal bovino
SG	Sobrevida Global
SLD	Sobrevida Livre de Doença
SSC	Solução Salina contendo citrato de sódio
SSH	Supression Subtractive Hybridization
TE	Tampão Tris-EDTA
TGF	Fator de crescimento transformante
TNM	do inglês *Tumor Node Metastasis*
Tris	Tris-Hidroximetil-Amino-Metano
TSA	Tricostatina A

ÍNDICE

1	**INTRODUÇÃO**	**1**
1.1	O câncer de mama	1
1.2	Alterações genéticas e epigenéticas no câncer	8
1.3	Metilação de DNA e o câncer de mama	16
1.4	Metilação de DNA como marcador molecular do câncer de mama	19
1.5	Utilização de agentes desmetilantes na identificação de genes diferencialmente metilados em tumores	21
1.6	Análise global da expressão gênica	28
1.7	Técnicas de hibridização subtrativa de cDNA	30
2	**OBJETIVOS**	**36**
2.1	Objetivo geral	36
2.2	Objetivos específicos	36
3	**MATERIAIS E MÉTODOS**	**38**
3.1	Linhagens celulares tumorais	38
3.2	Amostras teciduais de carcinoma mamário humano	38
3.3	Condições de cultura e manutenção das linhagens celulares	39
3.4	Tratamento das linhagens celulares com agente desmetilante	40
3.5	Extração de RNA total	40
3.6	Hibridização de subtração rápida - RaSH	41
3.6.1	Síntese de cDNA dupla-fita	42
3.6.2	Digestão da dupla-fita de cDNA e ligação nos adaptadores	43
3.6.3	Hibridização e Construção das Bibliotecas Subtrativas	44
3.7	Seleção de clones para seqüenciamento	47
3.8	Northern Blot Reverso	47
3.8.1	Fixação do DNA em membranas de Nylon	47

3.8.2 Reação de amplificação de RNA para a síntese de cDNA marcado com isótopo radioativo 48
3.8.3 Marcação de cDNA com isótopo radioativo 49
3.8.4 Hibridização das membranas com sonda radioativa 50
3.8.5 Detecção do sinal radioativo 50
3.9 Análise in silico das bibliotecas subtrativas 51
3.10 Seleção de genes para validação experimental 54
3.11 Validação dos genes por PCR em Tempo Real 55
3.11.1 Síntese de cDNA fita simples 55
3.11.2 Controle da qualidade do cDNA 55
3.11.3 Desenho de *primers* específicos 56
3.11.4 Reação de PCR quantitativa em Tempo Real 58
3.11.5 Cálculo da expressão gênica diferencial dos candidatos selecionados para a validação experimental 62

4 RESULTADOS e DISCUSSÃO 65
4.1 Tratamento das linhagens celulares tumorais e construção das bibliotecas subtrativas 65
4.2 Análise das seqüências geradas 72
4.3 Seleção de candidatos para validação experimental 80
4.4 Utilização de Northern Reverso para seleção em larga escala de genes candidatos para validação experimental 84
4.5 Validação experimental através de PCR em Tempo Real 87
4.6 Os genes candidatos alidados 94
4.7 Análise preliminar da expressão dos genes validados em amostras de tumores primários de mama 101

5 CONCLUSÕES **108**

6 REFERÊNCIAS BIBLIOGRÁFICAS **109**

ANEXO

Anexo 1 Curriculum Vitae

1 INTRODUÇÃO

1.1 O CÂNCER DE MAMA

Dentre os diversos tipos de câncer, o de mama tem alta incidência no Brasil, sendo uma das principais causas de morte entre as mulheres. Dos 472.050 novos casos de câncer previstos para o Brasil no ano de 2006, sendo 237.480 em mulheres, o câncer de mama será o segundo mais incidente entre a população feminina, sendo responsável por 48.930 novos casos (Tabela 1) (Ministério da Saúde 2005). A incidência deste tipo de câncer se elevou de forma constante até a década de 80, com discreto aumento desde então, devido ao uso mais freqüente da mamografia para o rastreamento da doença. Por outro lado, a mortalidade de pacientes com esse tipo de câncer apresentou um declínio considerável para mulheres com menos de 50 anos, principalmente devido às medidas de prevenção e de tratamento adotadas (STYBLO e WOOD 1998). Exames periódicos de mamografia e a implementação da terapia adjuvante, como, por exemplo, o tamoxifeno, permanecem como ferramentas fundamentais para o diagnóstico precoce e o aumento na sobrevida das pacientes, respectivamente.

Tabela 1 - Estimativas para o ano 2006 das taxas brutas de incidência por 100.000 e de número de casos novos por câncer, em mulheres, segundo localização primária.

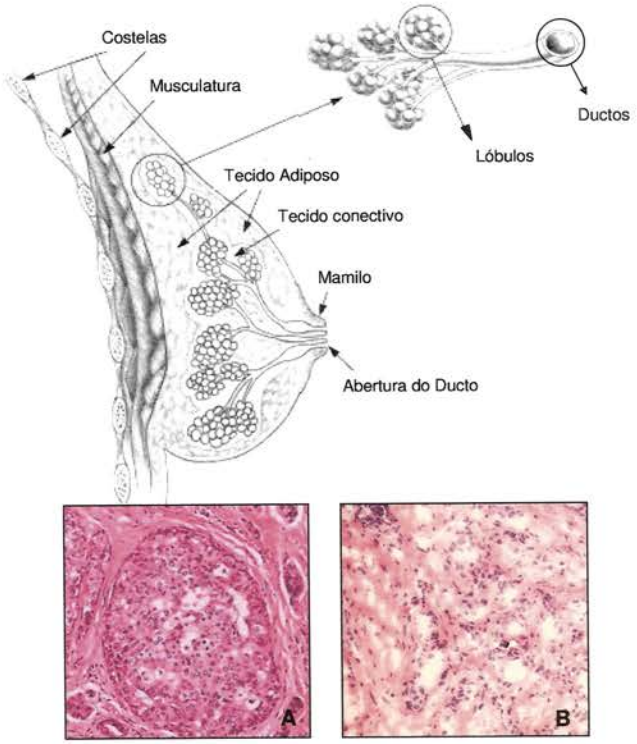
Localização Primária	*Estimativa de Casos Novos*	
Neoplasia Maligna	*Total*	*Taxa Bruta*
Mama Feminina	48.930	51,66
Traquéia, Brônquio e Pulmão	9.320	9,82
Estômago	8.230	8,65
Colo do Útero	19.260	20,31
Cólon e Reto	13.970	14,73
Esôfago	2.610	2,74
Leucemias	4.220	4,45
Cavidade Oral	3.410	3,58
Pele Melanoma	3.050	3,16
Outras Localizações	63.320	66,78
Subtotal	176.320	185,95
Pele Não Melanoma	61.160	64,53
Todas as Neoplasias	237.480	250,45

Fonte: Ministério da Saúde (2005)

O câncer de mama apresenta grande heterogeneidade clínica e histopatológica. Esta complexidade é gerada pela influência de diversos fatores, tanto exógenos (dieta e, ainda alvo de muita controvérsia, o uso de contraceptivos orais), quanto endógenos (predisposição genética e taxas hormonais) (BIÉCHE e LIDEREAU 1995). A ocorrência de menarca precoce, menopausa depois dos 50 anos de idade, nuliparidade ou primeira gravidez após os 30 anos de idade também parecem ser fatores importantes para a complexidade clinica e histopatológica desta doença. Além destes fatores, o consumo habitual de álcool parece estar associado ao risco aumentado do desenvolvimento do câncer de mama (RAHMAN e STRATTON 1998; BOYD 1999).

Os principais tipos histológicos dos tumores de mama são os carcinomas lobulares e ductais, manifestando-se na forma *in situ,* na qual o

tumor permanece confinado ao tecido mamário ou, em estádios mais avançados, rompendo a membrana basal e invadindo os tecidos adjacentes. O carcinoma ductal invasivo responde pela grande maioria (50 a 80%) dos casos de câncer de mama, seguido pelo carcinoma lobular invasivo, que responde por 5 a 15% dos casos. A Figura 1 representa esquematicamente a anatomia da mama feminina e as principais estruturas relacionadas ao câncer.



Fonte: Esquema modificado de http://www.cancervic.org.au/cancer1/prevent/breasthealth/. Fotomicrografias obtidas de seções de tumores de mama coradas com Hematoxilina-Eosina, gentilmente fornecidas pela Dra. Dirce M. Carraro, do Laboratório de Biologia Molecular do Hospital do Câncer A.C. Camargo (Aumento 40 vezes).

Legenda: (A) exemplo de carcinoma ductal *in situ* e (B) carcinoma ductal invasivo.

Figura 1 - Esquema representando a anatomia da mama feminina e as principais estruturas afetadas pelo câncer de mama.

FUNDAÇÃO ANTONIO PRUDENTE

Os métodos diagnósticos já estão muito avançados e, no geral, combinam o auto-exame, o exame clínico palpatório, a ultra-sonografia e a mamografia, chegando juntos a uma acuidade de até 95%. Apesar disso, no Brasil, 70% dos casos de câncer de mama ainda são diagnosticados tardiamente, o que dificulta muito a cura da doença (Hospital do Câncer 2006). Uma das principais razões da demora no diagnóstico é o alto custo financeiro dos exames. Além disso, outros fatores podem aumentar o número de pacientes com tumores diagnosticados tardia ou erroneamente. Como exemplo, o exame de mamografia (o principal método de rastreamento do câncer de mama na população brasileira), além de gerar certo desconforto para a paciente, pode não detectar uma fração de tumores. Isto porque sua sensibilidade é extremamente variável, de 46% a 88%, dependendo de diversos fatores, como tamanho e localização da lesão, densidade do tecido mamário, qualidade dos recursos técnicos e habilidade de interpretação do radiologista (Ministério da Saúde 2005). Ainda assim, o uso da mamografia é essencial na detecção precoce do câncer de mama. Atualmente, a ultra-sonografia mamária é uma alternativa mais sensível que a mamografia, porém requer equipamento especial de alta resolução, de preferência associado a Doppler colorido, assim como médico especializado na execução e interpretação das imagens. O que a torna mais cara sendo, portanto, utilizada como exame de confirmação dos exames de mamografia ou em casos onde a mamografia não atinge precisão satisfatória, como em casos de mulheres jovens, com tecido mamário denso. O exame clínico palpatório, quando realizado por médico ou enfermeiro

treinados, pode detectar tumores de até 1cm, se superficiais (Ministério da Saúde 2005). A sensibilidade varia entre 57% a 83% em mulheres de 50 a 59. O auto-exame das mamas não substitui o exame clínico palpatório e não é recomendado pelo INCa como estratégia única de rastreamento do câncer de mama.

O tratamento da doença envolve principalmente a ressecção cirúrgica do tumor, a quimioterapia e a radioterapia, sendo estes associados, ou não, a hormônioterapia, de acordo com características moleculares do tumor. Os tipos de cirurgia mais utilizados são a cirurgia conservadora e a mastectomia radical, com possível reconstrução da mama com tecidos autólogos. A quimioterapia contra o câncer de mama inclui uma variedade de agentes citotóxicos e hormonais usados principalmente para o controle rápido da massa tumoral, tratamento paliativo de carcinomas metastáticos e tratamento de tumores insensíveis a hormônios. Os agentes citotóxicos mais utilizados incluem drogas alquilantes, antimetabólitos, alcalóides de vinca, antraciclinas e taxanos. Individualmente, estes quimioterápicos produzem respostas efetivas em 20 a 80% dos pacientes com câncer de mama metastático, sendo que respostas completas são raras e menos de 20% dos pacientes que atingem estas respostas as mantêm em 5 anos.

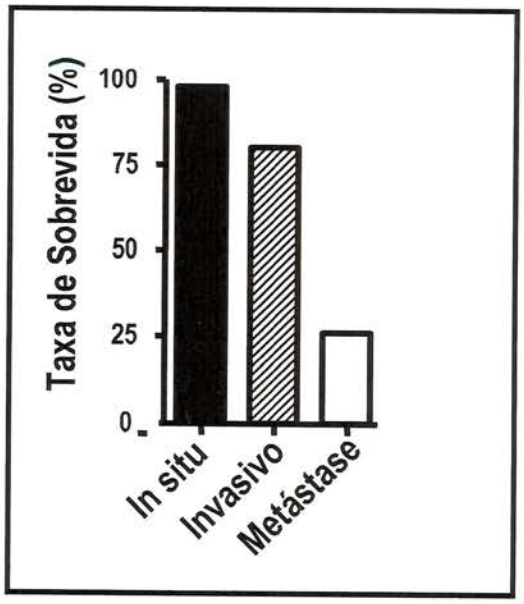
O sistema TNM (*Tumor-Node-Metastasis*) é utilizado para avaliar o tumor tanto clínica quanto anátomo-patologicamente, com o intuito de estabelecer o correto estadiamento da doença. Assim, são analisadas três características clínicas: (T) tamanho e localização do tumor primário; (N) a presença e, se houver, o número de linfonodos comprometidos por

metástases e (M) presença de metástases à distância (STYBLO e WOOD 1998). O tamanho do tumor tem sido utilizado como fator prognóstico secundário, dado que apenas 15% dos tumores com menos de 1,1cm de diâmetro apresentam comprometimento linfonodal axilar, enquanto mais de 60% dos tumores com mais de 5,5cm apresentam envolvimento de linfonodos axilares (STYBLO e WOOD 1998).

De modo semelhante à maioria dos pacientes que desenvolvem tumores sólidos, na maioria das pacientes com câncer de mama a causa de morte se deve à doença metastática que se desenvolve a partir do tumor primário. A disseminação de células tumorais a órgãos distantes pode ocorrer tanto por via sanguínea quanto por via linfática. Linfonodos axilares comprometidos são considerados o principal fator de risco para a propagação metastática sistêmica e alguns oncologistas acreditam que mulheres com linfonodos comprometidos devem sofrer terapia adjuvante. O número de linfonodos comprometidos por metástases e a presença de metástases à distância são os fatores prognósticos mais importantes no câncer de mama. Tumores de mama diagnosticados antes da disseminação metastática apresentam chances de cura elevadas. No entanto, em tumores metastáticos as taxas de sobrevida decaem significativamente, dependendo dos sítios secundários atingidos (LEE e WELCH 1997; WELCH et al. 2000; YANG et al. 2004). Nos E.U.A., entre 1995 e 2001, pacientes diagnosticados com câncer de mama *in situ*, ou seja, confinado ao tecido mamário apresentaram 98% de taxa de sobrevida em 5 anos, enquanto pacientes com câncer de mama com metástases detectadas ao diagnóstico

apresentaram queda significativa da taxa para apenas 26% (JEMAL et al. 2006) (Figura 2). Assim, torna-se clara a importância do desenvolvimento de metodologias sensíveis que permitam o diagnóstico precoce da doença.

O estadiamento pelo sistema TNM, combinado a outros aspectos anátomo-patológicos e moleculares, possibilitam a definição de um pior ou um melhor prognóstico. O grau de diferenciação celular, a presença de núcleos aberrantes, a freqüência de mitoses e a presença de mitoses atípicas, o tecido de origem do tumor primário e o grau de invasão das células neoplásicas no tecido adjacente são alguns dos critérios histológicos utilizados como fatores prognósticos. Características moleculares do tumor, como a expressão de receptores de estrógeno e progesterona, do oncogene *c-erbB2*, e das proteínas Ki-67 e PCNA podem revelar a natureza proliferativa do tumor, sendo então utilizados como marcadores moleculares de prognóstico, além de possibilitar também o tratamento diferencial para cada paciente. A expressão de receptores de estrógeno e progesterona pelo tumor indica, por exemplo, a sensibilidade do mesmo à hormônioterapia, possibilitando o tratamento com o Tamoxifeno (Hexal). Tumores com amplificação do oncogene *c-erbB2* podem ser tratados com anticorpo monoclonal específico anti-c-erbB2, o Herceptin® (Trastuzumab, Genentech), aumentando as chances de sobrevida da paciente.



Fonte: American Cancer Society-ACS (2006)

Figura 2 - Taxas de sobrevida em 5 anos de acordo com o estadiamento ao diagnóstico, durante o período de 1995 a 2001, nos E.U.A.

1.2 ALTERAÇÕES GENÉTICAS E EPIGENÉTICAS NO CÂNCER

Sabe-se que o câncer é causado pela proliferação anormal das células, formando, em um primeiro momento, um tumor localizado. Ao sofrerem modificações adicionais em suas propriedades de adesão ao substrato e de invasão, estas células tornam-se capazes de invadir tecidos adjacentes e, ao atingir a corrente sangüínea, podem colonizar tecidos distantes, originando as metástases à distância. A perda do controle do ciclo celular é uma conseqüência de alterações que ocorrem no material genético das células tumorais. Os genes afetados por essas alterações e envolvidos no processo de formação de tumores já estão bem caracterizados e, funcionalmente, podem ser divididos em duas classes: os proto-oncogenes e

os genes supressores de tumor. A aquisição de características associadas com a progressão tumoral e formação de metástases também é conseqüência de alterações que ocorrem no conteúdo genético das células que compõem o tumor localizado. No entanto, muito pouco se sabe sobre os genes que controlam estes processos.

Os genes envolvidos na formação e progressão tumoral podem ser afetados por vários tipos de alterações genéticas que incluem mutações pontuais (substituições, deleções e inserções) e cromossômicas (duplicações, translocações e deleções). Alterações epigenéticas, que alteram a expressão gênica sem modificar a seqüência de DNA propriamente dita, também têm sido associadas com o surgimento de patologias de origem tumoral, sendo a metilação um dos mecanismos mais estudados.

A metilação é uma alteração epigenética característica de genomas eucariotos e um dos principais mecanismos de regulação da expressão gênica. A metilação do DNA é constituída pela adição de radicais metil ao carbono 5 de uma citosina localizada à montante de uma guanina, formando um dinucleotídeo CpG. A adição de grupamentos metil é realizada por enzimas denominadas DNA Metiltransferases ou DNMTs (ATTWOOD et al. 2002).

A distribuição de regiões ricas em dinucleotídeos CpG no genoma dos vertebrados não é randômica. Regiões ricas em dinucleotídeos CpGs, denominadas "Ilhas de CpG", compõem aproximadamente 1-2% do genoma. A maioria das Ilhas de CpG são encontradas próximas às regiões

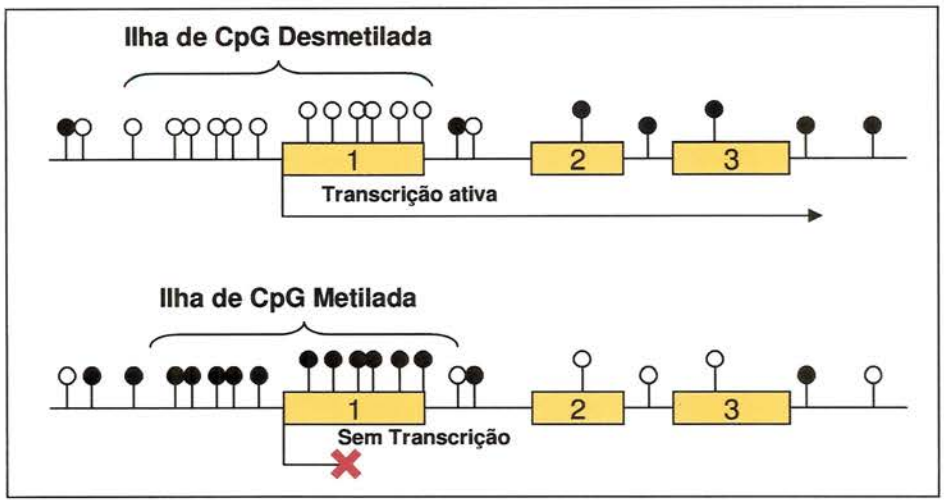
promotoras dos genes e em células normais não estão metiladas (JONES e BAYLIN 2002). O restante dos dinucleotídeos CpGs encontram-se fora das Ilhas de CpG e estão freqüentemente metilados.

De um modo geral, a presença de metilação em seqüências promotoras ou regulatórias está correlacionada com a repressão da expressão gênica. A metilação do DNA pode ser detectada, por exemplo, nos promotores de genes regulados por "*imprinting*" e no processo de inativação do cromossomo X. Além disso, a metilação também está associada com outros fatores, como envelhecimento, processos inflamatórios crônicos e infecções virais (USHIJIMA e OKOCHI-TAKADA 2005).

Diferentes mecanismos foram propostos para explicar o papel da metilação na diminuição da atividade transcricional. Em um primeiro modelo foi proposto que os resíduos de citosina metilados poderiam interferir diretamente na interação de fatores de transcrição aos seus sítios de ligação, uma vez que o grupo metil se projeta no sulco maior da dupla fita de DNA (GREGER et al. 1989). Além desta interferência direta, foi proposto que proteínas que se ligam a resíduos metilados (MBPs - *Methylcytosine Binding Proteins*) presentes nos sítios de ligação dos fatores de transcrição estariam competindo pela ligação com os próprios fatores de transcrição (CHRISTMAN 2002). Finalmente, sabe-se que as MBPs podem recrutar complexos protéicos contendo co-repressores e histona deacetilases e, com base nessa interação, foi proposto que a ligação destes complexos ao DNA levaria a uma mudança na estrutura de cromatina, impedindo a transcrição

(HERMAN 1994). Nenhum dos modelos propostos é exclusivo e, muito provavelmente todos esses modelos estão atuando simultaneamente na repressão da expressão gênica por metilação.

O processo de metilação pode silenciar genes fundamentais para a manutenção do correto funcionamento da célula. Foi observado que grande parte dos genes supressores de tumor e outros genes associados ao câncer contêm Ilhas de CpG nas suas seqüências regulatórias, que apresentam, geralmente, padrão de metilação alterado em células tumorais (BAYLIN et al. 1998; SCHMUTTE e JONES 1998). Um esquema representativo do papel da metilação na região promotora está mostrado na Figura 3. Este padrão de metilação alterado pode ser encontrado em, virtualmente, todos os tipos de neoplasias humanas (JONES e BAYLIN 2002). Há uma crescente lista de genes supressores de tumor que são silenciados por hipermetilação de seus promotores em tumores e, em muitos destes genes, esta hipermetilação é a única causa aparente para o seu silenciamento. Além destes mecanismos, a 5-metil-citosina pode sofrer deaminação hidrolítica, sendo convertida a uma timina. Estas mutações pontuais são freqüentemente observadas no gene supressor de tumor *TP53* (MAGEWU e JONES 1994) e em muitos outros genes de funções importantes como *Rb* e *c-H-ras-1* (GHAZI et al. 1990).



Fonte: Modificado de ESTELLER et al. (2002)

Legenda: As barras cor-de-laranja representam os exons do gene; círculos brancos e círculos pretos representam dinucleotídeos CpG não metilados e metilados respectivamente.

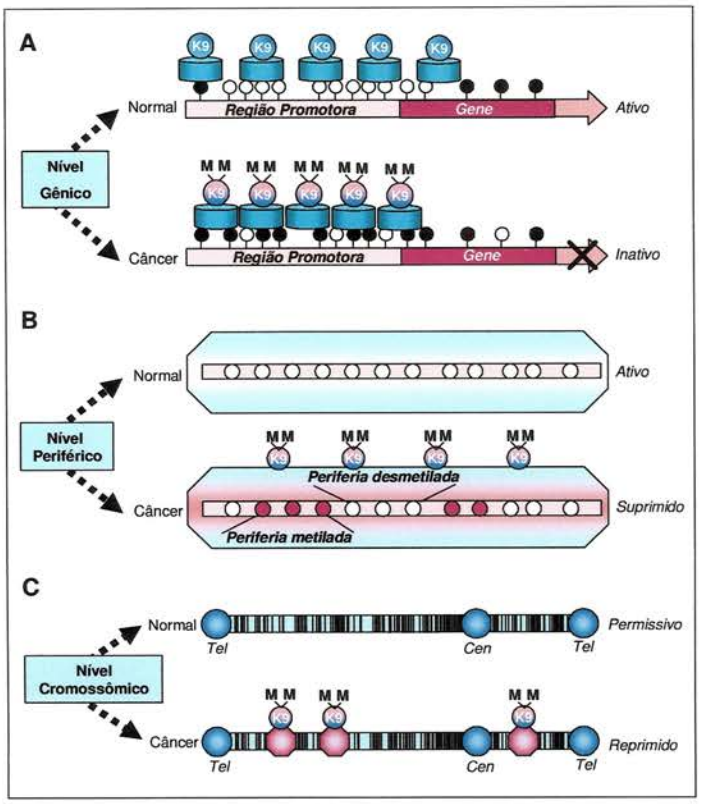
Figura 3 - Metilação de DNA na Ilha de CpG presente na região promotora de um gene.

Além da hipermetilação de regiões promotoras, uma diminuição global da metilação (hipometilação) tem sido observada em tumores. GAMA-SOSA et al. (1983), verificaram uma redução global de 10-20% no conteúdo de 5-metilcitosina no genoma de vários tipos de tumores quando comparados ao tecido normal correspondente. A perda de metilação ou a hipometilação de seqüências de DNA pode afetar diferentes tipos de seqüências repetitivas tais como os retrotransposons, os retrovírus endógenos e as repetições de DNA do tipo satélite, estando, nestes casos, associada à instabilidade cromossômica. Além disso, a perda da metilação também parece estar associada com a re-expressão de certos genes que estão silenciados em células normais. Por exemplo, o gene que codifica para a Ornitina

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenador de Ensino

descarboxilase e os proto-oncogenes *erb-A1* e *BCL-2* se mostraram hipometilados em leucemia linfocítica crônica (LIPSANEN et al. 1988). A ativação do gene *MAGE-1*, que codifica um antígeno tumoral freqüente em melanomas, também parece estar relacionada com a ausência de metilação na sua região promotora (DE SMET et al. 1996).

A hipermetilação ficaria, portanto, restrita apenas localmente, a promotores gênicos de um pequeno número de genes envolvidos de fato com a tumorigênese. No entanto, FRIGOLA et al. (2006) relataram metilação aberrante em uma extensa região do cromossomo 2, com redução ou ausência de expressão dos genes presentes na respectiva banda cromossômica. A metilação aberrante se estende por uma região que abrange cerca de 4Mb e é associada com metilação global dos resíduos de lisina 9 da histona H3. Interessantemente, alguns dos genes presentes nesta banda cromossômica não apresentaram expressão detectável, apesar de não apresentarem hipermetilação na Ilha de CpG presente em sua região promotora. Além disso, genes sem Ilha de CpG em sua região promotora situados entre duas regiões com densa metilação têm maior redução da expressão do que aqueles situados na periferia destas regiões. Assim, esse trabalho sugere que a metilação de uma determinada região do genoma pode provocar não só o silenciamento "pontual", ou seja, do gene alvo, como também de seus genes vizinhos, desencadeando a compactação da cromatina de toda a região e tornando o DNA inacessível para fatores de transcrição, inclusive para genes sem metilação em sua região promotora (Figura 4).



Fonte: Modificado de FRIGOLA et al. (2006)

Legenda: Possíveis alterações epigenéticas relacionadas com a transformação maligna. (A) Nível gênico, onde metilação de DNA e de resíduos de Lisina na histona H3 resultam em silenciamento gênico. (B) Nível periférico, onde o silenciamento epigenético em extensas regiões (LRES) resulta em silenciamento de genes com e sem metilação na região promotora. (C) Nível cromossômico, onde regiões extensas são suprimidas, convertendo eucromatina (transcricionalmente ativa) em heterocromatina (transcricionalmente inativa), resultando em aparente perda de heterosigosidade (LOH). M, metilação da lisina 9 da histona H3; círculos cheios, dinucleotídeo CpG metilado; círculos vazios, dinucleotídeos CpG desmetilados.

Figura 4 - Silenciamento epigenético no câncer.

Apesar dos numerosos estudos envolvendo metilação de DNA, o mecanismo responsável pela hipermetilação de determinados genes em detrimento de outros, que permanecem protegidos de metilação, ainda é

desconhecido. As primeiras evidências são provenientes de trabalhos que relatam o processo de indução de densa metilação a partir de regiões que apresentavam alguns dinucleotídeos CpG esparsos metilados (USHIJIMA e OKOCHI-TAKADA 2005). STIRZAKER et al. (2004), ao introduzir em linhagens celulares tumorais de próstata o gene *GSTP1* com promotor totalmente desmetilado, notaram baixa freqüência de metilação após manutenção das linhagens em cultura por 22 dias, enquanto a mesma construção, metilada esparsamente *in vitro* com metilase *Hpa*II, encontrava-se densamente metilada após o mesmo período de tempo. Este tipo de metilação em poucos dinucleotídeos CpG da região promotora funcionaria como um precursor para o aumento da freqüência de metilação nesta região. No entanto, a transcrição gênica ativa pôde impedir a metilação desta construção, mesmo quando esta se encontra esparsamente metilada (SONG et al. 2002; STIRZAKER et al. 2004). DE SMET et al. (2004) descreveram a ocorrência de re-metilação do promotor do gene *MAGE-A1* em casos onde havia o impedimento de atividade transcricional ou ausência de fatores de transcrição específicos. Uma evidência disto pode ser observada em análises globais para a identificação de genes metilados em câncer, onde muitos dos genes que apresentam metilação em Ilhas de CpG têm baixa ou nenhuma expressão detectada nas linhagens celulares normais (KANEDA et al. 2002). Assim, a diminuição da transcrição gênica também pode ser um dos mecanismos responsáveis pela indução de metilação da região promotora.

FUNDAÇÃO ANTONIO
PRUDENTE
BIBLIOTECA

1.3 METILAÇÃO DE DNA E O CÂNCER DE MAMA

A metilação de DNA já foi associada direta e indiretamente à regulação de mais de 40 genes em câncer de mama (WIDSCHWENDTER e JONES 2002). Por exemplo, a ausência de atividade do gene $p16^{ink4a}$, devido à hipermetilação de sua região promotora, foi relatada como um evento inicial na tumorigênese do câncer de mama, sendo observada em mais de 20% dos tumores de mama (PARRY e PETERS 1996). A hipermetilação do gene $p16^{ink4a}$ parece ser um evento anterior ao processo de tumorigênese, dado que, em células epiteliais mamárias humanas em cultura (HMEC), as variantes que possuem hipermetilação do mesmo em sua região promotora possuem concomitantemente maior instabilidade cromossômica e maior potencial angiogênico e inibitório de apoptose, sendo estas potenciais precursoras da progressão tumoral (LIU et al. 1998; GATELY 2000; HOWE et al. 2001; TLSTY et al. 2004).

Segundo HANAHAN e WEINBERG (2000) uma célula tumoral deve adquirir "novas habilidades" (*hallmarks*) para a transformação maligna. Estes *hallmarks* podem ser divididas em seis classes: (i) potencial ilimitado de replicação; (ii) auto-suficiência de fatores de crescimento; (iii) insensibilidade a sinais antiproliferativos; (iv) evasão da morte celular programada; (v) angiogênese sustentada e (vi) invasão de tecidos e metástase. A metilação aberrante de genes relacionados com qualquer uma destas seis classes pode ser encontrada no câncer de mama, contribuindo para a aquisição das

características tumorais (WIDSCHWENDTER e JONES 2002). Como exemplo, a hipermetilação aberrante da Ilha de CpG na região promotora do gene *CCND2*, envolvido no controle do ciclo celular, é encontrada em aproximadamente metade dos casos de câncer de mama (EVRON et al. 2001). Esta hipermetilação aberrante está relacionada com a ausência de expressão deste gene e pode, segundo os autores, possibilitar a aquisição de potencial ilimitado de replicação ao tumor de mama. A hipermetilação da região promotora do gene *CCND2* também foi encontrada em carcinomas ductais *in situ*, sugerindo a mesma como um evento inicial na tumorigênese da mama. Outros exemplos do envolvimento da metilação aberrante na aquisição de cada uma destes *hallmarks* da célula tumoral da mama estão resumidos na Tabela 2.

Tabela 2. Metilação e Câncer de Mama. Alguns exemplos do envolvimento da metilação aberrante de DNA no processo de aquisição de "novas habilidades" (*hallmarks*) no câncer de mama.

Gene	*Metilação e Câncer de Mama*	*Hallmarks*	*Referência*
CCND2	hipermetilação em aproximadamente metade dos tumores de mama, além de ser encontrada em carcinomas *in situ*	Potencial ilimitado de replicação	Evron, 2001
GSTP1	metilação aberrante da região promotora do gene *GSTP1* em cerca de 30% dos tumores primários de mama	Potencial ilimitado de replicação	Esteller, 1998
TGFβRII	silenciamento por metilação em células tumorais de mama ER^+, insensíveis a TGF-β	Insensibilidade à sinais inibitórios de proliferação	Kalkhoven et al. 1995
	Tratamento com 5-aza-dC leva à transcrição e acúmulo de TGF-β		Ammanamanchi, 1998
RASSF1A	hipermetilado em 42-65% de tumores primários de mama	Auto-suficiência em sinais de crescimento	Dammann, 2001 Burbee, 2001 Honorio, 2003
PR	hipermetilado em cerca de 40% dos tumores de mama negativos para PR além de muitas linhagens celulares tumorais também negativas para PR	Auto-suficiência em sinais de crescimento	Lapidus, 1996
E-Cad	hipermetilado em carcinomas pouco diferenciados em estágio avançado e desmetilada em tecido mamário normal	Invasão tecidual e formação de metástases	Graff, 2000
Maspina	Silenciado por metilação aberrante em seu promotor gênico em linhagens celulares tumorais de mama	Angiogênese sustentada	Domman, 2000
	Tratamento com 5-aza-dC leva à restauração da transcrição de Maspina na linhagem MCF-7		

1.4 METILAÇÃO DE DNA COMO MARCADOR MOLECULAR DO CÂNCER DE MAMA

É questionável se o rastreamento do câncer de mama somente por mamografia reduz eficazmente a mortalidade (OLSEN e GOTZSCHE 2001; MIETTINEN et al. 2002). Metodologias baseadas em análises do DNA podem ser ferramentas úteis para a detecção precoce do câncer de mama, dado que o DNA extraído de soro, plasma e outros fluidos corpóreos pode ser facilmente amplificado por PCR além de ser potencialmente mais sensível que os testes convencionais.

A descoberta de que DNA nu pode estar presente na corrente sangüínea e outros fluidos corpóreos despertou muito interesse para o desenvolvimento de diagnóstico e prognóstico molecular (MULLER et al. 2003). Muitos estudos detectaram alterações gênicas específicas de tumor em DNA recuperado a partir de plasma de pacientes com câncer. O DNA presente na corrente sangüínea e outros fluidos corpóreos pode revelar mutações oncogênicas, alterações de regiões satélites, rearranjos gênicos e cromossômicos, além de alterações epigenéticas (ANKER et al. 1999).

A detecção de metilação nas Ilhas de CpG possui muitas vantagens sobre outras alterações de DNA que podem ser detectadas em câncer (WIDSCHWENDTER e JONES 2002). O DNA metilado é facilmente detectado na presença de vasto excesso de DNA não metilado, sendo o limiar de detecção de 1 alelo metilado entre 1000 cópias não metiladas,

utilizando a tecnologia de MSP (HERMAN et al. 1996), e 10000 cópias não metiladas através da técnica de MethyLight (EADS et al. 2000).

Em especial, a utilização da metilação diferencial como marcador tumoral é vantajosa, dado que, ao contrário das diferentes alterações genéticas que ocorrem em diversas regiões do gene, as alterações no padrão de metilação são de um único tipo e ocorrem sempre na mesma região. Além disso, a metilação diferencial pode ser detectada mesmo em meio a um grande número de células normais, o que não acontece com alterações genéticas como perda de heterozigosidade (LOH), deleções homozigóticas ou mesmo mutações pontuais.

A detecção de alterações epigenéticas em fluidos corpóreos pode ser uma alternativa diagnóstica menos invasiva e mais sensível quando comparada aos métodos convencionais. Uma série de estudos relata a facilidade de identificação de metilação aberrante em DNA genômico extraído de células presentes em plasma, muco, urina, aspirado de ductos mamários entre outros fluidos. PALMISANO et al. (2000) detectaram metilação aberrante dos genes *p16* e *MGMT* em 100% das amostras de muco em pacientes de câncer de pulmão até três anos antes do seu diagnóstico clínico, mostrando que alterações moleculares podem ser detectadas previamente às alterações clínico-patológicas. Em câncer de mama, a metilação aberrante do gene $p16^{ink4a}$ pôde ser detectada em 8 (23%) dos 34 tumores analisados por SILVA et al. (1999), dos quais cinco (14%) apresentaram níveis de metilação detectáveis no plasma dos pacientes com câncer de mama. KRASSENSTEIN et al. (2004) utilizaram a

técnica de MSP para analisar o padrão de metilação da região promotora dos genes *GSTP1, RARβ2, p16^{ink4a}, p14^{ARF}, RASSF1A* e *DAP-kinase* em fluido aspirado de mamilo e amostras tumorais de pacientes com câncer de mama e detectaram hipermetilação de pelo menos um dos genes em todas (100%) as amostras de tumor e a metilação do mesmo gene em 82% das amostras de aspirado do mamilo, detectada inclusive em estágios iniciais da progressão tumoral. Além disso, a metilação aberrante dos genes *RASSF1A* e *APC* em DNA extraído de soro de pacientes com câncer de mama pôde ser considerada, em análise multivariada, um marcador prognóstico forte e independente (MULLER et al. 2003). EVRON et al. (2001) identificaram metilação aberrante nos genes *Ciclina D2, RAR-beta* e *Twist* em 17 de 20 amostras de fluido aspirado do mamilo de pacientes com carcinoma ductal invasivo, 2 de 7 amostras de carcinoma ductal *in situ* e apenas 5 de 45 amostras de doadoras saudáveis. Interessantemente, duas destas mulheres saudáveis vieram a ser diagnosticadas com câncer de mama posteriormente, demonstrando a detecção precoce da lesão no fluido aspirado de mamilo antes mesmo dos sintomas clínicos.

1.5 UTILIZAÇÃO DE AGENTES DESMETILANTES NA IDENTIFICAÇÃO DE GENES DIFERENCIALMENTE METILADOS EM TUMORES

Alterações no padrão de metilação têm sido amplamente utilizadas na identificação de novos genes envolvidos no processo de formação e

progressão tumoral. Além disso, através do emprego da tecnologia genômica para o estudo diferencial da metilação de genes em larga escala, tem sido possível correlacionar alterações globais no padrão de metilação com a progressão para a doença maligna, assim como com um pior prognóstico ou recorrência da doença após o tratamento.

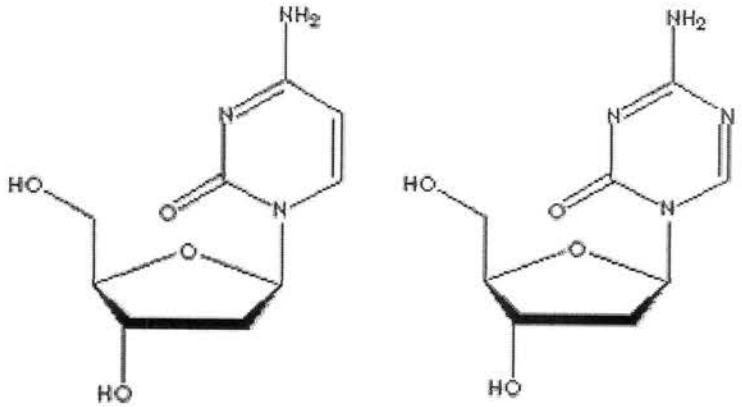
A identificação de novas regiões promotoras diferencialmente metiladas presentes no genoma tumoral é um passo importante para o entendimento do papel da metilação do DNA na tumorigênese e para o estabelecimento de um diagnóstico precoce e de um tratamento efetivo do câncer. Um fármaco sintetizado há cerca de 40 anos, descrito por SORM et al. (1964), revolucionou o estudo do processo de metilação do DNA. A 5-aza-2'-deoxicitidina é um análogo da citosina e, quando incorporada ao DNA, se liga irreversivelmente à DNA-Metiltransferase (DNMT), impedindo a ação desta enzima no processo de divisão celular (COLOT e ROSSIGNOL 1999) (Figura 5). Desta forma, o tratamento com 5-aza-2'-deoxicitidina é capaz de reduzir significativamente os níveis de metilação existentes induzindo a expressão de genes regulados por metilação. O tratamento de culturas de linhagens tumorais com esta droga pode, portanto, revelar genes silenciados pela metilação durante os processos de formação e progressão tumoral e, conseqüentemente, resultar na identificação de novos marcadores moleculares tumorais. A identificação em larga escala de genes diferencialmente metilados pode ser realizada combinando-se o tratamento de linhagens tumorais com agentes desmetilantes à utilização de

metodologias pré-estabelecidas para análise global do padrão de expressão gênica.

YAMASHITA et al. (2002) utilizaram a tecnologia de *microarrays* de cDNA para a identificação de genes supressores de tumor comumente inativados por metilação em carcinoma epidermóide de esôfago. O tratamento com 5-aza-2'-deoxicitidina foi utilizado com a finalidade de reativar genes epigeneticamente silenciados em três linhagens tumorais. Alterações no perfil de expressão gênica foram detectadas e quantificadas através da hibridização do RNA extraído das três linhagens antes e após o tratamento com o agente desmetilante em uma lâmina de vidro contendo 12.599 clones de cDNA. Ao todo mais de 565 genes mostraram-se reativados e hiper-expressos nas três linhagens após o tratamento com a droga. Um total de 58 genes conhecidos, cuja re-expressão foi detectada em pelo menos duas das três linhagens tratadas, foram selecionados para uma caracterização mais detalhada. Desses 58 genes, 44 (76%) apresentam Ilhas de CpG na região promotora sendo que destes, 25 genes foram selecionados para validação experimental por RT-PCR. A re-indução de todos estes 25 genes após o tratamento foi confirmada e o padrão de metilação da região promotora de 22 destes genes em linhagens celulares tumorais foi analisado através do tratamento com Bissulfito de Sódio seguido de seqüenciamento. O tratamento com Bissulfito de Sódio converte apenas as citosinas desprovidas de metilação em uracilas e o seqüenciamento da região permite a caracterização do seu padrão de metilação. Dos 22 genes analisados, 13 apresentaram alto grau de metilação em suas regiões

FUNDAÇÃO ANTONIO
DENTE

promotoras. A presença de metilação foi, então, avaliada em amostras tumorais e 10 dentre os 13 genes inicialmente avaliados se mostraram diferencialmente metilados em tumores de esôfago. Em estudos de análise funcional, 3 destes genes apresentaram atividade supressora de tumor.



Fonte: Modificado de CHRISTMAN (2002)

Figura 5 - Estruturas das moléculas de citidina e seu análogo 5-azacitidina.

FUKAI et al. (2003) também utilizaram o agente desmetilante 5-aza-2'-citidina e plataformas de *microarray* de cDNA para a identificação de genes silenciados por metilação em linhagens celulares de hepatoma. Dentre os 557 genes analisados, 14 mostraram indução de expressão de mais de 5 vezes após o tratamento. O trabalho resultou na identificação de hipermetilação do gene *Hepatocyte growth factor activator inhibitor 2/placental bikunin* (*HAI2/PB*) em linhagens celulares tumorais e amostras de tumor. Chiba 2005 analisaram o padrão de metilação dos outros 13 genes identificados por FUKAI et al. (2003) sendo que destes, 4 não possuíam Ilhas de CpG em sua região promotora, 4 apresentavam Ilhas de CpG totalmente desmetiladas e 5 apresentavam metilação aberrante em

muitas linhagens celulares e amostras tumorais, todas derivadas de hepatoma. Amostras de fígado normal analisadas não apresentaram hipermetilação na região promotora destes 5 últimos genes descritos.

YUAN et al. (2004), após o tratamento com 5-aza-2'-citidina de uma linhagem de câncer de pulmão de células não pequenas (NCI-H522), descreveram a expressão diferencial de 73 genes. Destes, 30 genes foram escolhidos para validação em outras linhagens celulares tratadas e não tratadas com 5-aza-2'-citidina. O gene *Hint/PCKI-1,* o qual encontrava-se silenciado, apresentou reativação de expressão após tratamento com 5-aza-2'-citidina também em outras linhagens de câncer de pulmão. Ensaios de análise funcional demonstraram atividade supressora de tumor para este gene. A super-expressão do gene nas linhagens NIC-H522 e NIC-H358 resultou em redução do potencial de proliferação celular. Camundongos do tipo "*nude*" inoculados com a linhagem NIC-H358 transfectada com o gene *Hint/PCKI-1* apresentaram tumores com crescimento significativamente mais lento.

KARPF et al. (2004) utilizaram a mesma estratégia para analisar qualitativamente o painel de genes reativados após tratamento com 5-aza-dC em 3 linhagens celulares de adenocarcinoma de cólon e cultura de células epiteliais de tecido mamário humano (HMEC). Cerca de 5,3% do total de 25.940 genes analisados apresentou diferenças no padrão de expressão após o tratamento em pelo menos uma das linhagens analisadas. Interessantemente, genes envolvidos com proliferação e diferenciação celular, apresentação de antígenos, via de sinalização de interferon, assim

como antígenos *Cancer/Testis* (CT), encontravam-se freqüentemente reativados após tratamento com agente desmetilante.

Vale ressaltar ainda o trabalho realizado por SHI et al. (2003) que, numa análise conjunta da expressão gênica, metilação do DNA e de proteínas que interagem com DNA, deu os primeiros passos na dissecção dos mecanismos epigenéticos de silenciamento gênico. Uma linhagem de câncer de ovário foi submetida ao tratamento combinado ou individual com 5-aza-dC e um inibidor da enzima Histona Deacetilase (HDAC), a Tricostatina A (TSA). Os genes que apresentaram indução da expressão puderam ser divididos em três classes distintas: grupo 1, re-induzidos após desmetilação de seus promotores gênicos; grupo 2A, cujos promotores encontravam-se desmetilados na amostra não tratada e foram re-induzidos após tratamento apenas com TSA ou TSA + 5-aza-dC, e grupo 2B, que contém a maioria dos genes reativados neste trabalho e que foram re-induzidos independentemente do tratamento com 5-aza-dC, TSA ou a combinação das duas drogas. Assim, os autores sugerem que o grupo 1 seja fortemente regulado por metilação de DNA, o grupo 2A seja regulado pela compactação da cromatina por HDACs, e o grupo 2B tenha sua reativação como efeito indireto do tratamento.

Uma série de estudos utilizando a mesma estratégia foi publicada nos últimos anos, relatando o sucesso na identificação de genes regulados por mecanismos epigenéticos. Um resumo destes estudos está mostrado na Tabela 3.

Tabela 3 - O tratamento de Linhagens celulares tumorais com 5aza-2'-deoxicitidina para a identificação de genes diferencialmente metilados.

Autores	*Tratamento*	*Linhagens celulares tratadas (tumor de origem)*	*Genes analisados*		*Valor de cutoff*	*Plataforma*
			dif. exp.(%)	*total*		
Suzuki et al. 2002	200nM 5AZA e/ou 300nM TSA	*RKO* (Câncer colorretal)	74 (0,74%)	~10000	n.d.	*microarray (m)*
Yuan et al. 2004	1 µM 5AZA	*NCI-H522* (Câncer de pulmão)	73 (6,2%)	1176	>5	*microarray (m)*
Chen et al. 2004	1 µM 5AZA	*RF1* e *RF48* (câncer de estômago)	101	n.d.	n.d.	*microarray (l)*
Takai et al. 2005	5 µmol/L 5AZA + 2,5 µmol/L SAHA	*Ishikawa* (câncer de endométrio)	676 (3%)	~22200	>2	*microarray (l)*
Yamashita et al. 2006	1 µM 5AZA	*AGS, MKN28, MKN45* e *KATOIII* (câncer de estômago)	1430 (3,6%)	~39000	>4	*microarray (l)*

(5AZA) 5-aza-2-deoxicitidina; (TSA) Tricostatina A; (SAHA) Suberoilanilida bis-hidroxamida; (l) lâmina de vidro; (m) membrana; (n.d.) não disponível

1.6 ANÁLISE GLOBAL DA EXPRESSÃO GÊNICA

Diversas metodologias foram desenvolvidas para a identificação em larga escala de genes diferencialmente expressos. Dentre as mais utilizadas, a técnica de *microarray* fornece uma poderosa plataforma de análise do padrão de expressão de milhares de genes simultaneamente, possibilitando a comparação entre diferentes amostras, podendo ser estas representativas de diferentes estágios do desenvolvimento, diferentes tipos celulares ou histológicos, diferentes sensibilidades a drogas ou entre tecidos normais e patologias. Muitos trabalhos na literatura relatam o uso de *microarrays* na pesquisa do câncer de mama. VAN 'T VEER et al. (2002) analisaram amostras de tumores primários de 98 pacientes através da técnica de *microarray* visando identificar uma assinatura molecular com forte valor preditivo para recaídas em pacientes com ausência de células tumorais nos linfonodos regionais ao diagnóstico. Um conjunto de aproximadamente 5000 genes pôde dividi-os em dois grupos, de melhor e de pior prognóstico. Além disso, o grupo também foi capaz de determinar uma assinatura molecular capaz de identificar os pacientes portadores de tumores com mutações em *BRCA1*.

Apesar do sucesso da utilização da técnica na identificação de genes diferencialmente expressos, a utilização de lâminas de microarray possui algumas limitações. A quantidade de genes analisados se restringe ao painel gênico fixado na lâmina. Além disso, a produção de lâminas de microarrays de cDNA é laboriosa e demanda alto investimento financeiro para a

aquisição de scanners e do robô utilizado para a fixação do DNA na lâmina. É necessária a participação de profissionais especializados para a realização da normalização e análise estatística dos dados gerados pelas hibridizações e há grande dificuldade na comparação entre experimentos diferentes.

Outra metodologia que permite a análise do padrão de expressão em larga escala foi desenvolvida por VELCULESCU et al. (1995). A técnica de SAGE (*Serial Analysis of Gene Expression*) baseia-se na quantificação de *tags*, ou seja, pequenas seqüências de 14 nucleotídeos exclusivas de cada transcrito, oriundas de bibliotecas construídas a partir de mRNA proveniente dos tecidos a serem estudados. Assim, uma amostra de tecido tem o painel de genes expressos representados por *tags* e, por conseqüência, o número de vezes que determinada *tag* aparece em cada amostra reflete sua expressão relativa. A expressão gênica de duas amostras pode então ser comparada utilizando um programa computacional que identifica e registra o número de vezes que determinada seqüência aparece na amostra. Outras vantagens desta técnica são o possível cruzamento de dados de experimentos diferentes e a criação de bancos de dados públicos de seqüências expressa em bibliotecas de SAGE. ABBA et al. (2005), analisando bibliotecas de SAGE construídas a partir de tumores de mama positivos e negativos para o receptor de estrógeno (RE), identificaram um grupo de genes capaz de separar amostras RE^+ e RE^-. Além disso, o trabalho ajudou a dissecar o painel de genes diferencialmente expressos entre as bibliotecas, classificando-os segundo sua função biológica e

revelando as principais classes de genes afetadas pela expressão do Receptor de Estrógeno. Além disso, identificaram 220 Elementos Responsivos ao Estrógeno (ERE) na região promotora de 163 dos 473 genes hiper-expressos em tumores RE^+. No entanto a construção de bibliotecas de SAGE, além de extremamente laboriosa e, por conseqüência disto, ter chances aumentadas de introdução de artefatos, demanda a utilização de um programa computacional especializado para o processamento dos dados gerados.

Assim, alternativas para a identificação em larga escala de genes diferencialmente expressos tornam-se necessárias. Nesse contexto, técnicas de subtração são promissoras e particularmente atrativas para a identificação de genes diferencialmente expressos, sem a necessidade de conhecimento prévio de suas características funcionais ou bioquímicas, bem como de utilização de ferramentas computacionais. Dentre elas, destacam-se as metodologias de SSH (DIATCHENKO et al. 1996) e de RaSH (JIANG et al. 2000).

1.7 TÉCNICAS DE HIBRIDIZAÇÃO SUBTRATIVA DE cDNA

A técnica de SSH (*Supression Subtractive Hybridization*) foi desenvolvida por DIATCHENKO et al. (1996) e consiste em analisar a expressão diferencial entre duas amostras de cDNA, subtraindo os transcritos comuns e isolando transcritos presentes exclusivamente em uma das amostras. A técnica permite, por exemplo, a análise de células

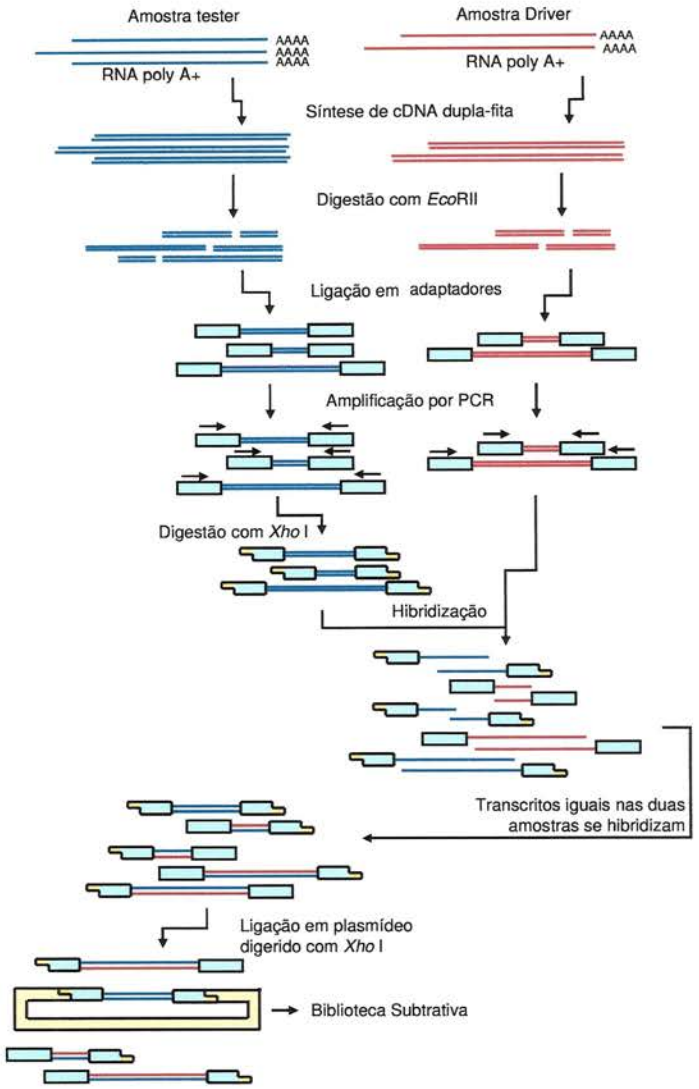
submetidas a diferentes condições de cultura (antes e depois do tratamento com drogas ou outros fatores) ou sincronizadas em diferentes etapas do ciclo celular, possibilitando a clonagem de genes diferencialmente expressos para posterior seqüenciamento, identificação e caracterização.

Nesta técnica, duas amostras de cDNA são sintetizadas a partir de mRNA. Uma das amostras, denominada *tester*, é originada a partir do RNA da linhagem celular ou tecido cujos genes se quer identificar e a outra amostra, denominada *driver*, é construída a partir do RNA da linhagem celular ou tecido utilizado para a comparação e cujos genes devem ser subtraídos. Seqüências adaptadoras são ligadas apenas à extremidade das moléculas de cDNA derivados da amostra *tester*. Em seguida, as moléculas de cDNA da amostra *tester* são desnaturadas e renaturadas na presença de um excesso de moléculas da amostra de cDNA *driver*. Três tipos de moléculas híbridas são formadas após a renaturação: *tester-tester*, *tester-driver* e *driver-driver*. As moléculas *tester-tester* representam transcritos expressos exclusivamente na amostra de interesse, as moléculas *tester-driver* representam transcritos comuns às duas amostras e as moléculas *driver-driver* correspondem a transcritos expressos apenas na amostra utilizada para comparação. As moléculas *tester-tester* são seletivamente amplificados em uma reação de PCR contendo *primers* correspondentes à seqüência adaptadora ligada apenas na extremidade das moléculas de cDNA derivadas da amostra *tester*. O produto dessa reação de PCR é então clonado, gerando uma biblioteca subtraída enriquecida para genes diferencialmente expressos na amostra em estudo. A comparação de

seqüências oriundas da biblioteca subtraída com seqüências depositadas em bancos de dados públicos permite então revelar a identidade dos genes diferencialmente expressos.

A técnica de RaSH (*Rapid Subtractive Hybridization*) (JIANG et al. 2000) foi desenvolvida a partir de uma adaptação da técnica SSH. Duas amostras de cDNA *tester* e *driver* são sintetizadas a partir de mRNA. As moléculas de cDNA de ambas as amostras são digeridas com enzimas de restrição para gerar fragmentos de menor tamanho e, ao contrário da técnica de SSH, seqüências adaptadoras são ligadas às extremidades das moléculas de cDNA de ambas as amostras. Após a ligação dos adaptadores é realizada uma reação de PCR utilizando *primers* complementares à seqüência contida nos adaptadores. Essa etapa permite a obtenção de uma maior massa dos fragmentos representados em ambas as amostras de cDNA, facilitando a etapa de subtração. As seqüências adaptadoras possuem o sítio para a enzima de restrição *Xho*I. Somente os produtos de PCR originados a partir da amostra *tester* são digeridos com a enzima *Xho*I, gerando extremidades coesivas. Após a digestão com a enzima de restrição, os produtos de PCR da amostra *tester* são desnaturados e renaturados na presença de um excesso de fragmentos amplificados a partir da amostra *driver*. Como descrito anteriormente para a técnica de SSH, três tipos de fragmentos híbridos são formados, mas apenas os híbridos *tester-tester* são selecionados através de clonagem em vetor com extremidades coesivas, previamente digerido com a enzima *Xho*I. Um esquema ilustrativo da técnica de RaSH está representado na Figura 6.

FUNDAÇÃO ANTONIO



Legenda: As bibliotecas de cDNA dupla-fita são obtidas das linhagens tratada e não tratada com 5-azacitidina. Adaptadores com sítio para enzima de restrição são ligados às duas bibliotecas. Somente a biblioteca utilizada como *tester* sofre digestão com enzima de restrição correspondente. A biblioteca *tester* é hibridizada com excesso da biblioteca *driver*. Assim, transcritos presentes nas duas bibliotecas formam híbridos. Somente transcritos presentes somente na biblioteca *tester* apresentarão os sítios complementares nas duas extremidades e poderão ser clonados e posteriormente seqüenciados.

Figura 6 - Esquema representando a técnica de RaSH.

A técnica de RaSH apresenta vantagens como redução no custo e tempo de preparo da biblioteca subtraída, além da facilidade na execução dos experimentos e maior eficácia de subtração. JIANG et al. (2000) utilizaram a técnica para analisar o padrão de expressão diferencial de células de melanoma humano tratadas com IFN-β e Mezereína, um ativador de proteína quinase C. Uma biblioteca de cDNA *tester* foi construída a partir de RNA polyA+ extraído de células tratadas com a droga e a biblioteca *driver* foi construída a partir de material de células não tratadas. A técnica de RaSH subtraiu as seqüências presentes nas duas bibliotecas identificando, dentre os 70 genes analisados, 23 (33%) diferencialmente expressos e associados ao tratamento. Outros estudos, utilizando a técnica de RaSH, mostraram a identificação e confirmação por Northern Blot Reverso de 15 genes com expressão elevada (SU et al. 2002) e 10 com expressão diminuída (SU et al. 2003) em astrócitos contaminados com HIV-1. Comparando duas linhagens celulares de melanoma humano com potenciais metastáticos diferentes, BOUKERCHE et al. (2004) também através da técnica de RaSH e validação por Northern Reverso, identificaram 8 genes diferencialmente expressos em linhagens celulares de alto potencial metastático, sendo 2 deles desconhecidos e outros 2 cuja participação na progressão tumoral ainda não havia sido descrita. A técnica de RaSH também foi utilizada na identificação do gene *HuUBP43* como diferencialmente expresso em melanomas em função do tratamento com IFN-β e Mezereína (KANG et al. 2001), e também na identificação de genes induzidos por soforose, um ativador de celulases em *Hypocrea jecorina* (SCHMOLL et al. 2004; SCHMOLL e KUBICEK 2005).

Os dados apresentados demonstram a grande utilidade da técnica de RaSH na identificação de genes diferencialmente expressos entre amostras de origem diversas sob diferentes contextos.

A metodologia de RaSH não necessita de programas computacionais para a identificação e validação dos genes diferencialmente expressos. Neste caso, estas etapas podem ser realizadas com a utilização do Northern Blot Reverso, o qual baseia-se na fixação de fragmentos amplificados a partir dos clones gerados pelo RaSH em membranas de Nylon e a subseqüente hibridização com cDNA das amostras analisadas marcado radioativamente. Os fragmentos que apresentarem diferenças de expressão podem ser recuperados e seqüenciados. Por outro lado, o seqüenciamento direto dos clones gerados pelo RaSH e a análise das seqüências geradas podem facilitar o processo de seleção de genes para a validação experimental.

O objetivo do presente estudo consistiu na utilização da técnica de RaSH em associação ao tratamento com agente desmetilante visando a identificação, em larga escala, de genes diferencialmente metilados em linhagens celulares tumorais de mama. A identificação destes genes certamente contribui para um melhor entendimento sobre a biologia dos tumores assim como para a identificação de novos marcadores tumorais.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Identificar genes diferencialmente metilados em linhagens celulares tumorais de mama através de tratamento com agente desmetilante seguido de análise global da expressão gênica através da técnica de RaSH e validação por PCR em Tempo Real.

2.2 OBJETIVOS ESPECÍFICOS

- Induzir a re-expressão gênica de genes silenciados por metilação em 3 linhagens celulares tumorais de mama através de tratamento com o agente desmetilante 5-aza 2'-deoxicitidina.
- Identificar os genes cuja expressão foi induzida pelo tratamento com 5-aza 2'-deoxicitidina através da construção de 6 bibliotecas subtrativas de cDNA derivadas de RNA total das 3 linhagens celulares, antes e após o tratamento, utilizando para tanto a técnica de RaSH.
- Realizar o seqüenciamento de 300 clones de cada uma das 6 bibliotecas subtrativas e analisar as seqüências geradas a fim de identificar genes expressos exclusivamente nas bibliotecas subtrativas das linhagens tratadas com 5-aza 2'-deoxicitidina.

- Validar experimentalmente o padrão de expressão dos genes induzidos por 5-aza 2' deoxicitidina utilizando para tanto a técnica de PCR em Tempo Real partindo de RNA total das linhagens celulares antes e após o tratamento com o agente desmetilante.
- Avaliar a expressão dos genes validados em 18 amostras tumorais de mama em relação a um pool de amostras de tecido mamário normal

3 MATERIAIS E MÉTODOS

3.1 LINHAGENS CELULARES TUMORAIS

As linhagens celulares tumorais utilizadas foram obtidas comercialmente (ATCC, American Type Culture *Collection*). As três linhagens celulares tumorais são provenientes de carcinoma ductal invasivo e foram isoladas a partir de sítio metastático. A linhagem MCF-7 (ATCC, HTB-22), foi isolada de uma mulher de 69 anos (SOULE et al. 1973). A linhagem MDA-MB-435 (ATCC, HTB-129), foi isolada de uma mulher de 31 anos (ENGEL et al. 1978). A linhagem MDA-MB-231 (ATCC, HTB-26), foi isolada de uma mulher de 51 anos (CAILLEAU et al. 1978).

3.2 AMOSTRAS TECIDUAIS DE CARCINOMA MAMÁRIO HUMANO

As amostras teciduais de carcinoma mamário humano de diferentes estádios clínicos (segundo o sistema de Classificação TNM dos tumores malignos da União Internacional contra o Câncer) provenientes do Banco de Tumores do Hospital do Câncer A.C. Camargo e do Hospital do Câncer Alfredo Abrão – MS (fornecidos pelo Prof. Dr. Ismael Guerreiro da Silva, Laboratório de Ginecologia Molecular – UNIFESP) foram mantidos a -70°C em solução RNA*later*™ (Ambion) para posterior obtenção do RNA total. O

Termo de Consentimento de doação, seguindo as normas institucionais dos hospitais envolvidos foi obtido de todos os pacientes. A Tabela 4 contém um resumo dos dados clínico-patológicos das amostras tumorias de mama.

Tabela 4 – Distribuição dos dados clínico-patológicos das 18 amostras de tumores primários de mama de pacientes do Hospital do Câncer Alfredo Abrão e do Hospital do Câncer A. C. Camargo.

Amostra	*Data Diag*	*TNM*	*T*	*N*	*M*	*Linf*	*N. Linf*	*RE*	*RP*	*Quim*
51T	15.03.04	1	1	0	0	0	0/15	1	1	N/R
110T	17.09.04	1	1	0	0	0	0/10	0	0	N/R
51	26.04.01	2a	2	1	0	1	1/23	1	1	N/R
60	08.05.03	1	1	0	0	0	0/15	0	0	N/R
64	06.09.01	1	1	0	0	0	0/23	1	0	N/R
12T	09.12.03	2b	2	1	0	1	14/16	-	-	N/R
91T	06.07.04	1	1	0	0	0	0/9	1	0	N/R
40T	19.02.04	2b	3	0	0	0	-	-	-	N/R
57T	24.03.04	2b	3	0	0	0	-	0	0	N/R
80T	03.06.04	1	1	0	0	0	0/13	1	1	N/R
56T	22.03.04	3a	3	1	0	1	14/15	1	0	N/R
50T	15.03.04	3b	4	1	0	1	4/6	1	1	N/R
123T	26.10.04	2a	1	1	0	1	12/14	1	1	N/R
52T	17.03.04	3b	4	1	0	1	3/7	-	-	N/R
126T	03.11.04	4	4	1	1	1	9/14	-	-	N/R
15T	15.12.03	4	4	1	1	1	11/13	-	-	N/R
39T	18.02.04	4	4	1	1	1	19/19	-	-	N/R
42	16.04.02	4	2	x	1	1	15/24	-	-	N/R

(Amostra) Identificação da amostra no laboratório; (Data Diag) data na qual o tumor foi diagnosticado; (TNM) estádio clínico do tumor segundo a Classificação TNM; (T) tamanho do tumor; (N) presença de linfonodos comprometidos, sendo x=não pôde ser realizado na paciente; (M) presença de metástase em sítios secundários; (Linf) presença de linfonodos comprometidos ou não, sendo 0=negativo e 1=positivo; (N.Linf) número de linfonodos comprometidos dentre os analisados em cada paciente; (RE) status do receptor de estrógeno, sendo 0=negativo e 1=positivo; (RP) status do receptor de progesterona, sendo 0=negativo e 1=positivo; (-) parâmetro não foi analisado na paciente ou não se encontra disponível no prontuário; (Quim) Tratamento quimioterápico neo-adjuvante, sendo R=realizado e N/R=não realizado.

3.3 CONDIÇÕES DE CULTURA E MANUTENÇÃO DAS LINHAGENS CELULARES

As linhagens de carcinoma mamário humano foram cultivadas a temperatura de 37°C em frascos plásticos descartáveis em atmosfera de 2%

CO_2/ 98% ar, para a manutenção do pH próximo ao fisiológico. As diferentes linhagens foram mantidas em meio de cultura RPMI (desenvolvido no "Roswell Park Memorial Institute") (Gibco) suplementado com 10% de soro fetal bovino (Cultilab Materiais para Cultura de Células), 1,2 g/L de bicarbonato de sódio, 25 mg/L de ampicilina e 100mg/L de estreptomicina.

3.4 TRATAMENTO DAS LINHAGENS CELULARES COM AGENTE DESMETILANTE

As linhagens celulares tumorais de mama foram tratadas com o agente desmetilante 5-Aza-2'-deoxicitidina (5Aza-dC) (SIGMA). 10^6 células foram plaqueadas em placas de 100 mm. 48h após o plaqueamento as células foram tratadas com 1 μM de 5-aza-dC. Após 48h de tratamento, a cultura foi lavada uma vez em PBS e novo meio de cultura foi adicionado às células, as quais permaneceram em cultura por mais 48h para a extração de RNA.

3.5 EXTRAÇÃO DE RNA TOTAL

O RNA total das linhagens celulares e dos tumores foi extraído através da técnica de Sedimentação em Cloreto de Césio (CHIRGWIN et al. 1979). Para extração de RNA total das linhagens celulares, as células foram lisadas em 9ml de solução de lise (4M Isotiocianato de Guanidina, 25mM Citrato de Sódio (pH7,0), 1M β-Mercaptoetanol). Em seguida, o lisado foi

aplicado sobre 4ml de uma solução de Cloreto de Césio (5,7M CsCl, 25mM NaAc) e submetido a ultracentrifugação a 150.000xg em rotor SW 50.1, a 20°C por, aproximadamente, 18h em ultracentrífuga Beckman L8-80M (Beckman Instruments). Após a centrifugação, o precipitado no fundo do tubo foi solubilizado em água bidestilada estéril tratada com Di-etil-pirocarbonato (DEPC).

Os tecidos de carcinoma mamário humano obtidos foram macerados manualmente em cadinhos de porcelana na presença de nitrogênio líquido e, posteriormente, submetidos aos procedimentos de extração e purificação de RNA total conforme descrito no parágrafo anterior.

A quantificação do RNA foi feita através da leitura de uma alíquota da amostra em espectrofotômetro com comprimento de onda equivalente a 260nm, considerando 1 Abs a 260nm equivale a 40mg/mL de RNA. A relação entre as leituras realizadas a 260 e 280nm foi utilizada como parâmetro na estimativa do grau de pureza do RNA. Para que fosse considerado puro, o valor desta relação deveria ser próximo a 2,0. A integridade do RNA total foi verificada através de eletroforese em gel de agarose.

3.6 HIBRIDIZAÇÃO DE SUBTRAÇÃO RÁPIDA - RaSH

A Hibridização de Subtração Rápida foi realizada em colaboração com o grupo do Dr. Wilson Araújo da Silva Jr da Faculdade de Medicina da

Universidade de São Paulo - Campus de Ribeirão Preto, conforme descrito por JIANG et al. (2000) e descrito a seguir.

3.6.1 Síntese de cDNA dupla-fita

A síntese do cDNA dupla-fita foi realizada a partir de RNA total das linhagens celulares acima descritas tratadas ou não com o agente desmetilante. Para a síntese foi utilizado o kit *Superscript™ Double-stranded cDNA Synthesis* (Invitrogen), seguindo as instruções do fabricante. Resumidamente, 1µg de RNA total foi misturado a 1µL de oligodT (0,5µg/µL) e a 4µL de H_2O DEPC, em um volume final de 12µL. Esta mistura foi incubada a 70°C durante 10min e em seguida foi mantida em gelo. À mistura foram adicionados 4µL de tampão 5X *First Strand Buffer*, 2µL de 0,1M DTT, 1µL de dNTPs 10mM, e 1µL de H_2O. Esta reação foi incubada a 37°C durante 2min e, em seguida, foi adicionada à mistura 3µL da enzima Transcriptase Reversa Superscript II (200U/µL), seguida de incubação a 37°C durante 30min. Nova alíquota de 3µL da enzima Transcriptase Reversa foi adicionada e a reação foi incubada a 37ºC por mais 30min. A segunda fita foi sintetizada adicionando-se 94µL de H_20 DEPC, 30µL de tampão 5X *Second Strand Buffer*, 3,0µL de 10mM dNTPs, 1,0µL de E. coli DNA ligase (10U/µL), 4,0µL de E. coli DNA polimerase (10U/µL) e 1,0µL de E. coli RNase H (2U/µL). Esta reação foi incubada a 16°C durante 2h, seguida de incubação em gelo. Dois microlitros de T4 polimerase (5U/µL) foram adicionados à reação e, em seguida, foi realizada nova incubação a 16°C por 5min. Dez microlitros de 0,5M EDTA (pH 8,0) foram adicionados. Uma incubação a 70°C foi realizada durante 10min e, em seguida, 50µL de H_20

DEPC e 200µL de fenol: clorofórmio: álcool isoamílico (25:24:1) foram adicionados. A reação foi homogeneizada e centrifugada a 15300xg durante 5min. A fase aquosa foi, então separada e precipitada adicionando-se 1/10 do volume de acetato de sódio 3M (pH 5,2) e o dobro do volume de 100% etanol.

3.6.2 Digestão da dupla-fita de cDNA e ligação nos adaptadores

As amostras de cDNA dupla fita das linhagens celulares tratadas e não tratadas com o agente desmetilante foram digeridas utilizando-se 20U da enzima de restrição *Mbo*I (Fermentas) por 3h a 37°C, com adição de nova alíquota de 20U da enzima e incubação durante toda a noite. Após a purificação do cDNA pelo método fenol/clorofórmio, seqüências adaptadoras em uma concentração final de 10uM (XDPN-14 CTGATCACTCGAGA; XDPN-12 GATCTCTCGAGT) contendo o sítio para a enzima de restrição *Xho*I foram adicionadas às amostras, que foram incubadas a 55ºC por 1min sendo então resfriadas 2ºC a cada 2min até atingirem a temperatura de 14ºC. Para a ligação dos adaptadores às extremidades dos cDNAs digeridos, foram adicionadas à amostra 15U da enzima T4 ligase (New England Biolabs) em volume final de 30µL, incubando-se durante toda a noite a 14°C. O volume das reações foi completado para 100µL adicionando-se 10mM Tris e 1mM EDTA (pH 7,0). Um volume mínimo de 40µL de cada ligação diluída foi utilizado para amplificação por PCR com oligonucleotídeo desenhado com base nas seqüências dos adaptadores (XDPN-18 CTGATCACTCGAGAGATC). Visando a obtenção de maior massa de

produto amplificado foram realizadas 20 reações idênticas, utilizando-se: 1µL da reação de ligação diluída, 1X PCR *Buffer*, 1µM dNTPs, 10 µM do oligonucleotídeo XDPN-18 e 1U da enzima Taq polimerase (Phoneutria). As reações foram submetidas a 1 ciclo de 5min a 72°C, seguido por 25 ciclos de 1min a 94°C, 1min a 55°C e 1min a 72°C e 1 ciclo de 3min a 72°C. Após a amplificação as 20 reações foram misturadas, e os fragmentos foram analisados em gel 1% agarose. Para a purificação das reações de amplificação, o volume final foi dividido em 4 tubos (250µL cada). A cada tubo foi adicionado o mesmo volume de fenol: clorofórmio: álcool isoamílico (25:24:1), sendo estes, em seguida, submetidos a centrifugação a 15300xg durante 5min. A fase aquosa foi transferida para novo tubo e precipitada através da adição de 62,5µL de 10M Acetato de Amônio, 3µL de Glicogênio (20mg/mL) e 1mL de 100% Etanol. Após precipitação durante toda a noite a -20ºC, a amostra foi centrifugada por 20min a 15300xg e o *pellet* foi lavado em seguida em 500µL de 70% Etanol. O mesmo procedimento foi realizado 3 vezes. Após a purificação, as amostras foram quantificadas e 10µg dos fragmentos derivados da amostra escolhida como *tester* foram digeridos utilizando-se 20U da enzima *Xho*I (New England Biolabs) durante 2h seguida de adição de mais 20U, e digestão por mais 2h. Após a digestão os fragmentos foram purificados com fenol: clorofórmio:álcool isoamílico (25:24:1) seguidos de precipitação na presença de Etanol como descrito anteriormente.

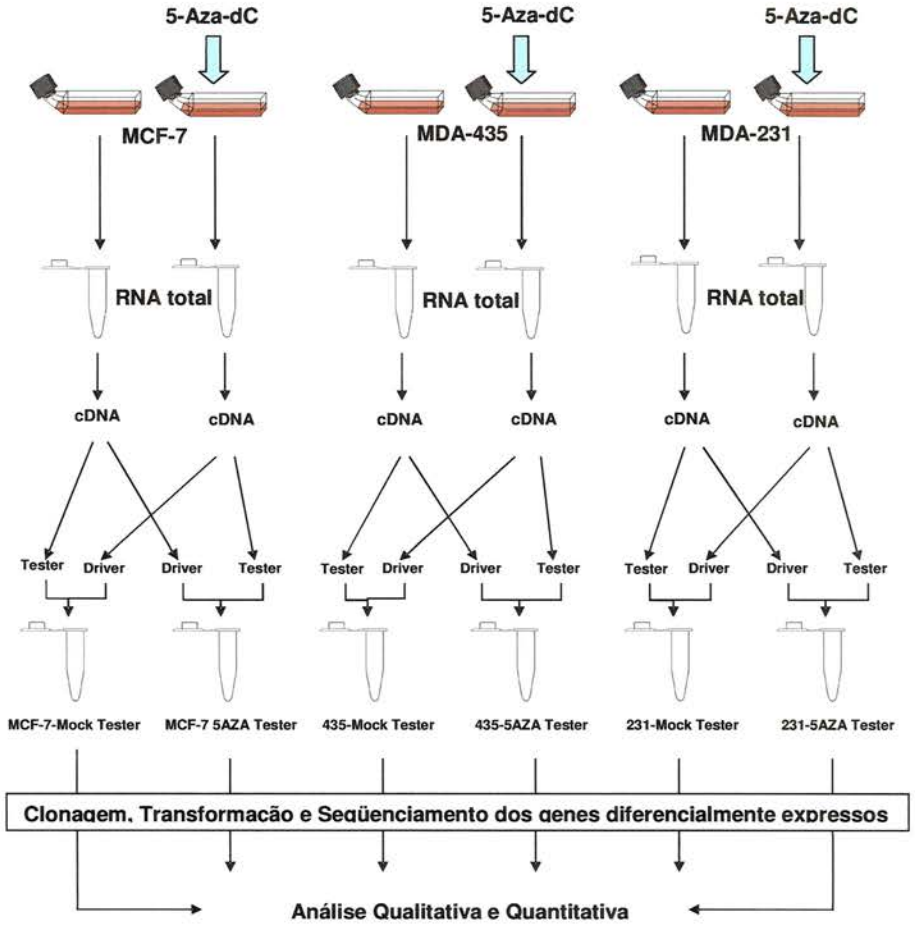
3.6.3 Hibridização e construção das bibliotecas subtrativas

Para cada linhagem celular utilizada, dois experimentos foram realizados. No primeiro experimento, o cDNA derivado da linhagem tratada com o agente desmetilante foi utilizado como *tester* e o cDNA da linhagem não tratada foi utilizado como *driver*. No segundo experimento, chamado "Experimento Reverso", o cDNA derivado da linhagem não tratada foi usado como *tester* e o cDNA derivado da linhagem tratada foi utilizado como *driver*. Ao todo foram construídas 6 bibliotecas subtrativas. Um esquema ilustrativo representando a estratégia utilizada está mostrado na Figura 7.

Através deste desenho experimental, foi possível a obtenção de um painel de genes expressos exclusivamente nas linhagens celulares tratadas com a droga (cuja expressão gênica foi reativada após tratamento) e outro painel representando os genes expressos apenas nas linhagens não tratadas. Após a comparação entre os painéis gênicos de cada uma das 6 bibliotecas, eventuais falhas na subtração puderam ser identificadas, como, por exemplo, a presença de genes tanto no painel da linhagem tratada quanto no painel da linhagem não tratada.

Para a hibridização e construção da biblioteca subtrativa, 100ng de cDNA *tester*, previamente digerido com enzima *Xho*I, foram hibridizados com 5µg de produto de PCR amplificado a partir do cDNA não digerido derivado da amostra *driver*. Os fragmentos foram misturados em 16,6µL de Solução de Hibridização constituída por 0,5M NaCl, 50mM Tris (pH 7,5), 0,2% SDS, 40% (vol/vol) Formamida. A mistura foi fervida por 5min e incubada por 48h a 42°C. Depois da purificação através do método fenol/clorofórmio, parte da mistura (1µL) foi ligada a 1µg de vetor (pZErO – I, Invitrogen) previamente

digerido com a enzima *Xho*I (New England Biolabs). Bactérias eletrocompetentes (DH10B) foram transformadas com 2µL da ligação. Os clones foram selecionados em placa contendo meio C.G. (Circle Grow) e antibiótico Zeocina (Invitrogen).



Legenda: Cada linhagem foi cultivada na presença ou ausência de 5 aza 2´ deoxicitidina. Dois experimentos individuais foram realizados para cada linhagem. Um dos experimentos utilizará a biblioteca tratada como tester e o outro utilizará como tester a biblioteca controle. Genes diferencialmente expressos foram clonados, transformados e seqüenciados. Em seguida, foram submetidos à análise quantitativa e qualitativa in silico.

Figura 7 - Estratégia adotada para a identificação de genes silenciados por metilação e reativados pelo tratamento com 5-aza-dC.

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenador de Ensino

FUNDAÇÃO ANTONIO PRUDENTE BIBLIOTECA

3.7 SELEÇÃO DE CLONES PARA SEQÜENCIAMENTO

Os clones foram selecionados randomicamente, amplificados por PCR de colônia utilizando *primers* complementares à seqüência do vetor pZErO – 1, sob condições pré-estabelecidas. Os fragmentos amplificados foram seqüenciados utilizando os mesmos *primers* usados para a amplificação, segundo a técnica de terminação com dideoxinucleotídeos (SANGER et al. 1977), utilizando-se o kit *Big Dye Terminator* (Applied Biosystems), seguindo as instruções do fabricante. Os produtos desta reação foram submetidos à eletroforese por capilaridade em seqüenciador automático de DNA ABI 3100 (Applied Biosystems).

3.8 *NORTHERN* BLOT REVERSO

A técnica de "*Northern* Blot Reverso" foi realizada como descrito por KANG et al. (1998), com pequenas alterações como descrito a seguir.

3.8.1 Fixação do DNA em membranas de Nylon

O produto obtido no PCR de colônia foi fixado em duas membranas de Nylon carregadas positivamente (HyBond N+, GE HealthCare), sendo uma das membranas hibridizada com cDNA proveniente da amostra tratada e a sua réplica com cDNA proveniente da amostra não tratada. 20µL da reação de PCR de cada clone foram adicionados a 180µL de Solução Desnaturante (1,6N NaOH; 100mM EDTA) e incubados a temperatura

ambiente por cerca de 10min. Este volume total foi aplicado nas membranas previamente embebidas em 2X SSC (0,3M Cloreto de Sódio; 30mM Citrato Trissódico Dihidratado, pH 7,0) com o auxílio do aparelho Vacuum *Manifold* (BioRad) de 96 poços. A membrana foi exposta a luz ultravioleta por cerca de 3min para fixação do DNA na mesma (etapa de *crosslink*).

3.8.2 Reação de amplificação de RNA para a síntese de cDNA marcado com isótopo radioativo

RNA total extraído da linhagem celular MDA-MB-231 submetida ou não ao tratamento foi amplificado segundo (WANG et al. 2000), visando o aumento da massa de RNA e enriquecimento da amostra total com mRNA. 3µg de RNA total foram adicionados à mistura contendo oligodT (0,5µg) (Integrated DNA Technologies) e volume final de 6µL. Esta mistura foi aquecida a 70ºC por 5min e depois resfriada em gelo. A primeira fita foi sintetizada utilizando o kit *SuperScript II Reverse Transcriptase* (Invitrogen), com 4µL de tampão 5X *First Strand Buffer*, 3µL de *template switch primer* (TS, 0,5 ug/µL), 2µL de 0,1M DTT, 1µL de inibidor de RNase (Invitrogen, 40U/µL), 2µL de 10mM dNTP e 2µL da enzima Transcriptase Reversa Superscript II (200U/µL), em um volume final de 20µL. Após incubação a 42ºC por 2h, a segunda fita foi sintetizada utilizando o kit *Advantage cDNA PCR* (BD Biosciences), adicionando-se ao cDNA sintetizado, 65,3µL de H_2O, 10µL do tampão *Advantage PCR Buffer*, 2µl de 10mM dNTP, 0,7µL de RNase H (2U/µL), 2µL da enzima DNA polimerase (*Advantage Polymerase,* 5U/µL), em um volume final de 100µL. A mistura foi incubada

sucessivamente a: 37ºC por 10min (digestão do RNA), 94ºC por 3min (desnaturação), 65ºC por 5min (anelamento dos *primers*) e 75ºC por 30min (extensão da segunda fita). A reação foi inativada através da adição de 5µL de solução contendo 1M NaOH e 2mM EDTA e incubando-se a 65°C por 10min. A transcrição *in vitro* de mRNA foi realizada utilizando o kit RiboMax (Promega). O cDNA dupla fita foi adicionado à mistura contendo 7,5µL de nucleotídeos a 25mM (ATP, GTP, CTP e UTP), 5µL de 5X *Reaction Buffer*, 2,5µL de mistura contendo inibidor de RNase e T7 *Phage Polymerase*, em um volume final de 25µL. A transcrição ocorreu durante incubação a 37ºC por um período de 5 a 6h. O RNA amplificado foi utilizado para a síntese de cDNA e simultânea marcação com isótopo radioativo.

3.8.3 Marcação de cDNA com isótopo radioativo

A marcação do cDNA das linhagens celulares com radioativo foi realizada segundo KANG et al. (1998) com pequenas alterações. 6µg de RNA amplificado conforme descrito no Item 3.7.2 foram misturados a 9µg de uma mistura de *primers* randômicos, em um volume final de 16µL. Esta mistura foi aquecida a 70ºC por 10min sendo, em seguida, adicionados 8,0µL de tampão 5X *First Strand Buffer*, 4µL de 0,1M DTT, 5,0µL de dNTPs (mistura de nucleotídeos dATP, dTTP e dGTP a 25mM e dCTP a 10mM) e 0,5µL de inibidor de RNase (40U/µL). A seguir foi adicionado à mistura 10µL de $\alpha dCTP^{32}$, sendo esta aquecida a 43ºC por 5min. Uma alíquota de 1µL da enzima Transcriptase Reversa SuperScript II (200U/µL) foi adicionada. Após 1h nova alíquota de 1µL de SuperScript II (200U/µL) foi adicionada seguida

de nova incubação a 43°C por mais 1h. A reação foi purificada utilizando colunas de Sefarose G50 (GE HealthCare).

A sonda radioativa foi desnaturada por 5min e então adicionada a um tampão de hibridização (50% [vol/vol] Formamida; 5X Denhardt; 5X SSPE; 0,1% SDS; 1,5% Sulfato de Dextrano) e foi misturada a DNA de esperma de salmão (100µg/mL) para subseqüente hibridização.

3.8.4 Hibridização das membranas com sonda radioativa

As membranas foram pré-incubadas em solução de pré-hibridização (50% [vol/vol] Formamida; 5X Denhardt; 5X SSPE; 0,1% SDS) por, no mínimo, 60min antes da incubação com a sonda marcada radioativamente.

A incubação com a sonda foi realizada durante toda a noite a 42ºC. Após a hibridização, as membranas foram lavadas duas vezes em tampão de lavagem (SSC 2x; SDS 0,1%) a temperatura ambiente por 30min e três vezes em tampão de lavagem (SSC 0,1X; SDS 0,1%) a 56°C por 20min.

3.8.5 Detecção do sinal radioativo

A detecção do sinal radioativo foi realizada através da sensibilização de uma tela de exposição composta por sais de európio (Molecular Dynamics, GE Healthcare). Estas telas são sensíveis a emissões radioativas, que são armazenadas na forma de moléculas de fosfato excitadas. Após a sensibilização, o rastreamento da tela foi realizado em aparelho de *Scanner Storm™ Phosphor Imager* (Molecular Dynamics, GE

Healthcare) e a análise dos resultados foi realizada com auxílio do programa computacional *ImageQuant for Windows NT* versão 5.0.

O gene *GAPDH* foi utilizado como normalizador do experimento. Diferentes diluições do produto amplificado a partir deste gene foram fixadas nas membranas. Um fator de correção foi calculado a partir da divisão dos valores densitométricos do gene *GAPDH* na membrana hibridizada com cDNA proveniente da linhagem tratada pelos valores densitométricos do mesmo na membrana com cDNA proveniente da linhagem não tratada. Assim, todos os valores obtidos dos *spots* na membrana hibridizada com cDNA proveniente da linhagem tratada foram divididos por este fator de correção.

Já o cálculo da expressão gênica diferencial foi realizado utilizando a linhagem não tratada como amostra referência. Assim os valores normalizados obtidos na membrana hibridizada com cDNA proveniente da linhagem tratada foram divididos pelos valores obtidos na membrana hibridizada com cDNA proveniente da linhagem não tratada. Ao final, obtivemos a razão de expressão tratada/ não tratada.

3.9 ANÁLISE *IN SILICO* DAS BIBLIOTECAS SUBTRATIVAS

O laboratório do Dr. Wilson Araújo da Silva Jr. (Centro de Terapia Celular - Faculdade de Medicina de Ribeirão Preto - USP) desenvolveu um conjunto de *softwares* que possibilita a análise e comparação das seqüências geradas a partir das bibliotecas subtrativas. As seqüências

geradas foram submetidas via Internet (http://guarani.fmrp.usp.br/cgi-bin/geap/index.cgi) e automaticamente analisadas por um conjunto de programas computacionais (*pipeline*), de forma seqüencial, conforme a ordem: (i) *Base-calling* e identificação de regiões de baixa qualidade; (ii) Identificação de vetores e regiões repetitivas, e (iii) Busca por similaridade em bancos de dados públicos.

Primeiramente, as seqüências submetidas foram analisados pelo programa de *base-calling* PHREDPHRAP® (*Washington University*) para extrair as seqüências de bases atribuindo um valor de qualidade à cada uma delas de acordo com a sua resolução. Além disso, o programa PHREDPHRAP® indica a localização da região de alta qualidade na seqüência, com essa informação o sistema irá descartar a região de baixa qualidade e determinar se a seqüência será aceita. O critério adotado para considerar a seqüência aceita é possuir no mínimo 100 bases de alta qualidade e o trim_*cutoff* adotado foi 0,05.

A presença de seqüências de vetor e *primers* foram verificadas utilizando o programa *cross_match*, através do alinhamento destas seqüências consideradas de boa qualidade com bancos de dados internos. As regiões repetitivas foram então mascaradas utilizando o programa *RepeatMasker*, com base de dados de seqüências repetitivas de mamíferos.

Após o mascaramento de seqüências de vetor e repetitivas, as demais foram então alinhadas a seqüências depositadas em bancos de dados públicos de seqüências mitocondriais, bacterianas, de levedura, de seqüências de mRNA completas (*RefSeq*), *Unigene* e ESTs humanas,

3.10 SELEÇÃO DE GENES PARA VALIDAÇÃO EXPERIMENTAL

Para validação experimental foram selecionados genes expressos exclusivamente nas linhagens tratadas, ou seja, genes identificados através dos experimentos nos quais o cDNA das linhagens tratadas foi utilizado como *tester* e o cDNA das linhagens não tratadas foi utilizado como *driver*. O experimento reverso, no qual o cDNA da linhagem não tratada foi usado como *tester* e o da linhagem tratada como *driver*, foi utilizado para identificar possíveis falhas na construção das bibliotecas subtrativas e excluir falso-positivos, aumentando a confiabilidade da análise.

Primeiramente, os genes a serem selecionados foram avaliados quanto a presença de Ilhas de CpG nas suas regiões promotoras. Para tanto foi utilizada a interface gráfica do genoma humano disponibilizada pela Universidade de Santa Cruz na Califórnia (www.genome.ucsc.edu). O programa de identificação de Ilhas de CpG analisa a seqüência base a base, estipulando valores a cada dinucleotídeo (+17 se for um CG, e -1 para qualquer outro). Assim, seleciona automaticamente regiões genômicas que possuem alta pontuação (*score*) e que obedeçam aos seguintes critérios pré-definidos: (i) conteúdo de GC de 50% ou mais; (ii) região maior que 200pb e (iii) razão entre o número observado de CG e o número esperado de Gs e Cs for maior que 0,6. Desta forma foram selecionados no total 41 candidatos com Ilha de CpG em sua região promotora para validação por PCR em Tempo Real.

3.11 VALIDAÇÃO DOS GENES POR PCR EM TEMPO REAL

3.11.1 Síntese de cDNA fita simples

Para a validação dos genes candidatos por PCR em Tempo Real, foi utilizado cDNA fita simples obtido a partir do mesmo RNA total utilizado para a construção das bibliotecas subtrativas, sendo, portanto proveniente do mesmo tratamento com 5-aza 2'-deoxicitidina. A síntese do cDNA simples fita foi realizada como descrito no item 3.6.1. 1µg de RNA total foi misturado a 1µL de oligodT (0,5 µg/µL) (Integrated DNA Technologies) e a 4µL de H_2O DEPC, em volume final de 12µL. Esta mistura foi incubada a 70°C durante 10min e, em seguida, mantida no gelo. À mistura foram adicionados 4µL de tampão 5X *First Strand Buffer*, 2µL de 0,1M DTT, 1µL de 10mM dNTPs, e 1µL de H_2O. A reação foi incubada a 37°C durante 2min e, em seguida, foi adicionada a mistura 3µL de enzima Transcriptase Reversa Superscript II (200 U/µL; Invitrogen), seguida de incubação a 37°C durante 30min. Nova alíquota de 3µL de enzima Transcriptase Reversa foi adicionada e a reação foi incubada a 37ºC por mais 30min.

3.11.2 Controle da qualidade do cDNA

Para avaliar a qualidade do cDNA sintetizado, foram realizadas amplificações de PCR específicas para o gene de expressão constitutiva *GAPDH* (Gliceraldeido 3-Fosfato Desidrogenase). Cada reação foi montada de forma que apresentasse uma concentração final de 1X para o tampão de reação da enzima *Taq* (Invitrogen), 1,5mM de $MgCl_2$ (Invitrogen), 0,2mM de

FUNDAÇÃO ANTONIO

dNTP (Invitrogen), 0,2µM de cada um dos *primers* ("*Forward*" ACCCACTCCTCCACCTTTGA e "*Reverse*" CTGTTGCTGTAGCCAAATTCGT), 0,5U da enzima Taq DNA Polimerase (5U/µl, Invitrogen) e 1µL de cDNA template. Nas reações foi utilizado um programa de 94°C por 2min, 35 ciclos de 94°C por 30seg, 56°C por 30seg e 72°C por 45seg e, após os ciclos, 72°C por 6min. As reações foram posteriormente analisadas em gel 1% agarose corado com Brometo de Etídeo.

3.11.3 Desenho de *primers* específicos

Para as reações de PCR em Tempo Real, foram desenhados, com o auxílio do programa *PrimerExpress* versão 3.0 (Applied Biosystems), um par de *primers* específico para cada candidato. Os parâmetros utilizados pelo programa *PrimerExpress* para o desenho dos *primers* são: amplificar fragmentos cujo tamanho varie entre 50 e 150bp, apresentar quantidade de CG entre 30 e 80%, não possuir a capacidade de anelar entre si ou formar estrutura secundária estável, apresentar temperatura de anelamento entre 58°C e 60°C. Além disso, os *primers* foram desenhados em *exons* diferentes do gene candidato visando evitar uma co-amplificação de DNA genômico contaminante e sempre na extremidade 3' do gene, visando minimizar artefatos gerados durante a síntese de cDNA. As seqüências dos *primers* desenhados para cada gene candidato estão representadas na Tabela 6.

Tabela 6 - Seqüências dos *primers* desenhados para a validação experimental. (*Amp, tamanho do amplicon*).

Candidato	*Nº Acesso*	*Anotação*	*Primers*		*Amp (pb)*
BM 0001	NM_003900	Homo sapiens sequestosome 1 (SQSTM1)	Fw	GCAGCTGCCCTTAGCCCT	63
			Rev	TCTTTTCCCTCCGTGCTCC	
BM 0002	NM_002276	Homo sapiens keratin 19 (KRT19)	Fw	AAGGATGCTGAAGCCTGGTTC	66
			Rev	CGTGTGGCCAGCGACC	
BM 0003	NM_003302	Homo sapiens thyroid hormone receptor interactor 6 (TRIP6)	Fw	TGGATCGAAGTTTTCACATTGG	69
			Rev	CGCCCTCAGAGGAGAGCA	
BM 0004	NM_015420	Homo sapiens DKFZP564O0463 protein (DKFZP564O0463)	Fw	CGCCTGTGGAAAGCTAATGC	77
			Rev	TGGTTATAATCCTTGGCTGCTTTT	
BM 0005	NM_033296	Homo sapiens T-cell activation protein (PGR1)	Fw	GAAGAGCGAGTCGTCGTGA	58
			Rev	GTTGACCAGAATCCATTGGC	
BM 0006	NM_000636	Homo sapiens superoxide dismutase 2, mitochondrial (SOD2)	Fw	AAATTGCTGCTTGTCCAAATCA	72
			Rev	CATCAATCCCCAGCAGTGG	
BM 0007	NM_005274	Homo sapiens guanine nucleotide binding protein (G protein), gamma 5 (GNG5)	Fw	CCCAGAAAGTCTGTTCCTTTTTGT	89
			Rev	GCTCTCTTGAATATCACTGGATCATAA	
BM 0008	NM_001862	Homo sapiens cytochrome c oxidase subunit Vb (COX5B)	Fw	CCCCTCCATCTCCAACAAGA	76
			Rev	CAGCCAAAACCAGACGACG	
BM 0009	NM_005749	Homo sapiens transducer of ERBB2, 1 (TOB1)	Fw	CCAGGTGACAGCCCCCTTA	100
			Rev	CCTCCATAGGCTGCAAACACA	
BM 0010	NM_030782	Homo sapiens cisplatin resistance related protein CRR9p (CRR9)	Fw	CGTGGTGTTTCTGGTCTACCTG	71
			Rev	CAAACTCGTTCACTCTGCGTTT	
BM 0101	NM_005566	Homo sapiens lactate dehydrogenase A (LDHA)	Fw	GAGTATAATGAAGAATCTTAGGCGGG	82
			Rev	GAAGACATCATCCTTTATTCCGTAAAG	
BM 0102	NM_016127	Homo sapiens hypothetical protein MGC8721 (MGC8721)	Fw	CCAGAACTGCATCAGGATATGG	83
			Rev	TCCAAAATTTCTGCATCCAGTG	
BM 0103	NM_014372	Homo sapiens ring finger protein 11 (RNF11)	Fw	GGGCTAGCTTTGGCGAGG	92
			Rev	AGGTGTTGGGTGGTAGACTGGA	
BM 0104	NM_003299	Homo sapiens tumor rejection antigen (gp96) 1 (TRA1)	Fw	GAATGCTTCGCCTCAGTTTGA	70
			Rev	TCAGGTTCTTCTTCGGGCTC	
BM 0105	NM_005348	Homo sapiens heat shock 90kDa protein 1, alpha (HSPCA)	Fw	GGATGATCAAACTTGGTCTGGG	81
			Rev	GCATTTCTTCAGTTACAGCAGCA	
BM 0106	NM_002894	Homo sapiens retinoblastoma binding protein 6 (RBBP6)	Fw	AATTTTTGGGAAGTTGGTTTTCC	87
			Rev	TGGACGAGGACAAGGATCAAG	
BM 0107	NM_139070	Homo sapiens mitogen-activated protein kinase 9 (MAPK9), transcript variant 4	Fw	GGGAAGAAAGAAGCAAGAATGG	77
			Rev	AGAAGGAGTGGCGTTGCTACTT	
BM 0108	NM_001253	Homo sapiens CDC5 cell division cycle 5-like (CDC5L)	Fw	TTTTGAAGAACTCAAGAAACATGAAGAT	81
			Rev	TGTTGTCGCTGAACGTCTTCTTT	
BM 0109	NM_000884	Homo sapiens IMP (inosine monophosphate) dehydrogenase 2 (IMPDH2)	Fw	CATGCCAGGACATTGGTGC	72
			Rev	CAAACTTAAGCTCCCCAGAGTACAT	
BM 0110	NM_005638	Homo sapiens synaptobrevin-like 1 (SYBL1)	Fw	GAAGAACCTCAAGCTCACTATTATCATC	103
			Rev	CACACAGCTTGGCCATGTAAA	
BM 201	NM_001901	Homo sapiens connective tissue growth factor (CTGF)	Fw	GAAGAGAACATTAAGAAGGGCAAAA	80
			Rev	CAGCCAGAAAGCTCAAACTTGA	
BM 202	NM_006937	Homo sapiens SMT3 suppressor of mif two 3 homolog 2 (yeast) (SUMO2), transcript variant 1	Fw	TGAACGACAGGGATTGTCAATG	88
			Rev	ATTTCCAACTGTGCAGGTGTGT	
BM 203	NM_001908	Homo sapiens cathepsin B (CTSB), transcript variant 1	Fw	CCAATAGCGAGAAGGACATCATG	111
			Rev	CGTGTTGGTACACTCCTGACTTG	
BM 204	NM_004339	Homo sapiens pituitary tumor-transforming 1 interacting protein (PTTG1IP),	Fw	GGACTACCCAGTTACAAGCGTCTT	89
			Rev	CCTCAAAGTTCACCCAACAAACTC	

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenador de Ensino

Cont/ Tabela 6

Candidato	*Nº Acesso*	*Anotação*	*Primers*		*Amp (pb)*
BM 205	NM_000201	Homo sapiens intercellular adhesion molecule 1 (CD54), human rhinovirus receptor (ICAM1)	Fw	CGCGAGGTGACCGTGAA	67
			Rev	CGGCTGCTACCACAGTGATG	
BM 206	NM_198941	Homo sapiens tumor differentially expressed 1 (TDE1), transcript variant 2	Fw	TGAACCTCTGAGTGCCAAGGA	75
			Rev	CCAGCAGTAATTAGGACATCACTGA	
BM 207	NM_005418	Homo sapiens suppression of tumorigenicity 5 (ST5), transcript variant 1	Fw	GCTGATCTCCCAGGACTCTGA	93
			Rev	TCCACAAAGAACCGGATAAACA	
BM 208	NM_000123	Homo sapiens excision repair cross-complementing 5 (ERCC5)	Fw	GGAACAGAACGAAGACAGATGAATC	118
			Rev	GCATCTTCTTTCTCCTGTTGTGCTA	
BM 209	NM_022121	Homo sapiens PERP, TP53 apoptosis effector (PERP)	Fw	CTTGTCTTCCTGAGAGTGATTGGA	103
			Rev	GGGTGAAGGTCTGGGTGTACTTC	
BM 210	NM_016213	Homo sapiens thyroid hormone receptor interactor 4 (TRIP4)	Fw	GCCTTTGGGAGTTCTGGTAAATC	79
			Rev	CTGTGAGGCTGCACCTGTGT	
BM 301	NM_001256	Homo sapiens cell division cycle 27 (CDC27)	Fw	TCCCAAAGAATCCCTCGTTTAC	105
			Rev	GGATCTAAATCCATAGCCCAAGAG	
BM 401	NM_017824	Homo sapiens membrane-associated ring finger (C3HC4) 5 (MARCH5)	Fw	CAAAGGACAATCTTGGGTGGAA	140
			Rev	AGTCAGTTTTATGCTTCTTCTTGTTCTG	
BM 402	NM_001349	Homo sapiens aspartyl-tRNA synthetase (DARS)	Fw	TGATTTGGAGAAAATTAAGGCTTACA	92
			Rev	AGTAACTCGTTCCAATCCAATGC	
BM 403	NM_006216	Homo sapiens serine (or cysteine) proteinase inhibitor 2 (SERPINE2)	Fw	ATCGCCTCCCTGGTTTATAGTAGA	96
			Rev	GGTTTGTTTATCTGCCCCATGA	
BM 404	NM_002026	Homo sapiens fibronectin 1 (FN1), transcript variant 1	Fw	GCCAGTCCTACAACCAGTATTCTCA	113
			Rev	GGAATCTTCTCTGTCAGCCTGTACA	
BM 405	NM_003155	Homo sapiens ATPase, H+ transporting, membrane associated protein M8-9 (ATP6M8-9)	Fw	CACCTCCCTCATTAGGAAGACAA	118
			Rev	CATGTTGAAAACCACGGAATATTC	
BM 406	NM_005765	Homo sapiens pitrilysin metalloproteinase 1 (PITRM1)	Fw	TGAAGCAGGCCCACAGAGA	82
			Rev	CTTCCCAGTGCCGAGGTATC	
BM 407	NM_014889	Homo sapiens MO25 protein (MO25)	Fw	TGAACCTGCTGCGAGACAAA	73
			Rev	TGAGTTTGGCCTGGTTCTTGA	
BM 501	NM_006888	Homo sapiens calmodulin 1 (phosphorylase kinase, delta) (CALM1)	Fw	TGACAAGGATGGCAATGGTTATA	120
			Rev	CCATCAATATCTGCTTCTCTGATCA	
BM 502	NM_014670	Homo sapiens basic leucine zipper and W2 domains 1 (BZW1)	Fw	CCTTCCAGAAAATAGTGGTGCTTT	99
			Rev	TCTTCCCCTTTGCAACATGTG	
BM 503	NM_003905	Homo sapiens amyloid beta precursor protein binding protein 1, 59kDa (APPBP1)	Fw	CACGAATTTTGCCGATATGGA	94
			Rev	TGGTGATTATTTTGATGACCTCTTG	

3.11.4 Reação de PCR quantitativa em Tempo Real

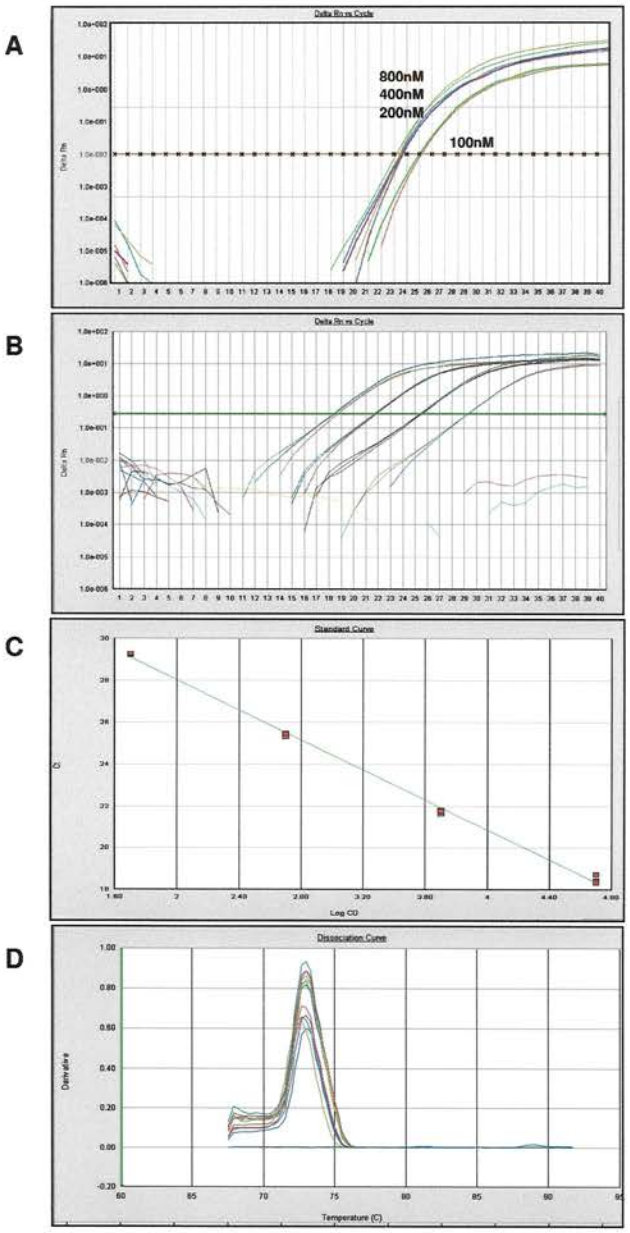
Para a quantificação do produto formado durante a reação de PCR em Tempo Real foi utilizado o reagente *SYBR® Green Dye* (Applied Biosystems). O corante *SYBR® Green Dye* possui afinidade pelo sulco menor da dupla fita de DNA. Quando não está ligado ao DNA, emite uma pequena fluorescência no comprimento de onda de 520ηm, entretanto, quando se liga a dupla fita de DNA, a fluorescência aumenta cerca de 100 vezes, permitindo então a detecção do produto amplificado.

As reações foram realizadas em volume final de 20µL, utilizando 5µL de cDNA diluído 100 vezes a partir do volume final da síntese de cDNA (~5ng), 5µL de *primers* com concentração variando entre 200 e 800nM, de acordo com o candidato, e 10µL do reagente 2X *Master Mix SYBR® Green* (Applied Biosystems). As reações foram realizadas em termociclador ABI Prism® 7300 *Sequence Detection System* (Applied Biosystems) nas seguintes condições: 50°C por 2min (etapa de ativação da enzima *AmpErase®*, um controle do fabricante contra eventuais contaminações com produto amplificado), 95°C por 10min (etapa de ativação da enzima DNA polimerase Taq Gold), e 40 ciclos de 95°C por 15seg (etapa de desnaturação) e 60°C por 1min (etapa de anelamento dos *primers* e extensão do *amplicon*). Após o término da reação, a especificidade da reação foi avaliada através da análise da dissociação do produto amplificado e, quando necessário, através da análise do produto final destas reações em gel de 8% poliacrilamida corado com prata.

O programa 7300 System SDS Software foi utilizado para a análise dos dados obtidos. O programa gera um gráfico onde está plotada a intensidade de fluorescência em função do número de ciclos de amplificação. Um limiar (*threshold*) é definido manualmente, onde a intensidade de fluorescência é estatisticamente diferente da fluorescência de fundo (*background*) e a curva encontra-se na fase logarítmica de amplificação. O ciclo da reação no qual a fluorescência da amostra intercepta o *threshold* é chamado Ct (*Cycle Threshold*). Assim, a comparação entre a expressão gênica de diferentes amostras é realizada

através da comparação entre os Cts. Amostras com Cts menores apresentam maior quantidade do gene alvo no cDNA *template* inicial na reação.

A concentração ideal de *primers* para as reações de PCR em Tempo Real foi obtida através de reações utilizando diferentes concentrações de *primers* e concentração fixa de cDNA. Foi escolhida a menor concentração que obtivesse maior fluorescência, sem a formação de dímeros e variação no Ct (Figura 8A). Pelo fato de ter sido escolhida a quantificação relativa entre as amostras, a eficiência de amplificação do amplicon pelos *primers* teve de ser calculada. Assim a eficiência foi calculada realizando-se reações utilizando quatro diluições seriadas de cDNA como *template* (1:1, 1:10, 1:100 e 1:1000, ver Figura 8B). Assim, foi construída uma curva-padrão, onde a média dos Ct obtidos em cada concentração variava em função do logaritmo da concentração de cDNA. O coeficiente angular da reta obtida (a, em y=ax+b) (Figura 8C) foi utilizado para cálculo da eficiência de amplificação dos *primers*, obtida através da seguinte fórmula: $ef = 10^{-1/a}$ (sendo ef = eficiência, a = coeficiente angular da reta). Foram aceitos valores de eficiência entre 1,70 e 2,30.



Legenda: Padronização das reações de PCR em Tempo Real do candidato BM 201. (A) Diferentes concentrações de *primers* testadas em concentração fixa de cDNA. (B) Diferentes diluições seriadas de cDNA *template* (1:1, 1:10, 1:100, 1:1000). (C) Curva-padrão obtida a partir das reações realizadas em B. (D) Curva de dissociação utilizada para avaliar a especificidade da reação e a possível formação de dímeros de *primers*.

Figura 8 – Exemplo de padronização das reações de PCR quantitativo em tempo real.

3.11.5 Cálculo da Expressão Gênica Diferencial dos Candidatos Selecionados para a Validação Experimental

Os valores de Ct obtidos pelo programa computacional 7300 System SDS (Applied Biosystems) foram utilizados para o cálculo da expressão gênica diferencial entre as linhagens submetidas ou não ao tratamento, assim como entre as amostras tumorais e o *pool* de amostras de tecido mamário normal. No caso das linhagens, a expressão gênica diferencial dos candidatos foi verificada através de quantificação relativa, sendo a linhagem não tratada utilizada como amostra referência. A diferença entre os valores de Ct (ΔCt) entre as amostras, após a normalização com genes de expressão estável, reflete sua expressão gênica diferencial. As reações foram realizadas sempre em duplicata e a média dos valores de Ct das duplicatas foi utilizada para o cálculo da expressão gênica diferencial. O gene que codifica a proteína *β-Actina* foi utilizado como gene normalizador (controle endógeno) e o cálculo da expressão diferencial relativa (razão) entre as amostras foi realizado segundo PFAFFL (2001) utilizando a fórmula descrita na Figura 9. Esta fórmula foi bastante adequada pois considera eficiências de amplificação menores que 100% e diferentes entre o gene alvo e o gene normalizador.

$$razão = \frac{Ef_{\text{gene alvo}}^{(Ct_{alvo\ mock} - Ct_{alvo\ tratada})}}{Ef_{\text{controle endógeno}}^{(Ct_{endógeno\ mock} - Ct_{endógeno\ tratada})}}$$

Fonte: PFAFFL (2001)

Legenda: Ef, eficiência, alvo, gene em análise; endógeno, gene normalizador.

Figura 9 - Fórmula proposta por PFAFFL (2001) para a quantificação relativa por PCR em Tempo Real.

A análise da expressão dos genes que apresentaram reativação maior que 2 vezes nas linhagens celulares foi realizada em 18 amostras de tumores de mama e um pool de amostras de tecido mamário normal obtido comercialmente (CLONTECH®). A quantificação foi realizada de modo semelhante à quantificação realizada nas linhagens celulares. A amostra de tecido mamário normal foi utilizada como amostra de referência e a escolha dos normalizadores ideais para o grupo de amostras analisado foi realizada com auxílio da ferramenta geNorm® (http://medgen.ugent.be/~jvdesomp/genorm/), desenvolvida por VANDESOMPELE et al. (2002) para o *Microsoft Excel*. Esta ferramenta ajuda a selecionar os genes normalizadores mais estáveis para um determinado grupo de amostras. A variação da expressão dos normalizadores nas amostras é realizada comparando-se cada normalizador aos demais normalizadores, sempre aos pares. Assim o programa indica os melhores e os piores normalizadores para o grupo de amostras em questão. Além disso, o programa calcula um fator de normalização para cada amostra individualmente, obtido através da média geométrica da expressão de cada normalizador na mesma amostra. O programa permite ainda avaliar o

impacto gerado pela adição de mais um normalizador ao grupo de amostras, e por conseqüência, escolher o número ideal de normalizadores para o grupo de amostras analisado. Assim, a expressão dos genes *HPRT*, *BCRP*, *β-Actina* e *GAPDH* foi analisada em todas as amotras tumorais e na amostra normal. As seqüências dos *primers* utilizados para amplificar os genes normalizadores estão representados na Tabela 7. O programa geNorm® considerou os quatro genes testados como sendo bons normalizadores para este grupo de amostras. Assim, os quatro genes foram utilizados para a normalização das reações de PCR em Tempo Real. O cálculo da expressão diferencial relativa (razão) entre as amostras tumorais e da amostra normal foi realizado também conforme PFAFFL (2001), utilizando a fórmula descrita na Figura 9.

Tabela 7 - Seqüências dos *primers* dos genes normalizadores utilizados.

Gene Normalizador		*Seqüência*
β-Actina	Fw:	GGCACCCAGCACAATGAAG
	Rev:	CCGATCCACACGGAGTACTTG
HPRT	Fw:	GAACGTCTTGCTCGAGATGTGA
	Rev:	TCCAGCAGGTCAGCAAAGAAT
GAPDH	Fw:	ACCCACTCCTCCACCTTTGA
	Rev:	CTGTTGCTGTAGCCAAATTCGT
BCRP	Fw:	CCTTCGACGTCAATAACAAGGAT
	Rev:	CCTGCGATGGCGTTCAC

4 RESULTADOS E DISCUSSÃO

4.1 TRATAMENTO DAS LINHAGENS CELULARES TUMORAIS E CONSTRUÇÃO DAS BIBLIOTECAS SUBTRATIVAS

Três linhagens celulares tumorais de mama, MCF-7, MDA-MB-435 e MDA-MB-231 foram submetidas ao tratamento com o agente desmetilante 5-aza-2'-deoxicitidina. Estas linhagens celulares apresentam diferentes características fenotípicas, mostrando desde características menos agressivas e mais próximas ao epitélio mamário normal, no caso da linhagem MCF-7, até fenótipos mais agressivos e metastáticos, no caso das linhagens MDA-MB-435 e MDA-MB-231, como mostrado na Tabela 8. A escolha de três linhagens celulares possibilitou a ampliação do painel gênico re-induzido pelo tratamento com o agente desmetilante, assim como a identificação de genes re-induzidos em mais de uma linhagem.

Tabela 8 - Características gerais das linhagens celulares utilizadas.

Linhagem Celular	*Nº ATCC*	*RE*	*RP*	*Potencial Invasivo*[a]	*Potencial Metastático*[b]	*Vimentina*	*Idade da Paciente*
MCF-7	HTB-22	+	+	+	-	-	69
MDA-MB-435	HTB-129	-	-	+++	++++	+	31
MDA-MB-231	HTB-26	-	-	++++	+	+	51

RE: status do receptor de estrógeno; RP: status do receptor de progesterona; Expressão gênica do filamento intermediário vimentina, marcador de fenótipo mesenquimal, detectado em ensaios de Northern Blot [Negativo (-), Positivo (+)].
[a] Potencial invasivo analisado em ensaios de quimioinvasão através de membrana basal reconstituível, MATRIGEL™, em placas do tipo Transwell®.
[b] Potencial metastático verificado através de ensaios de metástase espontânea em camundongo imunodeficiente do tipo *nude*.
Fonte: Modificada de PRICE et al. (1990); THOMPSON et al. (1992); BAE et al. (1993).

A linhagens celulares foram submetidas a 48h de tratamento com 1uM 5-aza-2'-deoxicitidina, seguido de 48h de incubação com novo meio antes da extração de RNA. É importante considerar que as linhagens podem apresentar diferentes sensibilidades ao tratamento. Assim a utilização de diferentes concentrações e diferentes tempos de tratamento estaria influenciando o painel de genes reativados. KARPF et al. (2004) relataram que o prolongamento do tempo de tratamento de células HMEC provocou um aumento de aproximadamente 3 vezes no número de genes diferencialmente expressos. O aumento da dose de 5-aza-dC de 0,5μM para 2μM no tratamento de células RKO também resultou num aumento de 3 vezes no painel de genes re-induzidos. Neste trabalho foram adotados o uso da concentração de 1μM da droga e de tempo de tratamento fixo para reduzir as variações do painel de genes re-induzidos, por exemplo pela alta citotoxicidade da droga. A concentração de 1μM de 5-aza-dC apresentou níveis de toxicidade toleráveis pelas três linhagens celulares submetidas ao tratamento.

A fim de reduzirmos o número de genes re-induzidos sem Ilha de CpG ou com Ilhas de CpG sem metilação na amostra analisada não foi realizado tratamento concomitante com Tricostatina A (TSA), um inibidor de Histona Deacetilase (HDAC). Desta forma, há maior probabilidade de identificação de genes efetivamente regulados pela metilação de seus promotores, e que esta possa futuramente ser utilizada como marcador molecular para o câncer de mama.

As linhagens celulares, MCF-7, MDA-MB-435 e MDA-MB-231, foram tratadas com a droga 5-aza-dC como descrito anteriormente, sendo que uma amostra proveniente da mesma linhagem, a qual não recebeu tratamento, foi utilizada como amostra controle. Para a seleção dos genes re-induzidos pelo tratamento com 5-aza-dC foi utilizada a metodologia de Hibridização de Subtração Rápida (RaSH). Esta metodologia permite a comparação de duas amostras excluindo genes expressos em ambas as amostras, tornando possível a identificação de genes expressos exclusivamente na amostra de interesse. O RaSH apresenta vantagens em relação a outras metodologias de subtração, como o SSH (DIATCHENKO et al. 1996) e o RDA (LISITSYN et al. 1993), principalmente por apresentar apenas um passo de amplificação por PCR, sendo possível até a não realização do mesmo. As estratégias que utilizam *arrays* como plataforma para a análise da expressão gênica, tanto em membrana quanto em lâminas de vidro, podem avaliar um número pré-determinado de genes, limitando assim o painel analisado. O RaSH, ao contrário, pode analisar todo painel transcrito nas amostras.

FUNDAÇÃO ANTONIO PRUDENTE

Para cada linhagem celular foram construídas duas bibliotecas subtrativas através da técnica de RaSH: uma delas utilizando cDNA da cultura tratada como *tester* e a outra utilizando cDNA da cultura não tratada como *tester*. A primeira contém os genes supostamente regulados por metilação reativados após tratamento com 5-Aza-dC, e a segunda foi construída com o intuito de servir de controle da primeira, pois, caso a subtração seja eficiente, um mesmo gene não pode estar presente nas duas bibliotecas. Assim, ao todo, foram construídas 6 bibliotecas subtrativas (Tabela 9).

Para a construção de cada biblioteca subtrativa, 25ug de RNA total foram utilizados para a síntese da 1ª e da 2ª fita de cDNA, utilizando o kit *SuperScript II Double-Stranded cDNA Synthesis* (Invitrogen). A qualidade de síntese foi verificada através de amplificação por PCR com *primers* específicos para um gene constitutivo (*GAPDH*) (Figura 10A). Em seguida, o cDNA sintetizado foi digerido com a enzima *Sau*3A1 e a ocorrência de digestão foi confirmada pela diferença de migração entre a amostra digerida e não digerida, por eletroforese em gel de agarose (Figura 10B). Aos fragmentos digeridos foi ligado um adaptador e a ligação deste foi confirmada através de amplificação por PCR com *primer* específico para o adaptador (Figura 10C). Para a obtenção de massa de DNA suficiente para a subtração foram realizadas 20 reações de PCR com *primer* específico para o adaptador. O produto dessas amplificações foi usado como amostra *driver*. Uma alíquota de 10µg do produto amplificado foi digerida com a enzima *Xho*I, e foi utilizada como amostra *tester*. Para confirmar a eficiência de

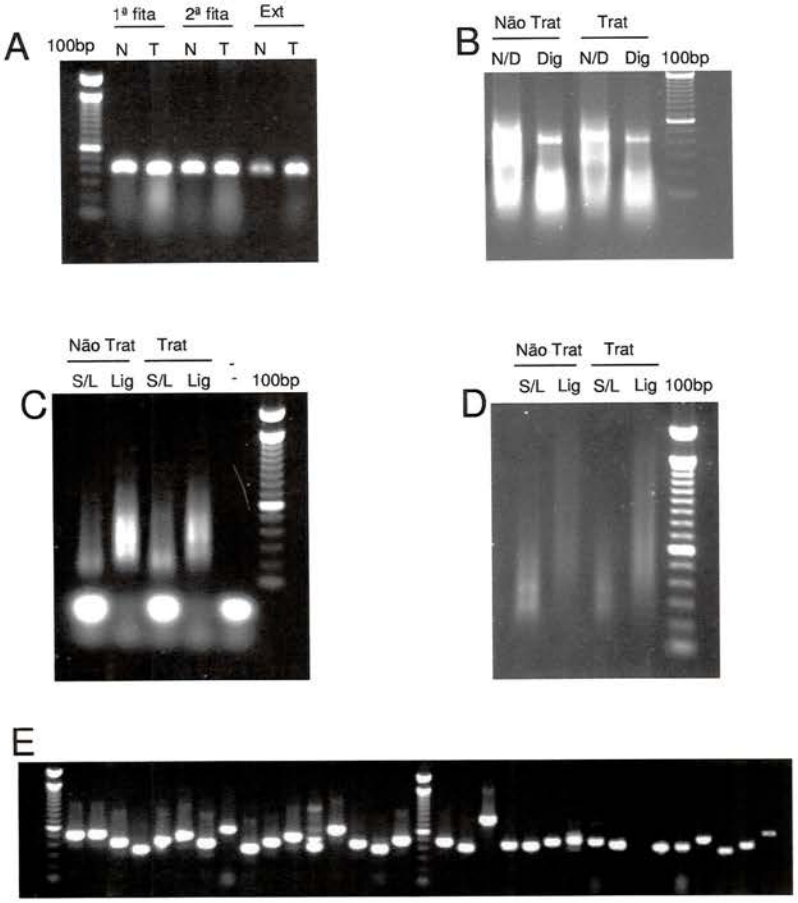
digestão com a enzima *Xho*I, uma alíquota da digestão foi incubada na presença da enzima T4 Ligase (New England Biolabs). Após a ligação, os fragmentos foram analisados através de eletroforese em gel de agarose para avaliar a formação de concatâmeros dos fragmentos digeridos e desta forma comprovar a digestão com *Xho*I (Figura 10D). Para a subtração, 100ng de cDNA da amostra *tester* foram misturados a 5ug de cDNA da amostra *driver* em Tampão de Hibridação, fervidos durante 5 minutos e incubados a 42ºC por 48 horas. Esta mistura foi então ligada no vetor pZErO - 1 (Invitrogen), transformada em bactérias competentes e plaqueada em meio L.B. Low Salt contendo Zeocina® (Invitrogen). Os clones foram isolados e seus insertos foram amplificados por PCR. Os fragmentos amplificados foram analisados através de eletroforese em gel de 1% agarose (Figura 10E) e submetidos a reações de seqüenciamento. As seqüências obtidas foram submetidas, via Internet, à página do Centro de Terapia Celular da Faculdade de Medicina de Ribeirão Preto – USP (http://guarani.fmrp.usp.br/cgi-bin/geap/index.cgi).

O protocolo original da metodologia de RaSH, descrito por JIANG et al. (2000), utiliza 100ng de amostra *tester* e 3ug de amostra *driver* para a subtração, proporcionando um excesso de 30 vezes mais *driver*. Neste trabalho foram utilizados 5ug de amostra *driver*, impondo assim maior estringência à subtração. Cada molécula *tester* teria, portanto, 50 vezes mais chances de hibridizar-se com moléculas *driver* do que com outras moléculas *tester*. Com isso, pretendíamos diminuir a geração de falso-positivos e obter maior certeza de que os genes identificados pelo RaSH estariam presentes exclusivamente na amostra *tester*.

Tabela 9 - Bibliotecas subtrativas construídas para as linhagens MCF-7, MDA-MB-435 e MDA-MB-231 e seqüências geradas.

MCF-7				MDA-MB-435				MDA-MB-231			
MCF-7-5Aza-Tester		MCF-7-Mock-Tester		435-5Aza-Tester		435-Mock-Tester		231-5Aza-Tester		231-Mock-Tester	
Tester (100ng)	Tratada	Tester (100ng)	Não tratada	Tester (100ng)	Tratada	Tester (100ng)	Não tratada	Tester (100ng)	Tratada	Tester (100ng)	Não tratada
Driver (5,0ug)	Não tratada	Driver (5,0ug)	Tratada	Driver (5,0ug)	Não tratada	Driver (5,0ug)	Tratada	Driver (5,0ug)	Não tratada	Driver (5,0ug)	Tratada
Seqüências		**Seqüências**		**Seqüências**		**Seqüências**		**Seqüências**		**Seqüências**	
Baixa Qualid	47	Baixa Qualid	85	Baixa Qualid	1	Baixa Qualid	2	Baixa Qualid	58	Baixa Qualid	136
Vetor	22	Vetor	30	Vetor	17	Vetor	1	Vetor	16	Vetor	143
Rep	1	Rep	1	Rep	4	Rep	5	Rep	21	Rep	4
Cont Mit	1	Cont Mit	1	Cont Mit	10	Cont Mit	2	Cont Mit	23	Cont Mit	24
Cont Lev	0	Cont Bac	0	Cont Lev	3	Cont Bac	1	Cont Lev	1	Cont Rib	19
RefSeq	269 (70)	RefSeq	77 (55)	RefSeq	170 (121)	RefSeq	70 (54)	RefSeq	596 (368)	RefSeq	213 (156)
UniGene	10 (8)	UniGene	9(9)	UniGene	19 (13)	UniGene	8(7)	UniGene	22 (22)	UniGene	13 (12)
EST	11(5)	EST	9(5)	EST	13	EST	26(11)	EST	18 (16)	EST	18 (15)
No match	78	No match	5	No match	5	No match	3	No match	7	No match	6

00-5AZA-*Tester*, 01-5AZA-*Tester* e 02-5AZA-*Tester* representam bibliotecas construídas com cDNA amostra tratada utilizado como *tester* para MCF-7, MDA-MB-435 e MDA-MB-231 respectivamente; 00-Mock-*Tester*, 01-Mock-*Tester* e 02-Mock-*Tester* representam bibliotecas construídas com cDNA amostra não tratada utilizado como *tester* para MCF-7, MDA-MB-435 e MDA-MB-231 respectivamente; *Baixa Qualid*, baixa qualidade; *Cont Mit*, contaminante com DNA mitocondrial; *Cont Lev*, contaminante com DNA de levedura; *Cont Bac*, contaminante com DNA bacteriano. *RefSeq*, alinhamento com seqüência referência, *UniGene*, alinhamento com *cluster* UniGene, *EST*, alinhamento com seqüências depositadas na forma de EST. O número representado entre parênteses corresponde ao número de transcritos diferentes.



Legenda: (A) Análise da qualidade da síntese de cDNA através de reações de RT-PCR com *primers* específicos para gene constitutivo GAPDH, utilizando como template cDNA das linhagens celulares não tratada (N) e tratada (T) após a síntese da primeira fita (1ª), da segunda fita (2ª) e após a extração com fenol clorofórmio (Ext). (B) Verificação da digestão com a enzima de restrição *Sau3A*1 através da comparação entre amostra digerida (Dig) com alíquota separada antes da digestão (N/D). (C) Verificação da ligação dos adaptadores através da amplificação com primer específico para os adaptadores (XDPN-18) e a subseqüente comparação com alíquota separada antes da adição de enzima DNA Ligase (S/L). (D) Verificação da digestão da amostra tester com *Xho*I através da adição de enzima DNA Ligase e subseqüente comparação com alíquota que não recebeu ligase (S/L). (E) Avaliação da qualidade dos produtos amplificados a partir dos insertos dos clones obtidos da transformação. Todos os controles foram visualizados em gel de 1% agarose.

Figura 10 - Construção das bibliotecas subtrativas.

4.2 ANÁLISE DAS SEQÜÊNCIAS GERADAS

Ao contrário do protocolo original descrito por JIANG et al. (2000), o qual utilizou o Northern Reverso para o *screening* dos clones gerados pelo RaSH, o presente trabalho utilizou um sistema *in silico*, formado por um conjunto de programas computacionais. O uso de um sistema *in silico* para a análise dos clones gerados pela técnica de RaSH, ao invés do *Northern* Reverso, possui muitas vantagens. A análise *in silico* possibilita a identificação de seqüências repetitivas e de seqüências contaminantes de outros organismos, além de possibilitar o alinhamento do produto amplificado com seqüências depositadas em bancos de dados públicos, revelando sua anotação e sua provável função. Ao ser depositada, cada seqüência foi analisada, de forma seqüencial, por diversos programas computacionais. Inicialmente a qualidade da seqüência é analisada através do programa PHREDPHRAP® (*Washington University*), excluindo seqüências consideradas de má qualidade. As seqüências consideradas de boa qualidade foram então alinhadas a bancos de dados internos contendo seqüências repetitivas, vetores e contaminantes de outros organismos. Em caso de similaridade as seqüências foram excluídas. As seqüências que não apresentaram similaridade com nenhuma seqüência do banco de dados interno foram então alinhadas a bancos de dados públicos de seqüências completas de mRNA (*RefSeq*) através do BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). Da mesma forma, seqüências sem similaridade com seqüências completas de mRNA foram alinhadas contra os

bancos de dados Unigene (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=unigene) e de ESTs (http://www.ncbi.nlm.nih.gov/dbEST/index.html). Além da caracterização dos transcritos, este conjunto de programas permite ainda comparar diferentes experimentos e identificar com maior facilidade os transcritos comuns a dois experimentos diferentes. Assim, foi possível a comparação entre um experimento e seu controle, facilitando a identificação de falso-positivos. Outra vantagem desta ferramenta é a comparação entre diferentes linhagens celulares. Assim, foi possível a identificação e seleção de genes comuns a duas ou mais linhagens (Figura 11).

No total 2367 seqüências foram geradas, sendo 1449 provenientes das bibliotecas subtrativas das linhagens celulares tratadas com 5-aza-dC e 918 provenientes das bibliotecas subtrativas de linhagens celulares não tratadas. Para a linhagem MCF-7 foram geradas 663 seqüências, sendo 439 do experimento MCF-7-5AZA-*tester* e 224 do experimento MCF-7-Mock-*tester*. As seqüências foram submetidas, via Internet, à *webpage* do Centro de Terapia Celular da Faculdade de Medicina de Ribeirão Preto – USP. No experimento MCF-7-5AZA-*tester*, 71 seqüências foram recusadas por apresentarem baixa qualidade de seqüência, algum tipo de contaminação (seqüências mitocondriais, bacterianas, de levedura) ou por apresentarem somente seqüências do vetor (para maiores informações ver Tabela 9). Das 368 seqüências restantes, quando alinhadas a seqüências depositadas em bancos de dados públicos, 269 seqüências apresentaram similaridade com seqüências completas de mRNA, representativas de 70 transcritos

diferentes, sendo 61 deles exclusivos desta biblioteca (MCF-7-5AZA-*tester*). Outras 10 seqüências apresentaram similaridade com *clusters* de ESTs do Unigene e 11 com *ESTs* humanas não catalogadas no Unigene. Para o experimento MCF-7-Mock-*tester*, 120 seqüências foram excluídas também por apresentarem baixa qualidade, seqüências contaminantes ou seqüências do vetor. Outras 77 seqüências mostraram similaridade a seqüências completas de mRNA, representando 55 transcritos diferentes, outras 9 mostraram similaridade com *clusters* de Unigene e 9 com ESTs humanas.

Para a linhagem MDA-MB-435, 242 seqüências foram geradas para o experimento 435-5AZA-*tester*, sendo 35 recusadas. 170 seqüências mostraram similaridade com RefSeqs, representativas de 121 transcritos diferentes, sendo 101 destes exclusivos da biblioteca subtrativa desta linhagem. Outras 19 seqüências apresentaram similaridade com *clusters* Unigene e 13 apresentaram similaridade com ESTs. Para o experimento 435-Mock-*tester*, das 118 seqüências geradas, 11 seqüências foram recusadas, e 70 apresentaram similaridade com seqüências completas de mRNA, representativas de 54 transcritos diferentes. Já 8 seqüências apresentaram similaridade com *clusters* Unigene e 26 com ESTs.

Para a linhagem MDA-MB-231, 1344 seqüências foram geradas no total, sendo 768 para o experimento 231-5AZA-*tester*. Destas, 58 foram recusadas por apresentarem baixa qualidade e 61 foram recusadas por apresentarem algum tipo de contaminação (com seqüências de levedura, mitocondriais ou ribossomais) ou apresentarem apenas seqüências do vetor.

596 seqüências apresentaram similaridade com seqüências completas de mRNA, representando 368 transcritos diferentes, sendo que destes 349 são transcritos exclusivos desta biblioteca subtrativa. 22 seqüências apresentaram similaridade com *clusters* Unigene e 18 com ESTs. Das 576 seqüências geradas no experimento 231-Mock-*tester, 213* apresentaram similaridade com seqüências completas de mRNA (156 transcritos diferentes), 13 apresentaram similaridade com *clusters* Unigene e 18 com ESTs.

FUNDAÇÃO ANTONIO

A

RaSH - LICR - Anamaria Camargo

Refused due to Low Quality	47
Refused due to Vector Presence	22
Refused due to Repeat Presence	1
Contaminant Mitochondrial	1
Contaminant Bacterium	0
Contaminant Yeast	0
Contaminant Ribosomal	0
Reference Sequences	269
Unigene	10
EST human	11
Non Redundant Human	0
Non Redundant non Human	0
No Matches	78
Total	439

Download | Reset

B

Vector(s)	**Id**	**Coordinate**
	pCR1-TOPO_Cloning_Vector	101..125
	pCR1-TOPO_Cloning_Vector	126..164

Nucleotides

Low Quality
Vector(s)
Primer(s)
Repeat(s)

```
ATGGACCCCNNCCCCAAAGNNNNNCCTGGATCACCAGCTATTTAGGTGG
AACTATTCAATACTCAAGATATGCATCAAGCTTGTCCCCGAGCTCGGAGC
CACTAGTAACGGCCGCCAGTGTGCTGGAATTCTGCAGATATCCATCACAC
TGGCGGCCGCTCGAGAGATCCCCAGAAGGACAAGCGCTTCTCGGGCACCG
TCAGGCTTAAGTCCACTCCCCGCCCTAAGTTCTCTGTGTGTGTCCTGGGG
GACCAGCAGCACTGTGACGAGGCTAAGGCCGTGGATATCCCCCACATGGA
CATCGAGGCGCTGAAAAAACTCAACAAGAATAAAAAACTGGTCAAGAAGC
TGGCCAAGAAGTATGATGCGTTTTTGGCCTCAGAGTCTCTGATCTCTCGA
GAGATCTGTGGTCCTTTGGGACTCATCAAAGCCTTGGTTTAGCATTTTGT
CAGTTTTATCTTCAGAAATTCTCTGCGATTAAGAAGATAATTTATTAAAG
GTGGTCCTTCCTACCTCTGTGGTGTGTGTCGCGCACACAGCTTAGAAGTG
CTATAAAAAAGGAAAGAGCTCCAAATTGAATCACCTTTATAATTTACCCA
TTTCTA
```

Accession NM_007104

External Links Links

SAGE Tag GGCAAGCCCC

H2G Libraries

```
CTC_UmbilicalCord2_normal_cmc_GC52
SAGE_Prostate_carcinoma_B_LN-1
SAGE_Brain_medulloblastoma_B_DL5
SAGE_Retina_Peripheral_normal_B_1
SAGE_Brain_astrocytoma_grade_III_B_407
```

View Expression

BesTags NM_007104.3

Description Homo sapiens ribosomal protein L10a (RPL10A), mRNA

Query Start 66

C

RaSH - LICR - Anamaria Camargo

Mitochondrial		RH00AZ 2	RH00CT 1
J01415	Human mitochondrion, complete genome	1	4
		1	4

Yeast Est		RH00AZ 2	RH00CT 1
T37322	T37322 EST102420 Saccharomyces cerevisiae cDNA 5' end. 1/95		1
		0	1

RefSeq		RH00AZ 2	RH00CT 1
XM_171544	Homo sapiens similar to sgCP9304 [Anopheles gambiae str. PEST] (LOC253141), mRNA	2	
NM_024719	Homo sapiens hypothetical protein FLJ22474 (FLJ22474), mRNA	1	
XM_117287	Homo sapiens LOC200830 (LOC200830), mRNA	1	
NM_005558	Homo sapiens ladinin 1 (LAD1), mRNA	1	
NM_002952	Homo sapiens ribosomal protein S2 (RPS2), mRNA		1
XM_044434	Homo sapiens KIAA1458 protein (KIAA1458), mRNA	1	
NM_001569	Homo sapiens interleukin-1 receptor-associated kinase 1 (IRAK1), mRNA	2	
NM_014153	Homo sapiens HSPC055 protein (HSPC055), mRNA		1
NM_001356	Homo sapiens DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 3 (DDX3), transcript variant 2, mRNA	1	
NM_003295	Homo sapiens tumor protein, translationally-controlled 1 (TPT1), mRNA		1
NM_133376	Homo sapiens integrin, beta 1 (fibronectin receptor, beta polypeptide, antigen CD29 includes MDF2, MSK12) (ITGB1), transcript variant 1E, mRNA		1
XM_030865	Homo sapiens M-phase phosphoprotein 10 (U3 small nucleolar ribonucleoprotein) (MPHOSPH10), mRNA	1	
NM_000999	Homo sapiens ribosomal protein L38 (RPL38), mRNA	3	
NM_003900	Homo sapiens sequestosome 1 (SQSTM1), mRNA	4	
NM_031992	Homo sapiens Williams-Beuren syndrome chromosome region 1 (WBSCR1), transcript variant 2, mRNA	2	

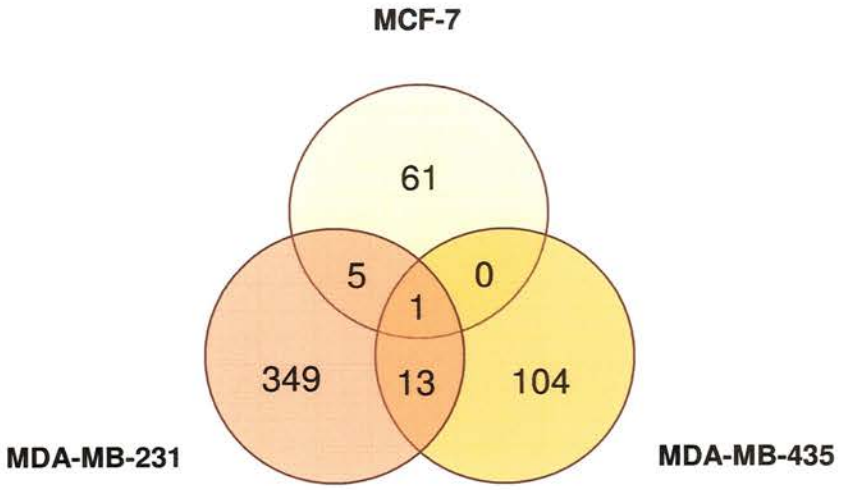
Legenda: (A) Análise qualitativa geral das seqüências submetidas. (B) Análise detalhada de uma única seqüência, visualizando informações como a porção da seqüência com baixa qualidade (1), vetor (2), número de acesso (3), SAGE tag virtual (4), anotação (5) e etc., apontadas pelas setas. (C) Comparação entre dois experimentos diferentes, possibilitando a visualização de seqüências comuns a diferentes experimentos e a quantidade seqüências por transcrito.

Figura 11 - *Webpage* utilizada para análise *in silico*.

Como previamente descrito, a página possibilita também a comparação entre diferentes experimentos, visando identificar genes comuns a dois ou mais experimentos (Figura 11C). Assim, comparando-se um experimento com o seu reverso, por exemplo, MCF-7-5Aza-*tester* com MCF-7-Mock-*tester*, foi possível verificar a presença de falsos positivos, ou seja, seqüências não subtraídas da amostra de interesse. Para a linhagem celular MCF-7, do total de 122 genes identificados, apenas 3 estavam presentes tanto no experimento MCF-7-5Aza-*tester* quanto no MCF-7-Mock-*tester*, confirmando a eficiência da subtração. A linhagem celular MDA-MB-435 também apresentou apenas 3 transcritos, dos 172 identificados,

comuns aos dois experimentos. Para a linhagem celular MDA-MB-231, 44, dos 524 genes identificados, estavam presentes tanto no experimento 231-5AZA-*tester* como em seu reverso, 231-Mock-*tester*.

Utilizando este mesmo recurso da *webpage,* é possível a identificação dos transcritos comuns entre as bibliotecas subtrativas das diferentes linhagens. Assim, foi possível a identificação de 1 transcrito reativado pelo tratamento com 5-aza-dC nas três linhagens celulares analisadas, 5 comuns às linhagens MCF-7 e MDA-MB-231 e 13 comuns às linhagens MDA-MB-435 e MDA-MB-231. Não foram identificados genes comuns entre as biblioteca subtrativas das linhagens MCF-7 e MDA-MB-435. Estes dados estão esquematizados na Figura 12.



Legenda: O esquema acima representa o número de genes presentes no experimento que utiliza cDNA da amostra tratada como tester na hibridização de subtração. As intersecções entre os círculos representam seqüências comuns a duas ou mais linhagens celulares.

Figura 12 - Análise global das seqüências geradas pelos experimentos 5AZA-*tester*.

No total foram identificados 533 genes presentes exclusivamente nas bibliotecas subtrativas que utilizaram cDNA da linhagem tratada com 5-aza-dC como *tester*. Como em outras estratégias que utilizam bibliotecas para a identificação de genes diferencialmente expressos em larga escala, é possível que este número aumente à medida que mais clones são seqüenciados. Como exemplo, 61 genes re-induzidos pelo tratamento foram identificados na biblioteca MCF-7-AZA-*tester*, para a qual foram seqüenciados 439 clones. Por outro lado, para a biblioteca 231-AZA-*tester,* os 768 clones seqüenciados permitiram a identificação de 349 genes re-induzidos pelo tratamento. O seqüenciamento de um número muito grande

de colônias poderia cobrir o painel real de genes reativados, mas elevaria o custo e o tempo de realização do trabalho.

Assim, a *webpage* do Centro de Terapia Celular da Faculdade de Medicina de Ribeirão Preto – USP possibilitou a identificação de um grande número de genes re-induzidos pelo tratamento com 5-aza-dC e, possivelmente, regulados por metilação.

4.3 SELEÇÃO DE CANDIDATOS PARA VALIDAÇÃO EXPERIMENTAL

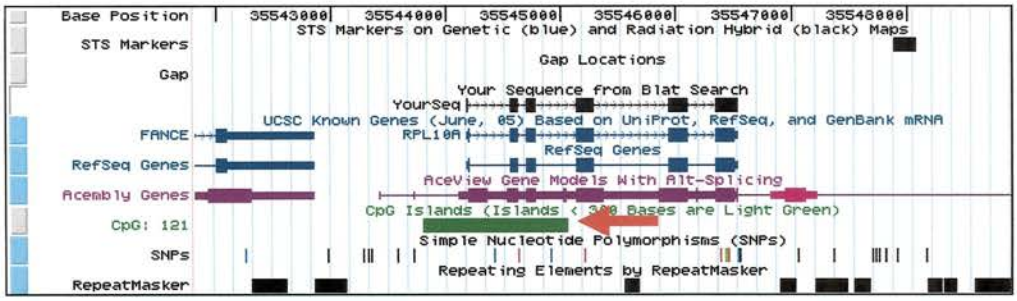
A técnica de RaSH foi utilizada em cerca de uma dezena de trabalhos desde seu desenvolvimento. Assim, ainda não há muita informação disponível na literatura sobre a eficiência desta metodologia. Mesmo realizando-se a subtração em condições altamente estringentes e após a eliminação de genes presentes nas bibliotecas controles através da análise *in silico*, os genes identificados pela técnica de RaSH ainda podem corresponder a falso-positivos. De acordo com os autores que descreveram a técnica (JIANG et al. 2000), apenas 23 (~33%) dos 70 clones que foram gerados pelo RaSH e analisados por *Northern* Reverso apresentaram diferença de expressão superior a 1,8 vezes.

Assim, torna-se clara a necessidade de validação experimental dos resultados obtidos pelo RaSH por outra metodologia. Neste trabalho o PCR em Tempo Real foi escolhido para a validação experimental de genes diferencialmente expressos selecionados através da técnica de RaSH.

A partir do número de acesso de cada transcrito foi obtida sua seqüência completa (*FASTA*), utilizando o *Entrez Nucleotide Databases*, fornecido pelo *National Center for Biotechnology Information* (www.ncbi.nlm.nih.gov). Esta seqüência foi utilizada para a verificação da presença de Ilhas de CpG na região promotora dos genes identificados, através do programa BLAT (http://www.genome.ucsc.edu/cgi-bin/hgBlat) (Figura 13). Cerca de 85,6% dos genes presentes nas bibliotecas subtrativas construídas com cDNA proveniente da linhagem celular tratada usado como *tester*, escolhidos randomicamente e analisados no programa BLAT, possuem Ilha de CpG em sua região promotora. O número obtido é superior ao estimado, dado que segundo a literatura, aproximadamente 60% dos genes no genoma humano possuem Ilha de CpG em sua região promotora (ATTWOOD et al. 2002; WIDSCHWENDTER e JONES, 2002). Assim, este dado mostra que o tratamento com agente desmetilante foi eficiente no enriquecimento de nossa amostra com genes possivelmente regulados por metilação (Figura 14). Levando-se em consideração o fato de que uma fração dos genes presentes nas bibliotecas de subtração não possui Ilhas de CpG, fica clara a importância da inspeção manual dos candidatos, incluindo a presença da Ilha de CpG na região promotora do gene como critério de seleção de candidatos para validação.

A reativação da expressão de genes candidatos que não apresentam Ilha de CpG em sua região promotora pode se dever a diferentes fatores: (i) pela desmetilação e subseqüente reativação da expressão gênica de fatores de transcrição; (ii) pela desmetilação e reativação de outros genes situados

"*upstream*" na cascata metabólica ativando a expressão gênica do gene que não possui Ilha; (iii) alterações inespecíficas devidas a processos biológicos decorrentes da ativação de uma via de transdução de sinais através da desmetilação de um gene ou ainda (iv) a desmetilação de toda a região cromossômica onde o gene candidato está situado, a qual encontrava-se silenciada por um evento epigenético denominado *Long-Range Epigenetic Silencing* (LRES).



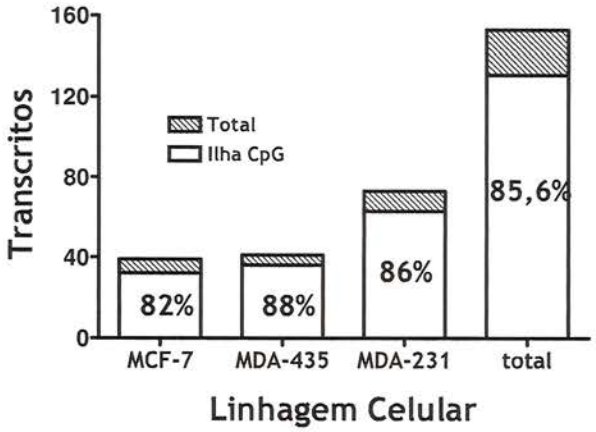
Legenda: Interface gráfica fornecida pelo programa BLAT (http://www.genome.ucsc.edu/cgi-bin/hgBlat), mostrando a presença de Ilha de CpG (apontada pela seta em vermelho) na região promotora do gene e, em alguns casos, invadindo os primeiros exons.

Figura 13 - Verificação da presença de Ilha de CpG nos genes identificados pela técnica de RaSH.

Além dos mecanismos acima descritos, KARPF et al. (1999) relataram alterações do padrão de expressão de genes situados na via de sinalização de Interferon após o tratamento com agente desmetilante. Apesar de não serem individualmente regulados por metilação em seus promotores, os genes desta via são afetados indiretamente pelo tratamento com agente desmetilante. LIANG et al. (2002) também relataram ativação de genes situados na via de transdução de sinal de interferon após tratamento de

células de câncer de bexiga com 5-aza-dC. Assim, uma busca por genes situados nessa via de transdução foi realizada em nossas bibliotecas subtrativas com o intuito de eliminá-los de nossas análises. Os genes que codificam Interferon-α, -β e -γ, além de seus receptores, JAK-1 e -2, Tyk2, STAT-1 e -2 não estão presentes nas bibliotecas subtrativas geradas a partir do tratamento das linhagens tumorais de mama com 5-aza-dC.

Para a validação experimental, foram selecionados 41 genes candidatos, sendo 8 genes candidatos exclusivos da linhagem MCF-7, 7 da linhagem MDA-MB-435, 10 da linhagem MDA-MB-231, além dos 5 genes em comum entre as linhagens MCF-7 e MDA-MB-231, 10 em comum entre as linhagens MDA-MB-435 e MDA-MB-231 e do único gene re-induzido pelo tratamento nas três linhagens celulares. Os genes selecionados e suas respectivas anotações estão representados na Tabela 10.



Legenda: Utilizando o programa BLAT, os genes re-induzido pelo tratamento foram analisados quanto à presença de Ilhas de CpG em sua região promotora. Barras hachuradas representam o total de genes analisados; barras brancas representam o número de genes com Ilha de CpG em sua região promotora.

Figura 14 - Presença de Ilhas de CpG na região promotora nos genes reativados por 5-aza-dC.

4.4 UTILIZAÇÃO DE *NORTHERN* REVERSO PARA SELEÇÃO EM LARGA ESCALA DE GENES CANDIDATOS PARA VALIDAÇÃO EXPERIMENTAL

O protocolo original da técnica de RaSH utiliza o *Northern* Reverso como método de *screening* dos genes diferencialmente expressos identificados pela subtração. A utilização do *Northern* Reverso possibilita a validação experimental em larga escala, o que não acontece com o uso da *webpage* utilizada neste trabalho, onde cada candidato deve ser validado individualmente, demandando maior tempo e trabalho manual. Com o intuito de realizar a validação experimental em larga escala dos genes identificados na linhagem MDA-MB-231 pela metodologia de RaSH, foi realizado também um experimento de *Northern* Reverso. Para tal experimento, todos os produtos amplificados dos clones gerados pela construção da biblioteca subtrativa da linhagem MDA-MB-231 foram transferidos para membrana de nylon e hibridizados com RNA marcado radioativamente da respectiva linhagem celular tratada e não tratada com 5-Aza-dC. O gene *GAPDH* foi utilizado como normalizador deste experimento. Dos 368 transcritos analisados (596 clones fixados), 42 (cerca de 11,4%) apresentaram expressão diferencial maior ou igual a 1,8 vezes, valor de *cutoff* utilizado previamente por KANG et al. (1998). Quando confrontados à análise realizada *in silico* pela análise direta das seqüências, 2 destes 42 genes correspondiam a contaminações com DNA mitocondrial, 2 correspondiam a seqüências repetitivas, 3 correspondiam a seqüências de baixa qualidade e

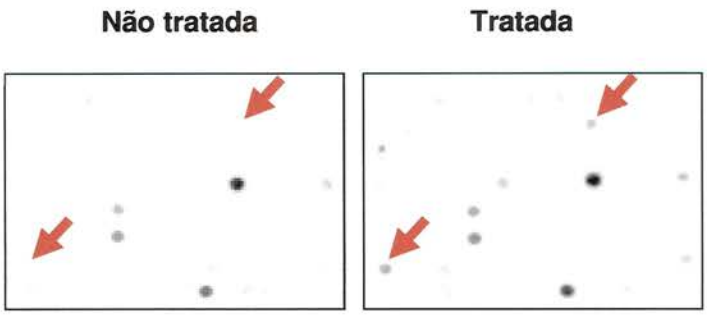
1 correspondia a seqüência do vetor, mostrando a praticidade e real vantagem do uso da análise *in silico* em relação ao Northern Reverso. Essas seqüências foram então excluídas de nossas análises. Os quatro candidatos que apresentaram expressão superior a 3 vezes foram selecionados para validação por PCR em Tempo Real (BM 201, 202, 203 e 204). Um exemplo da hibridização está mostrado na Figura 15.

Tabela 10 - Genes selecionados para validação experimental.

Candidatos			Repres.		
Nome	Descrição	Anotação GeneOntology1	MCF	435	231
BM 001	*Homo sapiens sequestosome 1 - SQSTM1*	Metabolismo de proteínas, Catabolismo	4		
BM 002	*Homo sapiens keratin 19 - KRT19*	Sem anotação	3		
BM 004	*Homo sapiens DKFZP564O0463 protein*	Sem anotação	5		
BM 005	*Homo sapiens T-cell activation protein - PGR1*	Sem anotação	4		
BM 007	*Homo sapiens guanine nucleotide binding protein, gamma 5 - GNG5*	Comunicação Celular	5		
BM 008	*Homo sapiens cytochrome c oxidase subunit Vb - COX5B*	Processos fisiológicos celulares	1		
BM 009	*Homo sapiens transducer of ERBB2, 1 - TOB1*	Regulação de processos fisiológicos	1		
BM 010	*Homo sapiens cisplatin resistance related protein CRR9p - CRR9*	Sem anotação	1		
BM 101	*Homo sapiens lactate dehydrogenase A – LDHA*	Metabolismo		3	
BM 102	*Homo sapiens hypothetical protein MGC8721*	Sem anotação		4	
BM 103	*Homo sapiens ring finger protein 11 - RNF11*	Metabolismo		3	
BM 106	*Homo sapiens retinoblastoma binding protein 6 - RBBP6*	Regulação de processos fisiológicos		1	
BM 107	*Homo sapiens mitogen-activated protein kinase 9 - MAPK9*	Metabolismo, resposta a estímulo		1	
BM 108	*Homo sapiens CDC5 cell division cycle 5-like (S. pombe) - CDC5L*	Sem anotação		1	
BM 109	*Homo sapiens inosine monophosphate dehydrogenase 2 - IMPDH2*	Metabolismo, processos fisiológicos		1	
BM 201	*Homo sapiens connective tissue growth factor - CTGF*	Regulação de proliferação celular			4
BM 202	*Homo sapiens SMT3 suppressor of mif two 3 homolog 2 - SUMO2*	Metabolismo, processos fisiológicos			1
BM 203	*Homo sapiens cathepsin B - CTSB*	Metabolismo, processos fisiológicos			20
BM 204	*Homo sapiens pituitary tumor-transf. 1 interac. protein - PTTG1IP*	Processos fisiológicos celulares			1
BM 205	*Homo sapiens intercellular adhesion molecule 1 - ICAM1*	Comunicação celular			1
BM 206	*Homo sapiens tumor differentially expressed 1 - TDE1*	Sem anotação			1
BM 207	*Homo sapiens suppression of tumorigenicity 5 - ST5*	Sem anotação			1
BM 208	*Homo sapiens excision repair cross-complementing - ERCC5*	Sem anotação			1
BM 209	*Homo sapiens PERP, TP53 apoptosis effector - PERP*	Sem anotação			3
BM 210	*Homo sapiens thyroid hormone receptor interactor 4 - TRIP4*	Regulação de processos fisiológicos			1
BM 301	*Homo sapiens cell division cycle 27 - CDC27*	Regulação do ciclo celular	3	3	1
BM 401	*Homo sapiens membrane-associated ring finger 5 - MARCH5*	Processos fisiológicos celulares		1	1
BM 402	*Homo sapiens aspartyl-tRNA synthetase (DARS), mRNA*	Processos fisiológicos celulares		1	1
BM 403	*Homo sapiens serine (or cysteine) proteinase inhibitor - SERPINE2*	Sem anotação		5	2
BM 404	*Homo sapiens fibronectin 1 (FN1), transcript variant 1, mRNA*	Comunicação Celular		2	3
BM 405	*Homo sapiens ATPase, membr. associated protein M8-9 - ATP6M8-9*	Sem anotação		1	1
BM 406	*Homo sapiens pitrilysin metalloproteinase 1 - PITRM1*	Regulação de processos fisiológicos		1	1
BM 407	*Homo sapiens MO25 protein (MO25), mRNA*	Sema anotação		2	1
BM 104	*Homo sapiens tumor rejection antigen (gp96) 1 - TRA1*	metabolismo, resposta a estímulo		4	1
BM 105	*Homo sapiens heat shock 90kDa protein 1, alpha – HSPCA*	metabolismo		3	2
BM 110	*Homo sapiens synaptobrevin-like 1 - SYBL1*	metabolismo, resposta a estímulo		2	1
BM 501	*Homo sapiens calmodulin 1 (phosphorylase kinase, delta) - CALM1*	Comunicação Celular	1		1
BM 502	*Homo sapiens basic leucine zipper and W2 domains 1 - BZW1*	Sem anotação	1		1
BM 503	*Homo sapiens amyloid beta precursor prot. binding prot. - APPBP1*	Metabolismo, comunicação celular	1		1
BM 003	*Homo sapiens thyroid hormone receptor interactor 6 - TRIP6*	Processos Fisiológicos Celulares	5		1
BM 006	*Homo sapiens superoxide dismutase 2, mitochondrial - SOD2*	Resposta a Estímulo, Metabolismo	4		2

[1]GeneOntology (http://www.fatigo.org/)

A tabela mostra o nome do candidato, sua anotação de acordo com *PubMed* (Descrição), anotação de acordo com o *Gene Ontology* (Anotação GeneOntolgy) e o número de seqüências presentes na biblioteca (Repres).



Legenda: Amostra representativa de clones originados da biblioteca 231-5AZA-tester fixados em membranas de Nylon e hibridizados com cDNA marcado radioativamente ([32P]dCTP) proveniente da linhagem MDA-MB-231 tratada e não tratada com 5-aza-dC. Após o tratamento, 42 genes apresentaram reativação da expressão maior ou igual a 1,8 vezes. Setas vermelhas indicam dois clones mostrando diferença de expressão maior que 3 vezes como conseqüência do tratamento. O gene *GAPDH* foi utilizado com normalizador.

Figura 15 - Análise em larga escala por *Northern* Reverso.

4.5 VALIDAÇÃO EXPERIMENTAL ATRAVÉS DE PCR EM TEMPO REAL

A técnica de PCR em Tempo Real é uma ferramenta muito sensível e possibilita análise a quantitativa da expressão gênica. Assim, neste trabalho foi realizada a quantificação relativa, onde a re-indução do gene pelo tratamento foi calculada utilizando a linhagem celular não tratada com 5-aza-dC como referência, normalizada pela expressão de um controle endógeno, baseado na fórmula proposta por PFAFFL (2001), descrita em Material e Métodos.

Os *primers* foram desenhados com o auxílio do software *PrimerExpress* versão 3.0 (Applied BioSystems) para cada candidato, como

descrito em Material e Métodos. O método de detecção escolhido foi o sistema *SYBR® Green* (Applied BioSystems) o qual baseia-se na fluorescência emitida pelo corante ao intercalar-se em DNA dupla fita, à medida em que ocorre a amplificação do DNA. Devido à sua inespecificidade, as reações utilizando *SYBR® Green* foram devidamente padronizadas para cada gene candidato, a fim de se reduzir a ocorrência de amplificação inespecífica e a formação de dímeros de *primers*. Com o propósito de avaliar a eficiência de amplificação dos produtos pelos *primers* desenhados e escolher a melhor concentração de cDNA *template*, a expressão dos candidatos foi avaliada primeiramente em linhagem celular de mama normal imortalizada (HB4A). Assim, foram escolhidas as concentrações ideais de *primers* e de cDNA nas quais obtivéssemos o maior valor de fluorescência e maior especificidade de amplificação. A análise da curva de dissociação do produto amplificado e a visualização do mesmo em gel de poliacrilamida 8% possibilitaram a identificação de problemas na reação de amplificação como, por exemplo, a presença de amplificação inespecífica, ou formação de dímeros de *primers*. Dos 41 genes candidatos selecionados para validação experimental, quatro candidatos apresentaram problemas de amplificação (BM 103, 107, 205 e 405) e não puderam ter suas reações padronizadas.

Do total de 533 genes re-induzidos pelo tratamento com 5-aza-dC, 37 (~7%) foram testados nas linhagens celulares nas quais apresentaram re-indução da expressão pela metodologia de RaSH. Assim, os candidatos exclusivos de uma linhagem celular foram testados apenas nas amostras de

FUNDAÇÃO ANTONIO PRUDENTE

cDNA provenientes desta linhagem, sendo que os genes candidatos comuns a duas ou mais linhagens foram testados nas amostras de cDNA provenientes das linhagens nas quais foi identificado.

Para a linhagem MCF-7, dos 8 candidatos testados, apenas um (~12,5%) candidato (BM 001) apresentou indução de expressão superior a 2 vezes na cultura tratada em relação à não tratada, confirmando a re-indução da expressão gênica após o tratamento. Nenhum candidato, dos 7 testados, apresentou reativação de expressão superior a 2 vezes na linhagem MDA-MB-435 tratada em relação à não tratada. Dois, dos 10 (~20%) candidatos (BM 201 e BM 202) apresentaram expressão maior que 2 vezes na amostra tratada em relação à não tratada na linhagem MDA-MB-231. Dos cinco genes testados nas linhagens MCF-7 e MDA-MB-231, o gene candidato BM 006 (~20%), apresentou re-indução de expressão superior a 2 vezes nas duas linhagens celulares. O gene BM 404 também apresentou re-indução da expressão nas duas linhagens testadas, MDA-MB-435 e MDA-MB-231, num total de 10 genes testados (10%).

Os cinco genes que apresentaram reativação de sua expressão acima de 2 vezes nas linhagens celulares nas quais foram identificados como diferencialmente expressos pelo RaSH (BM 001, 006, 201, 202 e 404) foram testados também nas outras linhagens. Os valores da re-indução da expressão gênica de todos os candidatos testados estão mostrados na Tabela 11 e na Figura 16.

Aparentemente não há correlação entre a quantidade de seqüências de cada transcrito presentes na biblioteca de RaSH e a eficiência de

subtração e por, conseqüência, a ocorrência de falso-positivos na biblioteca subtrativa. Como exemplo, genes representados por até cinco seqüências (candidato BM 007, por exemplo) não apresentaram re-indução da expressão superior a 2 vezes quando analisado por PCR em Tempo Real enquanto outros com apenas um único clone (como exemplo o candidato BM 202) apresentaram maior nível de re-indução.

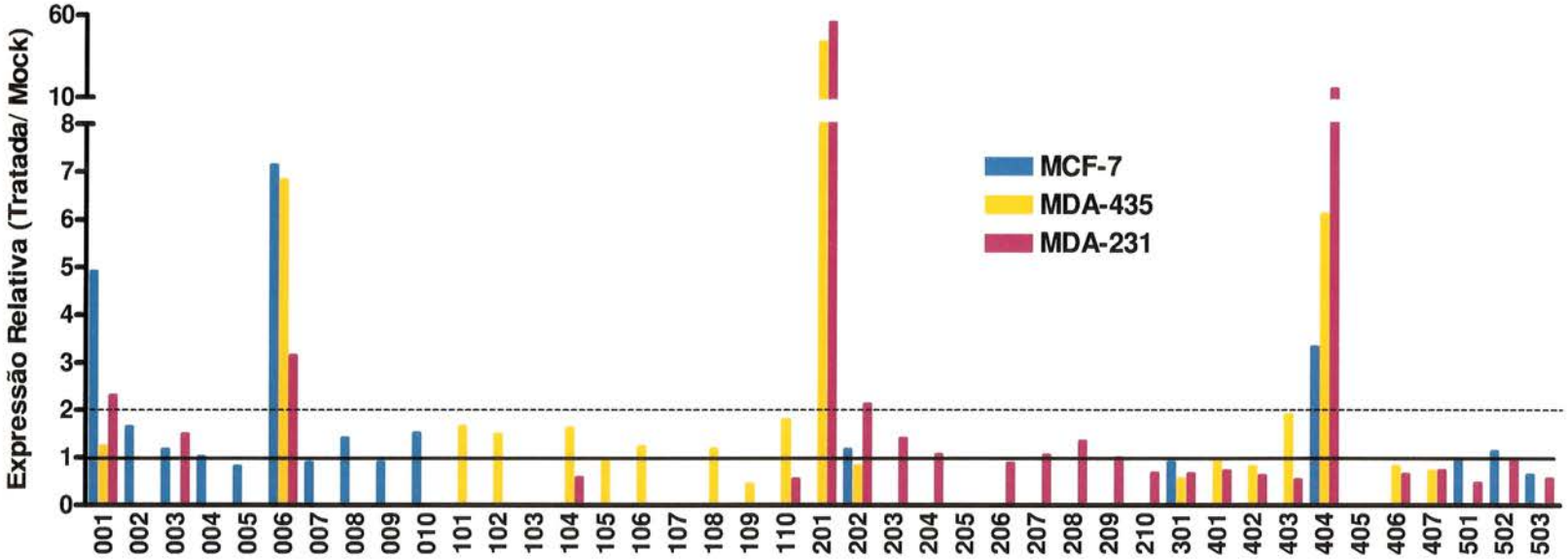
Ao todo, dos 37 candidatos testados, 5 (cerca de 13%) apresentaram reativação da expressão acima de 2 vezes após tratamento com 5-aza-dC. Estes resultados são comparáveis a resultados encontrados na literatura. JIANG et al. (2000), ao analisarem o padrão de expressão diferencial de células de melanoma humano tratadas ou não com IFN-β e mezereína verificaram, por experimentos de *Northern* Reverso, a expressão diferencial acima de 1,8 vezes, decorrente do tratamento, de 23 de 70 (~33%) clones originados de bibliotecas subtrativas de RaSH, escolhidos randomicamente. No entanto, no mesmo trabalho, ao aumentar o número de clones analisados, através da hibridização com as mesmas bibliotecas subtrativas de 1000 clones de cDNA originadas do RaSH fixados em lâminas de vidro e 400 clones em membranas, a porcentagem de genes diferencialmente expressos diminuiu para 11% e 16%, respectivamente. Estes valores aproximam-se dos valores obtidos no experimento de *Northern* Reverso (~11,4%) e PCR em Tempo Real (~13%), ambos realizados neste projeto Em um trabalho comparando duas linhagens celulares de melanoma humano com potenciais metastáticos diferentes, BOUKERCHE et al. (2004) ao analisarem clones gerados da construção de uma biblioteca subtrativa

através de *Northern* Reverso, obtiveram diferença de expressão maior que 2 vezes somente em 8 de 72 (~11%) genes escolhidos. Em outro trabalho utilizando a técnica de RaSH, de 224 clones originados de biblioteca subtrativa de RaSH, 22 clones representando 20 genes (cerca de 9,8%) apresentaram diferença de expressão por *Northern* Reverso (SCHMOLL et al. 2004).

Tabela 11 - Reativação da expressão gênica após tratamento com o agente desmetilante 5-aza-dC.

Candidato		Razão 5-aza-dC / Mock		
		MCF-7	*MDA-435*	*MDA-231*
001	SQSTM1	***4,91***	1,24	***2,31***
002	KRT19	1,65		
003	TRIP6	1,17		1,50
004	DKFZP564O0463	1,02		
005	PGR1	0,82		
006	SOD2	***7,14***	***6,84***	***3,14***
007	GNG5	0,89		
008	COX5B	1,41		
009	TOB1	0,90		
010	CRR9	1,52		
101	LDHA		1,65	
102	MGC8721		1,49	
103	RNF11			
104	TRA1		1,62	0,58
105	HSPCA		0,91	
106	RBBP6		1,23	
107	MAPK9			
108	CDC5L		1,19	
109	IMPDH2		0,45	
110	SYBL1		1,80	0,56
201	CTGF	0,00	***43,66***	***55,32***
202	SUMO2	1,17	0,83	***2,12***
203	CTSB			1,41
204	PTTG1IP			1,06
205	ICAM1			
206	TDE1			0,88
207	ST5			1,05
208	ERCC5			1,35
209	PERP			0,99
210	TRIP4			0,68
301	CDC27	0,90	0,56	0,66
401	MARCH5		0,94	0,72
402	DARS		0,81	0,62
403	SERPINE2		1,91	0,54
404	FN1	***3,32***	***6,12***	***14,93***
405	ATP6M8-9			
406	PITRM1		0,82	0,66
407	MO25		0,72	0,73
501	CALM1	0,96		0,47
502	BZW1	1,13		0,94
503	APPBP1	0,64		0,56

A tabela mostra os valores da expressão de cada gene candidato na linhagem celular tratada em relação à linhagem celular não tratada com 5-aza-dC. Os valores superiores a 2 vezes estão destacados do texto.



Legenda: A diferença de expressão gênica foi obtida com base na razão entre a expressão na linhagem tratada e na linhagem não tratada (mock), obtidas em experimentos de PCR quantitativa em Tempo Real. Os candidatos foram testados no mesmo cDNA utilizado para a construção da biblioteca subtrativa de origem e nas outras linhagens apenas nos casos que apresentaram reativação de expresso superior a 2 vezes. A linha cheia representa razão (Tratada / Mock) igual a 1 e a linha tracejada representa o valor de *cutoff* adotado (>2).

Figura 16 - Reativação da expressão gênica após tratamento com 5-aza-dC.

4.6 OS GENES CANDIDATOS VALIDADOS

O gene *SQSTM1* (*p62*, *Sequestosome 1*, candidato BM 001) apresentou re-indução de sua expressão, através da metodologia de PCR em Tempo Real, de 4,9 vezes na linhagem MCF-7 e 2,3 vezes na linhagem MDA-MB-231 após o tratamento com agente desmetilante. Este gene codifica uma proteína multifuncional que se liga a ubiquitina livre (VADLAMUDI et al. 1996) e pode representar um mecanismo regulatório mediado por ubiquitinação durante a proliferação e diferenciação celular (LEE et al. 1998). Diversos tecidos humanos adultos apresentam expressão constitutiva do gene *SQSTM1* e a expressão aumentada do mesmo foi detectada em diversos tumores, entre eles tumores de próstata (KITAMURA et al. 2006) e mama (THOMPSON et al. 2003). No entanto esses trabalhos não relatam a expressão alterada do gene *SQSTM1* na linhagem MCF-7, uma das linhagens nas quais o mesmo apresentou re-indução em nossas análises. Sabe-se que o *SQSTM1* é regulado por PDEF (*Prostate-Derived ETs Factor*) (THOMPSON et al. 2003), mas o papel da metilação na regulação de sua expressão permanece desconhecido, sendo possível que as diminuições no nível de metilação da região promotora sejam responsáveis pelo aumento da expressão observado em tumores

O gene *SOD2* (*Superoxide dismutase 2*, BM 006) apresentou re-indução de sua expressão superior a 2 vezes nas três linhagens celulares analisadas. Este gene codifica uma proteína mitocondrial responsável pela diminuição do estresse causado por espécies reativas de oxigênio (GUO et

al. 2003). Camundongos "*knock out*" para *SOD2* apresentaram maior susceptibilidade à injúria causada por O_2 (WILLIAMS et al. 1998). O gene *SOD2* também parece estar envolvido na proteção contra apoptose induzida por TNF-α (WONG et al. 1989) e apresenta expressão reduzida em muitos tipos de câncer e linhagens celulares tumorais (LI et al. 1995; YAN et al. 1996; OBERLEY, 2001). Na linhagem de mieloma múltiplo IM-9 o gene *SOD2* encontra-se hipermetilado (HODGE et al. 2005) e a transfecção do mesmo em linhagens de melanoma resulta na reversão do fenótipo tumoral (HUANG et al. 1999). No entanto, não há relatos de sua regulação por metilação em tumores de mama.

O gene *CTGF* (*Connective tissue growth factor*, BM 201) apresentou re-indução de expressão de 55,3 e 43,6 vezes nas linhagens MDA-MB-231 e MDA-MB-435, respectivamente. Não foram dectados transcritos de mRNA correspondentes a este gene na linhagem MCF-7, mesmo após o tratamento com 5-aza-dC. A proteína codificada por este gene parece estar envolvida diretamente na condrogênese e na ossificação, apresentando a habilidade de promover proliferação, maturação e hipertrofia de células de cartilagem em cultura (TAKIGAWA et al. 2003). O gene *CTGF* apresenta expressão aumentada em tecidos afetados por injúria ou patologias e áreas fibro-proliferativas e marcadamente reduzida em tumores de mama, havendo correlação estatística com recorrência local, pior prognóstico, metástase e mortalidade (JIANG et al. 2004). Em linhagens celulares de hepatoma, a expressão de *CTGF* foi re-induzida por tratamento com 5-aza-dC e a redução de sua expressão pode ser atribuída à metilação aberrante de seu

FUNDAÇÃO ANTONIO PRUDENTE BIBLIOTECA

promotor gênico (CHIBA et al. 2005). Assim como em nosso trabalho, (HISHIKAWA et al. 1999) não detectaram níveis de mRNA de *CTGF* na linhagem tumoral MCF-7 e a sua super-expressão deste gene na mesma linhagem provocou altos índices de morte celular por apoptose, sugerindo que *CTGF* pode atuar como um gene supressor de tumor.

O gene *SUMO-2* (*Small ubiquitin-like modifier 2*, BM 202) apresentou re-indução e teve reativação acima de duas vezes (2,12 vezes) apenas na linhagem MDA-MB-231. A proteína codificada por este gene faz parte de uma família, na qual estão incluídas *SUMO-1, -2* e *-3*, que realizam a modificação pós-traducional de proteínas de maneira análoga à ubiquitinação. Diferentemente da última, que direciona proteínas-alvo à sua degradação, as proteínas da família SUMO participam de muitos processos celulares como transporte nuclear, regulação transcricional, apoptose e estabilidade protéica (SU e LI 2002). Os genes *SUMO-2* e *-3* apresentam entre si 95% de identidade e têm 46% e 48% de identidade com *SUMO-1*, respectivamente (ROSENDORFF et al. 2006). A proteína SUMO-2 está geralmente relacionada com a repressão transcricional de co-repressores, co-ativadores, histonas, enzimas modificadoras de histonas ou fatores de transcrição. Em experimentos nos quais foram realizadas mutações que reprimem ligações a proteínas SUMO, houve aumento de atividade dos fatores de transcrição Elk-1, Sp3, c-myb, c-jun e AP2 (ROSS et al. 2000; SAPETSCHNIG et al. 2002; GILL 2003; GIRDWOOD et al. 2003; YANG e SHARROCKS 2005). Não há relatos do envolvimento desta proteína com patologias, como o câncer, por exemplo, mas seu papel repressor pode

regular a transcrição de outros genes sabidamente relacionados à tumorigênese. Até o momento não há relatos do papel da metilação em sua regulação.

O gene *FN1* (*Fibronectin 1,* BM 404) apresentou re-indução de expressão 3,32 vezes superior na linhagem MCF-7, 6,11 vezes na linhagem MDA-MB-435 e 14,93 vezes na linhagem MDA-MB-231. Este gene codifica uma glicoproteína de alto peso molecular, a fibronectina. Representa cerca de 1% das proteínas presentes no soro, além de ser o principal componente da matriz extracelular presente no tecido mamário. Esta proteína está ausente ou com expressão muito reduzida em diversas células transformadas. Até o momento foi relatado o papel desta proteína na adesão celular, morfologia, e arquitetura da superfície celular. Sua ausência está relacionada à perda da inibição da proliferação por contato. Muitos tipos celulares dependem da fibronectina para a ligação ao colágeno. A fibronectina é codificada por um único gene, em duas isoformas, uma celular e transcrita por fibroblastos em cultura e células tumorais e identificada no presente estudo, e outra produzida pelo fígado e secretada ao plasma. O gene FN1 apresenta densa metilação em linhagens celulares de hepatoma e sofre reativação da expressão gênica após tratamento com 5-aza-dC (CHIBA et al. 2005).

As alterações genéticas freqüentes no câncer incluem mutações pontuais, deleções gênicas e perda da heterosigosidade (LOH), além de alterações epigenéticas (COSTELLO e PLASS 2001). Os mecanismos genéticos e epigenéticos poderiam contribuir, individualmente ou

combinados, para o silenciamento de genes supressores de tumor, de acordo com as seguintes combinações: (i) um dos alelos poderia sofrer mutações pontuais, enquanto o outro alelo poderia sofrer hipermetilação de sua região promotora; (ii) hipermetilação homozigótica; (iii) deleção gênica de um dos alelos e metilação do outro; (iv) mutações ou deleções homozigóticas ou combinadas entre si. Assim, uma região que apresenta alta freqüência de deleções em tecidos tumorais pode sugerir a localização de um importante gene supressor de tumor. A perda da expressão deste supressor de tumor, seja por sua mutação pontual, deleção cromossômica ou por silenciamento epigenético, poderia contribuir para a tumorigênese e para a progressão tumoral (JONES e BAYLIN 2002).

A região cromossômica dos genes candidatos validados foi então analisada quanto à recorrência de aberrações cromossômicas freqüentes no câncer de mama, mais especificamente deleções e LOH. A *webpage* do "*Cancer Genome Anatomy Project*" (CGAP, http://cgap.nci.nih.gov/chromosomes/RecurrentAberrations) possibilita a análise de uma determinada região cromossômica com alterações recorrentes no câncer existentes na literatura. Através da página é possível realizar o rastreamento da literatura de acordo com o tipo de alteração (numérica ou estrutural), por banda cromossômica ou por cromossomo individualmente, tipo de tecido e diferentes tipos de tumor, além de citar e direcionar o usuário para o número de casos onde a aberração foi encontrada.

A ocorrência de monossomia envolvendo o cromossomo onde se encontra o potencial supressor de tumor poderia ser um do passos na tumorigênese, onde a ausência deste cromossomo causaria LOH, e o silenciamento epigenético causaria a perda do segundo alelo. A monossomia do cromossomo 5, onde está localizado o gene *SQSTM1* (BM 001), foi relatada em 50 casos de câncer de mama, sendo 2 deles carcinoma *in situ* e 48 casos de adenocarcinoma. O cromossomo 2, onde encontra-se o gene *FN1* (BM 404), encontra-se em monossomia em mais de 50 casos, a maioria adenocarcinomas de mama, assim como o cromossomo 6, que contém os genes *SOD2* (BM 006) e *CTGF* (BM 201), e que também encontra-se em monossomia em mais de 50 casos envolvendo adenocarcinomas, tumores benignos epiteliais e lesões não neoplásicas de mama. Além disso, a perda de heterozigosidade (LOH) deste último cromossomo é encontrada em mais de 40% dos melanomas malignos (HUANG et al. 1997). O cromossomo 17, onde se localiza o gene *SUMO2* (BM 202), encontra-se em monossomia em mais de 100 casos envolvendo adenocarcinomas, *Phyllodes* malignos, hiperplasias da mama, e carcinomas *in situ*.

A busca por deleções freqüentes nas bandas cromossômicas onde se encontram os candidatos validados revelou a ocorrência de deleções em mais de 30 casos de câncer de mama na região 6q25. Nesta banda cromossômica localiza-s o gene *SOD2* (BM 006) e dentre os tumores mais freqüentemente encontrados com deleção nesta região temos tumores benignos epiteliais da mama, os quais respondem por mais de 50% dos

casos, seguidos por carcinomas da mama. A banda 6q23, onde se localiza o gene *CTGF* (BM 201) encontra-se deletada em 15 casos na literatura, sendo 13 deles casos de adenocarcinomas do tecido mamário. A deleção da região 17q25, onde está localizado o gene *SUMO2* (BM 202), foi observada em 2 casos de adenocarcinoma da mama. As regiões 5q35, onde está localizado o gene *SQSTM1* (BM 001) e 2q34, onde localiza-se o gene *FN1* (BM 404), não apresentam deleções freqüentes em tumores de mama. No entanto outros tipos de aberrações cromossômicas numéricas e estruturais também relacionadas ao câncer foram relatadas nestas duas regiões como, por exemplo, trissomias e adições. O resumo das aberrações estruturais recorrentes em tumores de mama na região onde estão situados os genes validados está representado na Tabela 12.

Tabela 12 - Aberrações estruturais encontradas regiões cromossômicas dos cinco genes re-induzidos pelo tratamento com 5-aza-dC.

Gene	*Banda*	*Anormalidade*	*Morfologia*	*Topografia*	*Casos*
SQSTM1	5q35	add(5)(q35)	Adenocarcinoma	Mama	4
SOD2	6q25	add(6)(q25)	Adenocarcinoma	Mama	2
	6q25	***del(6)(q25)***	***Adenocarcinoma***	***Mama***	***3***
	6q25	***del(6)(q25)***	***Tumor Benigno Epitelial, NOS***	***Mama***	***18***
	6q25	***del(6)(q25)***	***Tumor Benigno Epitelial, tipo especial***	***Mama***	***2***
	6q25	***del(6)(q25)***	***Carcinoma, NOS***	***Mama***	***8***
CTGF	6q23	t(6;12)(q23;q13)	Adenocarcinoma	Mama	2
	6q23	add(6)(q23)	Adenocarcinoma	Mama	4
	6q23	***del(6)(q21q23)***	***Adenocarcinoma***	***Mama***	***3***
	6q23	***del(6)(q23)***	***Adenocarcinoma***	***Mama***	***10***
	6q23	***del(6)(q23)***	***Tumor Benigno Epitelial, tipo especial***	***Mama***	***2***
SUMO2	17q25	add(17)(q25)	Adenocarcinoma	Mama	3
	17q25	***del(17)(q25)***	***Adenocarcinoma***	***Mama***	***2***
	17q25	hsr(17)(q25)	Adenocarcinoma	Mama	3
FN1	2q34	add(2)(q34)	Leucemia Leucoblástica Aguda/Linfoma		4
	2q34	add(2)(q34)	Leucemia Mieloblástica		2

As deleções estão destacadas em negrito itálico. (del) deleção; (add) adição; (t) translocação; (hsr) regiões homogeneamente coradas; (NOS) não especificado.

Estes dados podem ser considerados indícios da participação destes genes no câncer de mama. Porém, é importante frisar que células tumorais geralmente apresentam grande número de aberrações estruturais e numéricas causadas simplesmente pelas disfunções em sua maquinaria de replicação. Portanto, estudos mais detalhados destas regiões no câncer de mama são de extrema importância para o entendimento do real papel destes genes no câncer.

4.7 ANÁLISE PRELIMINAR DA EXPRESSÃO DOS GENES VALIDADOS EM AMOSTRAS DE TUMORES PRIMÁRIOS DE MAMA

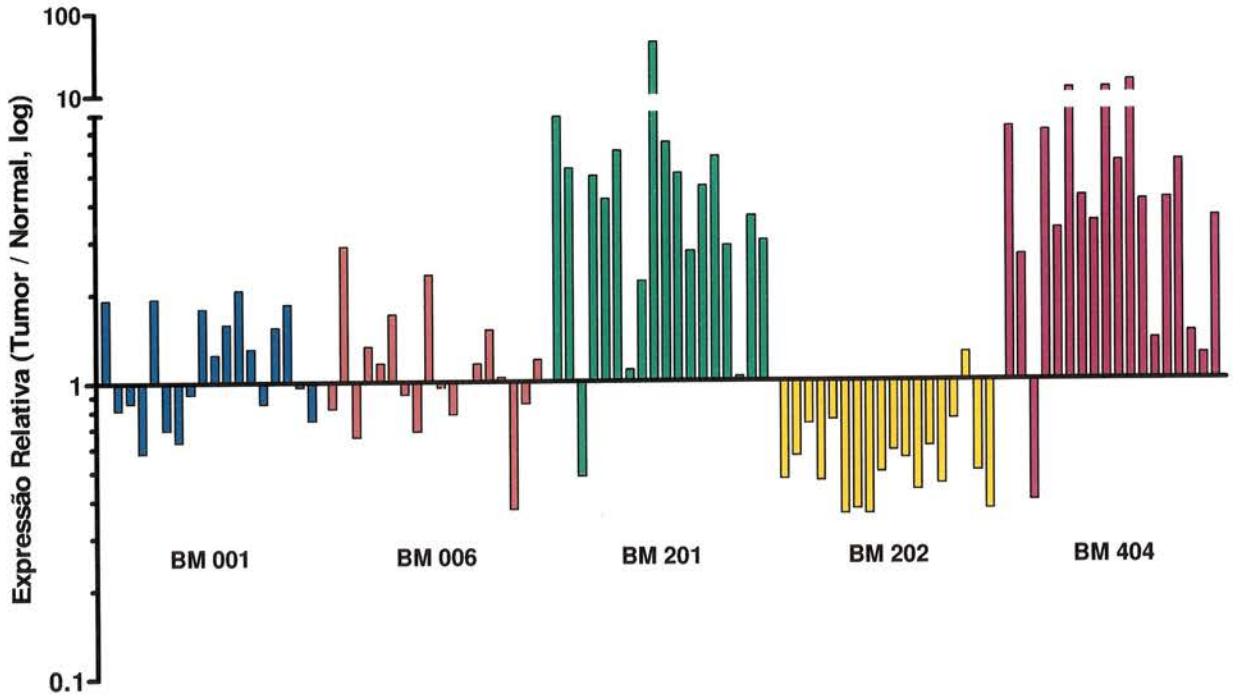
O estudo de expressão gênica em linhagens celulares possui muitas vantagens, como a facilidade de extração de RNA e DNA de boa qualidade, possibilidade de cultura em diferentes condições e maior facilidade na realização de ensaios biológicos *in vitro* e *in vivo*. Por outro lado, sabe-se que células em cultura podem apresentar variações em seu perfil de expressão gênica devido a artefatos ocasionados pelo novo microambiente. Diferentes condições de cultura e as sucessivas passagens às quais as células são submetidas podem formar diferentes tipos de pressão seletiva, as quais podem gerar não só a perda de características do tumor de origem como também diferenças entre a mesma linhagem tumoral cultivada em dois laboratórios diferentes. Fica, portanto, clara a necessidade de análise da expressão dos candidatos validados em amostras de tumores de pacientes com câncer de mama.

Para uma análise preliminar, o padrão de expressão dos cinco candidatos validados em linhagens celulares foi analisado em 18 amostras de tumor de mama, obtidas do Banco de Tumores do Hospital do Câncer A.C. Camargo e do Hospital do Câncer Alfredo Abrão, comparado a um *pool* de 3 amostras de tecido mamário normal obtido comercialmente (CLONTECH®). A partir desta análise pretendíamos apenas verificar se os candidatos validados nas linhagens celulares possuíam expressão ausente ou reduzida em tumores quando comparados com amostra de tecido mamário normal, o que seria esperado de um marcador molecular silenciado por metilação.

As amostras de tumores de mama eram representativas dos 4 estádios clínicos da doença: 6 amostras do estádio I, 5 amostras do estádio II, 3 do estádio III e 4 do estádio IV. Na quantificação relativa por PCR em Tempo Real, cDNA proveniente de tecido mamário normal foi usado como amostra referência. Os quatro genes de expressão constitutiva disponíveis em nosso laboratório (*ACTB, GAPDH, BCRP* e *HPRT*) foram utilizados como normalizadores. A ferramenta computacional geNorm®, disponível gratuitamente na Internet (http://medgen.ugent.be/~jvdesomp/genorm), permite a análise de diversos controles endógenos dentro do painel de amostras utilizado, possibilitando a escolha dos genes que melhor se comportam como normalizadores para aquele grupo de amostras. Os quatro genes foram considerados normalizadores adequados e um fator de normalização, baseado na média geométrica dos valores da expressão de cada gene, foi calculado para cada uma das amostras. Os fatores de

normalização obtidos foram utilizados para o cálculo dos valores de expressão relativa (razão tumor/normal), representados na Figura 17.

Os genes *SQSTM1* e *SOD2* apresentaram pouca variação de expressão entre as amostras tumorais em relação ao tecido mamário normal, sendo que o último apresentou redução da expressão superior a 2 vezes em apenas uma amostra tumoral. Já os genes *CTGF* e *FN1* apresentaram aumento de expressão superior a 2 vezes em 15/18 e 14/18 amostras analisadas, respectivamente. O gene *SUMO2* apresentou redução em 17/18 amostras tumorais analisadas, sendo esta superior a duas vezes em 10 destas amostras (cerca de 55,5%).



Legenda: A expressão dos genes BM 001, 006, 201, 202 e 404 foi quantificada por PCR em Tempo Real, em 18 amostras de tumor de mama de diferentes estádios clínicos. A expressão dos genes candidatos nos tumores estão representadas em relação a tecido de mama normal (Razão Tumor/ Normal). Como controle endógeno, foi utilizado um fator de normalização calculado a partir da expressão dos genes *HPRT, BCRP, β-Actin* e *GAPDH*. O eixo Y está representado em escala logarítmica (log10).

Figura 17 - Análise preliminar da expressão gênica dos cinco genes validados em amostras tumorais de mama.

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenador de Ensino

FUNDAÇÃO ANTÔNIO PRUDENTE

A re-indução dos genes *SQSTM1*, *SOD2*, *CTGF* e *FN1* nas linhagens celulares testadas, após tratamento das mesmas com agente desmetilante, não indica, necessariamente, que os mesmos encontravam-se silenciados por metilação na linhagem não tratada e que o tratamento tenha provocado sua re-ativação. Nas linhagens foi constatada apenas a re-indução da expressão gênica após o tratamento. Assim, é possível que estes genes apresentassem nas linhagens pequenos níveis de metilação e expressão gênica ativa ou moderadamente ativa, podendo a desmetilação causada pelo tratamento ter contribuído apenas para o aumento de sua expressão. Portanto, a falta de correlação dos níveis de expressão destes quatro genes entre as linhagens celulares e as amostras tumorais pode ser resultado simplesmente da ausência de silenciamento por metilação nas primeiras.

Vale ressaltar ainda a diferença entre linhagens celulares e amostras tumorais quanto ao seu padrão de metilação. Linhagens celulares podem apresentar níveis elevados de metilação quando comparadas ao seu tumor de origem (SMIRAGLIA et al. 2001). Os autores, ao analisarem o padrão de metilação de 1184 regiões promotoras pela técnica de *Restriction Landmark Genomic Scanning* (RLGS), de 114 amostras tumorais e 24 linhagens celulares tumorais estabelecidas originadas de 8 tecidos tumorais diferentes, descrevem o aumento da taxa de metilação em linhagens celulares. Os resultados indicam graus elevados de metilação de no mínimo 5 e no máximo 93 vezes em linhagens celulares quando comparadas com seu tumor de origem. Linhagens celulares de câncer de mama apresentam, por exemplo, aumento de 43 vezes na metilação em relação ao tecido mamário

tumoral. Assim, segundo os autores, a maioria da metilação encontrada em linhagens celulares seria derivada de artefatos gerados pela cultura *in vitro* e devem ser analisadas com cautela.

Outro fator a ser relevado neste tipo de análise é a influência que o estroma exerce no tumor (BIRGERSDOTTER et al. 2005). Células que o circundam, como fibroblastos, células endoteliais, células do sistema imune, juntamente com a matriz extracelular são participantes ativas na formação, crescimento e estimulação do tumor. Estas células podem apresentar níveis de expressão gênica diferentes do gene em questão, fornecendo assim uma interpretação errônea de sua expressão. Portanto, devido à alta sensibilidade da metodologia de PCR em Tempo Real, pode se tornar difícil a visualização da expressão gênica diferencial em amostras passíveis de possuir contaminação com células estromais. A microdissecção das amostras tumorais por um profissional bem treinado pode excluir os eventuais contaminantes estromais como linfócitos infiltrantes, células normais da borda tumoral ou células endoteliais.

A falta de correlação entre os resultados apresentados nas linhagens e aqueles apresentados em tumores para a maioria dos candidatos pode, portanto, refletir a presença de artefatos gerados simplesmente pela cultura *in vitro*, pelos métodos de extração de RNA - os quais não excluem células normais contaminates - ou até pelas características do tipo de célula predominante na cultura celular e na amostra tumoral. É importante ressaltar que este grupo de amostras tumorais é pequeno e um número maior de

FUNDAÇÃO ANTONIO
DENTE

amostras tumorais e tecidos normais poderia fornecer uma análise mais precisa do padrão de expressão destes genes.

Não há, na literatura, relatos do envolvimento do gene *SUMO2* com a tumorigênese e progressão tumoral. A indução do mesmo pelo tratamento com 5-aza-dC e sua expressão reduzida nas amostras tumorais sugerem o mesmo como um potencial marcador molecular de metilação para o câncer de mama. Sendo assim, estudos mais detalhados devem ser realizados para verificar a ocorrência de correlação significativa entre os níveis de expressão deste gene e dados clínico-patológicos como, por exemplo, o estadiamento clínico, a sobrevida global e livre de doença, surgimento de metástases linfonodais e à distância e a expressão de marcadores clássicos de prognóstico como o Receptor de Estrógeno e o oncogene *c-erbB2*.

O tratamento do DNA genômico das linhagens celulares com o Bissulfito de Sódio, seguido de seqüenciamento da Ilha de CpG da região promotora do gene *SUMO2*, poderá revelar o padrão de metilação nesta região comprovando a participação da metilação na redução da expressão e na re-indução provocada pela desmetilação de seu promotor. A análise do padrão de metilação em um grande número de amostras tumorais de mama possibilitará a verificação da ocorrência de metilação aberrante no câncer de mama. A existência de correlação destes resultados com dados clínico-patológicos possibilitará a utilização da metilação como marcador molecular desta doença.

5 CONCLUSÕES

- A análise *in silico* das bibliotecas subtrativas e suas bibliotecas reversas possibilitou a identificação e exclusão de falso-positivos, presentes tanto em amostras tratadas quanto em amostras não tratadas com agente desmetilante, além de gerar dados importantes, como a comparação entre os painéis de genes reativados nas diferentes linhagens;
- Mais de 2300 clones foram seqüenciados e possibilitaram a identificação de 533 genes possivelmente induzidos pelo tratamento com 5aza-dC;
- 41 genes foram selecionados para a validação experimental por PCR em Tempo Real. 37 deles tiveram sua expressão comparada entre as linhagens tratada e não tratada. Cinco (~13%) deles apresentaram reativação superior a 2 vezes (*SQSTM1, SOD2, CTGF, SUMO2* e *FN1*);
- O gene *SUMO2* apresentou redução de expressão em amostras de tumores de mama quando comparada à expressão em um *pool* de amostras de tecido mamário normal, sendo este um possível candidato a marcador molecular de mama.

6 REFERÊNCIAS BIBLIOGRÁFICAS

Abba M, Hu Y, Sun H, et al. Gene expression signature of estrogen receptor α status in breast cancer. **BMC Genomics** 2005; 6:1471-2164/6/37.

[ACS] American Cancer Society. **Cancer facts and figures**. Available from: <URL:http://www.cancer.org/downloads/STT/CAFF2005BrF.pdf> [2006 jun 12].

Ammanamanchi S, Kim SJ, Sun LZ, Brattain MG. Induction of transforming growth factor-beta receptor type II expression in estrogen receptor-positive breast cancer cells through SP1 activation by 5-aza-2'-deoxycytidine. **J Biol Chem** 1998; 273:16527-34.

Anker P, Mulcahy H, Chen XQ, Stroun M. Detection of circulating tumour DNA in the blood (plasma/serum) of cancer patients. **Cancer Metastasis Rev** 1999; 18:65-73.

Attwood JT, Yung RL, Richardson BC. DNA methylation and the regulation of gene transcription. **Cell Mol Life Sci** 2002; 59:241-57.

Bae SN, Arand G, Azzam H, et al. Molecular and cellular analysis of basement membrane invasion by human breast cancer cells in Matrigel-based in vitro assays. **Breast Cancer Res Treat** 1993; 24:241-55.

Baylin SB, Herman JG, Graff JR, Vertino PM, Issa JP. Alterations in DNA methylation: a fundamental aspect of neoplasia. **Adv Cancer Res** 1998; 72:141-96.

Biéche I, Lidereau R. Genetic alterations in breast cancer. **Genes Chromosomes Cancer** 1995; 14:227-51.

Birgersdotter A, Sandberg R, Ernberg I. Gene expression perturbation in vitro--a growing case for three-dimensional (3D) culture systems. **Semin Cancer Biol** 2005; 15:405-12.

Boukerche H, Su ZZ, Kang DC, Fisher PB. Identification and cloning of genes displaying elevated expression as a consequence of metastatic progression in human melanoma cells by rapid subtraction hybridization. **Gene** 2004; 343:191-201.

Boyd N. Câncer de mama. In: Love RR, editor. **Manual de oncologia clínica**. Trad de I Carelli Filho. 6 ed. São Paulo: Fundação Oncocentro de São Paulo São Paulo; 1999. p.360-81.

Cailleau R, Olive M, Cruciger QV. Long-term human breast carcinoma cell lines of metastatic origin: preliminary characterization. **In vitro** 1978; 14:911-5.

Chen J, Rocken C, Klein-Hitpass L, et al. Microarray analysis of gene expression in metastatic gastric cancer cells after incubation with the methylation inhibitor 5-aza-2'-deoxycytidine. **Clin Exp Metastasis** 2004; 21: 389-97.

Chiba T, Yokosuka O, Fukai K, et al. Identification and investigation of methylated genes in hepatoma. **Eur J Cancer** 2005; 41:1185-94.

Chirgwin JM, Przybyla AE, MacDonald RJ, Rutter WJ. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. **Biochemistry** 1979; 18:5294-9.

Christman JK. 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. **Oncogene** 2002; 21:5483-95.

Colot V, Rossignol JL. Eukaryotic DNA methylation as an evolutionary device. **Bioessays** 1999; 21:402-11.

Costello JF, Plass C. Methylation matters. **J Med Genet** 2001; 38:285-303.

Dammann R, Yang G, Pfeifer GP. Hypermethylation of the cpG island of Ras association domain family 1A (RASSF1A), a putative tumor suppressor gene from the 3p21.3 locus, occurs in a large percentage of human breast cancers. **Cancer Res** 2001; 61:3105-9.

De Smet C, De Backer O, Faraoni I, Lurquin C, Brasseur F, Boon T. The activation of human gene MAGE-1 in tumor cells is correlated with genome-wide demethylation. **Proc Natl Acad Sci U S A** 1996; 9;93:7149-53.

De Smet C, Loriot A, Boon T. Promoter-dependent mechanism leading to selective hypomethylation within the 5' region of gene MAGE-A1 in tumor cells. **Mol Cell Biol** 2004; 24:4781-90.

Diatchenko L, Lau YF, Campbell AP, et al. Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. **Proc Natl Acad Sci U S A** 1996; 93:6025-30.

Domman FE, Rice JC, Hendrix MJ, Futscher BW. Epigenetic silencing of maspin gene expression in human breast cancers. **Int J Cancer** 2000; 85:805-10.

Eads CA, Danenberg KD, Kawakami K, et al. MethyLight: a high-throughput assay to measure DNA methylation. **Nucleic Acids Res** 2000; 28:E32.

Engel LW, Young NA, Tralka TS, Lippman ME, O'Brien SJ, Joyce MJ. Establishment and characterization of three new continuous cell lines derived from humam breast carcinomas. **Cancer Res** 1978; 38:3352-64.

Esteller M, Corn PG, Urena JM, Gabrielson E, Baylin SB, Herman JG. Inactivation of glutathione S-transferase P1 gene by promoter hypermethylation in human neoplasia. **Cancer Res** 1998; 58:4515-8

Esteller M, Guo M, Moreno V, et al. Hypermethylation-associated Inactivation of the Cellular Retinol-Binding-Protein 1 Gene in Human Cancer. **Cancer Res** 2002; 62:5902-5.

Evron E, Umbricht CB, Korz D, et al. Loss of cyclin D2 expression in the majority of breast cancers is associated with promoter hypermethylation. **Cancer Res** 2001; 61:2782-7.

Frigola J, Song J, Stirzaker C, Hinshelwood RA, Peinado MA, Clark SJ. Epigenetic remodeling in colorectal cancer results in coordinate gene suppression across an entire chromosome band. **Nat Genet** 2006; 38:540-9.

Fukai K, Yokosuka O, Chiba T, et al. Hepatocyte growth factor activator inhibitor 2/placental bikunin (HAI-2/PB) gene is frequently hypermethylated in human hepatocellular carcinoma. **Cancer Res** 2003; 63:8674-9.

Gama-Sosa MA, Slagel VA, Trewyn RW, et al. The 5-methylcytosine content of DNA from human tumors. **Nucleic Acids Res** 1983; 11:6883-94.

Gately S. The contributions of cyclooxygenase-2 to tumor angiogenesis. **Cancer Metastasis Rev** 2000; 19:19-27.

Ghazi H, Magewu AN, Gonzales F, Jones PA. Changes in the allelic methylation patterns of c-H-ras-1, insulin and retinoblastoma genes in human development. **Dev Suppl** 1990; 115-23.

Gill G. Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity. **Curr Opin Genet Dev** 2003; 13:108-13.

Girdwood D, Bumpass D, Vaughan OA, et al. P300 transcriptional repression is mediated by SUMO modification. **Mol Cell** 2003; 11:1043-54.

Graff JR, Gabrielson E, Fujii H, Baylin SB, Herman JG. Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. **J Biol Chem** 2000; 275:2727-32.

Greger V, Passarge E, Hopping W, Messmer E, Horsthemke B. Epigenetic changes may contribute to the formation and spontaneous regression of retinoblastoma. **Hum. Genet** 1989; 83:155-8.

Guo G, Yan-Sanders Y, Lyn-Cook BD, et al. Manganese superoxide dismutase-mediated gene expression in radiation-induced adaptive responses. **Mol Cell Biol** 2003; 23:2362-78.

Hanahah D, Weinberg RA. Hallmarks of cancer. **Cell** 2000; 100: 57-70.

Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. **Proc Natl Acad Sci U S A** 1996; 93:9821-6.

FUNDAÇÃO ANTONIO
DE ... NTE
... TECA

Herman JG. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. **Proc Natl Acad Sci U S A** 1994; 91:9700-9704.

Hishikawa K, Oemar BS, Tanner FC, Nakaki T, Luscher TF, Fujii T. Connective tissue growth factor induces apoptosis in human breast cancer cell line MCF-7. **J Biol Chem** 1999; 274:37461-6.

Hodge DR, Xiao W, Peng B, Cherry JC, Munroe DJ, Farrar WL. Enforced expression of superoxide dismutase 2/manganese superoxide dismutase disrupts autocrine interleukin-6 stimulation in human multiple myeloma cells and enhances dexamethasone-induced apoptosis. **Cancer Res** 2005; 65:6255-63.

Honorio S, Agathanggelou A, Schuermann M, et al. Detection of RASSF1A aberrant promoter hypermethylation in sputum from chronic smokers and ductal carcinoma in situ from breast cancer patients. **Oncogene** 2003; 22:147-50.

Hospital do Câncer A. C. Camargo. **O que é câncer?** <URL: http://www.hcanc.org.br/intro/crom1.html> [2006 fev 10].

Howe LR, Subbaramaiah K, Brown AM, Dannenberg AJ. Cyclooxygenase-2: a target for the prevention and treatment of breast cancer. **Endocr Relat Cancer** 2001; 8:97-114.

Huang Y, He T, Domann FE. Decreased expression of manganese superoxide dismutase in transformed cells is associated with increased cytosine methylation of the SOD2 gene. **DNA Cell Biol** 1999; 18:643-52.

Huang Y, Peng J, Oberley LW, Domann FE. Transcriptional inhibition of manganese superoxide dismutase (SOD2) gene expression by DNA methylation of the 5' CpG island. **Free Radic Biol Med** 1997; 23:314-20.

Jemal A, Siegel R, Ward E, et al. Cancer Statistics. **CA Cancer J Clin** 2006; 56:106-30.

Jiang H, Kang DC, Alexandre D, Fisher PB. RaSH, a rapid subtraction hybridization approach for identifying and cloning differentially expressed genes. **Proc Natl Acad Sci U S A** 2000; 97:12684-9.

Jiang WG, Watkins G, Fodstad O, Douglas-Jones A, Mokbel K, Mansel RE. Differential expression of the CCN family members Cyr61, CTGF and Nov in human breast cancer. **Endocr Relat Cancer** 2004; 11:781-91.

Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. **Nat Rev Genet** 2002; 3:415-28.

Kalkhoven E, Roelen BA, de Winter JP, et al. Resistance to transforming growth factor beta and activin due to reduced receptor expression in human breast tumor cell lines. **Cell Growth Differ** 1995; 6:1151-61.

Kaneda A, Kaminishi M, Yanagihara K, Sugimura T, Ushijima T. Identification of silencing of nine genes in human gastric cancers. **Cancer Res** 2002; 62:6645-50.

Kang D, Jiang H, Wu Q, Pestka S, Fisher PB. Cloning and characterization of human ubiquitin-processing protease-43 from terminally differentiated human melanoma cells using a rapid subtraction hybridization protocol RaSH. **Genes Chromosomes Cancer** 2001; 267:233-42.

Kang DC, LaFrance R, Su ZZ, Fisher PB. Reciprocal subtraction differential RNA display: an efficient and rapid procedure for isolating differentially expressed gene sequences. **Proc Natl Acad Sci U S A** 1998; 95:13788-93.

Karpf AR, Lasek AW, Ririe TO, Hanks AN, Grossman D, Jones DA. Limited gene activation in tumor and normal epithelial cells treated with the DNA methyltransferase inhibitor 5-aza-2'-deoxycytidine. **Mol Pharmacol** 2004; 65:18-27.

Karpf AR, Peterson PW, Rawlins JT, et al. Inhibition of DNA methyltransferase stimulates the expression of signal transducer and activator of transcription 1, 2, and 3 genes in colon tumor cells. **Proc Natl Acad Sci U S A** 1999; 96:14007-12.

Kitamura H, Torigoe T, Asanuma H, et al. Cytosolic overexpression of p62 sequestosome 1 in neoplastic prostate tissue. **Histopathology** 2006; 48:157-61.

Krassenstein R, Sauter E, Dulaimi E, et al. Detection of breast cancer in nipple aspirate fluid by CpG island hypermethylation. **Clin Cancer Res** 2004; 10:28-32.

Lapidus RG, Ferguson AT, Ottaviano YL, et al. Methylation of estrogen and progesterone receptor gene 5' CpG islands correlates with lack of estrogen and progesterone receptor gene expression in breast tumors. **Clin Cancer Res** 1996; 2:805-10.

Lee JH, Welch DR. Supression of metastasis in human breast carcinoma MDA-MB-435 cells after transfection with the metastasis supressor gene, KiSS-1. **Cancer Res** 1997; 57:2384-7.

Lee YH, Ko J, Joung I, Kim JH, Shin J. Immediate early response of the p62 gene encoding a non-proteasomal multiubiquitin chain binding protein. **FEBS Lett** 1998; 438:297-300.

Li JJ, Oberley LW, St Clair DK, Ridnour LA, Oberley TD. Phenotypic changes induced in human breast cancer cells by overexpression of manganese-containing superoxide dismutase. **Oncogene** 1995; 10:1989-2000.

Liang G, Gonzales FA, Jones PA, Orntoft TF, Thykjaer T. Analysis of gene induction in human fibroblasts and bladder cancer cells exposed to the methylation inhibitor 5-aza-2'-deoxycytidine. **Cancer Res** 2002; 62: 961-6.

Lipsanen V, Leinonen P, Alhonen L, Janne J. Hypomethylation of ornithine decarboxylase gene and erb-A1 oncogene in human chronic lymphatic leukemia. **Blood** 1988; 72:2042-4.

Lisitsyn N, Lisitsyn N, Wigler M. Cloning the differences between two complex genomes. **Science.** 1993; 12;259:946-51.

Liu XH, Yao S, Kirschenbaum A, Levine AC. NS398, a selective cyclooxygenase-2 inhibitor, induces apoptosis and down-regulates bcl-2 expression in LNCaP cells. **Cancer Res** 1998; 58:4245-9.

Magewu AN, Jones PA. Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. **Mol Cell Biol** 1994; 14:4225-32.

Miettinen OS, Henschke CI, Pasmantier MW, Smith JP, Libby DM, Yankelevitz DF. Mammographic screening: no reliable supporting evidence? **Lancet** 2002; 359:404-5.

Ministério da Saúde. **Estimativa 2006: incidência de câncer no Brasil**. Rio de Janeiro: INCA; 2005.

Ministério da Saúde. **Câncer da mama: epidemiologia.** Disponível em: <URL:http//www.inca.gov.br/cancer/mama> [2006 jun 22]

Muller HM, Widschwendter A, Fiegl H, et al. DNA methylation in serum of breast cancer patients: an independent prognostic marker. **Cancer Res** 2003; 63:7641-5.

Oberley, LW. Anticancer therapy by overexpression of superoxide dismutase. **Antioxid Redox Signal** 2001; 3:461-72.

Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. **Lancet** 2001; 358:1340-2.

Palmisano WA, Divine KK, Saccomanno G, et al. Predicting lung cancer by detecting aberrant promoter methylation in sputum. **Cancer Res** 2000; 60:5954-8.

Parry D, Peters G. Temperature-sensitive mutants of p16CDKN2 associated with familial melanoma. **Mol Cell Biol** 1996; 16:3844-52.

Pfaffl MW. A new mathematical model for relative quantification in Real-Time RT-PCR. **Nucleic Acids Res** 2001; 29:2002-7.

Price JE, Polyzos A, Zhang RD, Daniels LM. Tumorigenicity and metastasis of human breast carcinoma cell lines in nude mice. **Cancer Res** 1990; 50:717-21.

Rahman N, Stratton MR. The genetic of breast cancer susceptibility. **Annu Rev Genet** 1998; 32:95-121.

Rosendorff A, Sakakibara S, Lu S, et al. NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. **Proc Natl Acad Sci U S A** 2006; 103:5308-13.

Ross DT, Scherf U, Eisen MB, et al. Systematic variation in gene expression patterns in human cancer cell lines. **Nat Genet** 2000; 24:227-35.

Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain terminating inhibitors. **Proc Natl Acad Sci U S A** 1977; 74:5463-7.

Sapetschnig A, Rischitor G, Braun H, et al. Transcription factor Sp3 is silenced through SUMO modification by PIAS1. **Embo J** 2002; 21:5206-15.

Schmoll M, Kubicek CP. ooc1, a unique gene expressed only during growth of Hypocrea jecorina (anamorph: Trichoderma reesei) on cellulose. **Curr Genet** 2005; 48:126-33.

Schmoll M, Zeilinger S, Mach RL, Kubicek CP. Cloning of genes expressed early during cellulase induction in Hypocrea jecorina by a rapid subtraction hybridization approach. **Fungal Genet Biol** 2004; 41:877-87.

Schmutte C, Jones PA. Involvement of DNA methylation in human carcinogenesis. **Biol Chem** 1998; 379:377-388.

Shi H, Wei SH, Leu YW, et al. Triple analysis of the cancer epigenome: an integrated microarray system for assessing gene expression, DNA methylation, and histone acetylation. **Cancer Res** 2003; 63:2164-71.

Silva JM, Dominguez G, Villanueva MJ, et al. Aberrant DNA methylation of the p16INK4a gene in plasma DNA of breast cancer patients. **Br J Cancer** 1999; 80:1262-4.

Smiraglia DJ, Rush LJ, Fruhwald MC, et al. Excessive CpG island hypermethylation in cancer cell lines versus primary human malignancies. **Hum Mol Genet** 2001; 10:1413-9.

Song JZ, Stirzaker C, Harrison J, Melki JR, Clark SJ. Hypermethylation trigger of the glutathione-S-transferase gene (GSTP1) in prostate cancer cells. **Oncogene** 2002; 21:1048-61.

Sorm F, Piskala A, Cihak A, Vesely J. 5-Azacytidine, a new, highly effective cancerostatic. **Experientia** 1964; 15:202-3.

Soule HD, Vazguez J, Long A, Albert S, Brennan M. A human cell line from a pleural effusion derived from a breast carcinoma. **J Natl Cancer Inst** 1973; 51:1409-16.

Stirzaker C, Song JZ, Davidson B, Clark SJ. Transcriptional gene silencing promotes DNA hypermethylation through a sequential change in chromatin modifications in cancer cells. **Cancer Res** 2004; 64:3871-7.

Styblo TM, Wood WC. Traditional prognostic factors for breast cancer. In: Bland KI, Copeland III EM, editors. **The breast: comprehensive management of benign and malignant diseases.** 2nd. Philadelphia: W B Saunders; 1998. p.419-25.

Su HL, Li SS. Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. **Genes Chromosomes Cancer** 2002; 21:296:65-73.

Su ZZ, Chen Y, Kang DC, et al. Customized rapid subtraction hybridization (RaSH) gene microarrays identify overlapping expression changes in human fetal astrocytes resulting from human immunodeficiency virus-1 infection or tumor necrosis factor-alpha treatment. **Gene** 2003; 306:67-78.

Su ZZ, Kang DC, Chen Y, et al. Identification and cloning of human astrocyte genes displaying elevated expression after infection with HIV-1 or exposure to HIV-1 envelope glycoprotein by rapid subtraction hybridization, RaSH. **Oncogene** 2002; 21:3592-602.

Takai N, Kawamata N, Walsh CS, et al. Discovery of epigenetically masked tumor suppressor genes in endometrial cancer. **Mol Cancer Res** 2005; 3:261-9.

Takigawa M, Nakanishi T, Kubota S, Nishida T. Role of CTGF/HCS24/ecogenin in skeletal growth control. **J Cell Physiol** 2003; 194:256-66.

Thompson EW, Paik S, Brunner N, et al. Association of increased basement membrane invasiveness with absence of estrogen receptor and expression of vimentin in human breast cancer cell lines. **J Cell Physiol** 1992; 150:534-44.

Thompson HG, Harris JW, Wold BJ, Lin F, Brody JP. p62 overexpression in breast tumors and regulation by prostate-derived Ets factor in breast cancer cells. **Oncogene** 2003; 22:2322-33.

Tlsty TD, Crawford YG, Holst CR, et al. Genetic and epigenetic changes in mammary epithelial cells may mimic early events in carcinogenesis. **J Mammary Gland Biol Neoplasia** 2004; 9:263-74.

Ushijima T, Okochi-Takada E. Aberrant methylations in cancer cells: where do they come from? **Cancer Sci** 2005; 96:206-11.

FUNDAÇ ANTONIO PRUDENTE BIBLIOTECA

Vadlamudi RK, Joung I, Strominger JL, Shin J. p62, a phosphotyrosine-independent ligand of the SH2 domain of p56lck, belongs to a new class of ubiquitin-binding proteins. **J Biol Chem** 1996; 23:20235-7.

van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. **Nature** 2002; 415:530-6.

Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biol** 2002; 3:Research0034

Velculescu VE, Zhang, L Vogelstein B, Kinzler, KW. Serial analysis of gene expression. **Science** 1995; 20;484-7.

Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. **Nature Biotechnology** 2000; 18:457-9.

Welch DR, Steeg PS, Rinker-Schaeffer CW. Molecular biology of breast cancer metastasis. Genetic regulation of human breast carcinoma metastasis. **Breast Cancer Res** 2000; 2:408-16.

Widschwendter M, Jones PA. DNA methylation and breast carcinogenesis. **Oncogene** 2002; 21:5462-82.

Williams MD, Van Remmen H, Conrad CC, Huang TT, Epstein CJ, Richardson A. Increased oxidative damage is correlated to altered mitochondrial function in heterozygous manganese superoxide dismutase knockout mice. **J Biol Chem** 1998; 273:28510-5.

Wong GH, Elwell JH, Oberley LW, Goeddel DV. Manganous superoxide dismutase is essential for cellular resistance to cytotoxicity of tumor necrosis factor. **Cell** 1989; 8:58:923-31.

Yamashita K, Upadhyay S, Osada M, et al. Pharmacologic unmasking of epigenetically silenced tumor suppressor genes in esophageal squamous cell carcinoma. **Cancer Cell** 2002; 2:485-95.

Yamashita S, Tsujino Y, Moriguchi K, Tatematsu M, Ushijima T. Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray. **Cancer Sci** 2006; 97:64-71.

Yan T, Oberley LW, Zhong W, St Clair DK. Manganese-containing superoxide dismutase overexpression causes phenotypic reversion in SV40-transformed human lung fibroblasts. **Cancer Res** 1996; 15;56:2864-71.

Yang J, Mani SA, Donaher JL, et al. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. **Cell** 2004; 117:927-39.

Yang SH, Sharrocks AD. PIASx acts as an Elk-1 coactivator by facilitating derepression. **Embo J** 2005; 24:2161-71.

Yuan BZ, Jefferson AM, Popescu NC, Reynolds SH. Aberrant gene expression in human non small cell lung carcinoma cells exposed to demethylating agent 5-aza-2'-deoxycytidine. **Neoplasia** 2004; 6:412-9.

Fundação Antônio Prudente
Luiz Fernando Lima Reis
Coordenadoria de Ensino

CURRICULUM VITAE

Dados Pessoais:

Nome: Murilo Vieira Geraldo
RG: 32568273-2
CPF: 222908798-35
E-mail: mu_vg@yahoo.com.br
Endereço: R. Dionísio Pedrellini, 86 Jd. Olímpia
São Paulo/SP C.E.P.: 05542-130
Data de Nascimento: 09 de Outubro de 1979
Local de Nascimento: São Paulo
Filiação: Ivaldir Geraldo e Elvira Rosa Vieira Geraldo

Formação Acadêmica:

Mestre em Ciências – Área de Concentração: Oncologia
Início: Agosto de 2004. Término: Outubro de 2006
Instituto Ludwig de Pesquisa sobre o Câncer, ILPC, São Paulo, Brasil
Título "Identificação de Genes Diferencialmente Metilados em Linhagens Tumorais de Mama através do Tratamento com Agente Desmetilante e Construção de Bibliotecas Subtrativas de cDNA".
Orientadora: Dra. Anamaria Aranha Camargo.
Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo.
Processo de número: 04/9574-0

2003-2004

Iniciação Científica em Biologia Molecular de Neoplasias Humanas, no Laboratório de Biologia Molecular e Genômica, Instituto Ludwig de Pesquisa sobre o Câncer.

2000-2003

Iniciação Científica em Biologia Molecular de *Leishmania sp*, no Laboratório de Leishmanioses, Departamento de Parasitologia, Instituto de Ciências Biomédicas da Universidade de São Paulo.

1999-2004

Bacharel em Ciências Biológicas pela Universidade de São Paulo.

Idiomas

Inglês – Teste de Proficiência Completo, União Cultural Brasil Estados Unidos. Média de 76%

Publicações Científicas:

GERALDO, Murilo V, SILBER, Ariel M, PEREIRA, Claudio A, ULIANA, Silvia RB

Characterisation of a developmentally regulated amino acid transporter gene from Leishmania amazonensis.
FEMS Microbiology Letters 242 (2005) 275–280.

Participação em Eventos:

Apresentação de Painel na XIII Reunião Anual da Sociedade Brasileira de Biologia Celular (SBBC), realizado junto ao Simpósio Internacional de Matriz Extra-Celular, Armação dos Búzios, RJ. 26 a 30 de Julho de 2006.

Apresentação de painel no IV São Paulo Research Conference - From Molecular Biology to Treatment, São Paulo, 10 a 12 de novembro de 2005. Applied Cancer Research Supplement 2005, v. 2, p. 70.

Apresentação de painel (2º autor) na 97ª Reunião Anual da Associação Americana de Pesquisa em Câncer (A.A.C.R.), março de 2006, Washington D.C. - EUA.

Apresentação de painel na XXXIV Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular (SBBq), Águas de Lindóia, 2 a 5 de julho de 2005.

Palestrante no WorkShop - Universidade Mackenzie "*O câncer na era da Biologia Molecular*" realizado pela Empresa Junior de Biologia da Universidade Mackenzie no dia 30 de março de 2005. Duração 90 minutos.

Apresentação de painel (2º autor) na Reunião Anual da Associação Americana de Pesquisa em Câncer (A.A.C.R.), março de 2005, Anaheim, California - EUA.

Apresentação de painel (2º autor) no II International Conference on Bioinformatics and Computational Biology (ICoBiCoBi), Angra dos Reis, 25 a 28 de outubro de 2004.

Apresentação de painel (2º autor) na XXXIII Reunião Anual da Sociedade Brasileira de Bioquímica e Biologia Molecular (SBBq), Caxambu, 16 A 19 de maio de 2004.

Simpósio Internacional – "50 anos do Hospital do Câncer – How Close Are We From Cancer Cure?" – São Paulo, agosto de 2003

Apresentação de painel na XXIX Reunião Internacional em Pesquisa Básica em Doença de Chagas, Caxambu, 4 a 6 de novembro de 2002