

IDENTIFICAÇÃO DE MARCADORES MOLECULARES EM CÂNCER DE MAMA RELACIONADOS À PROGRESSÃO E METÁSTASE

THIAGO FELCAR SARAIVA

**Dissertação apresentada à Fundação Antônio
Prudente para obtenção do título de mestre**

Área de concentração: Oncologia

Orientadora: Dra. Dirce Maria Carraro

**São Paulo
2008**

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Saraiva, Thiago Felcar

Identificação de marcadores moleculares em câncer de mama relacionados à progressão e metástase / Thiago Felcar Saraiva – São Paulo, 2008.

94p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

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

Orientadora: Dirce Maria Carraro

Descritores: 1. CÂNCER DE MAMA/genética. 2. MARCADORES GENÉTICOS. 3. METASTASE. 4. ANALISES *MICROARRAY*.

EPÍGRAFE

“Não existe verdadeira inteligência sem bondade.”

Ludwig van Beethoven

DEDICATÓRIA

Aos meus pais José e Clarice
e irmãos Karolina e Diogo

AGRADECIMENTOS

Primeiramente agradeço de coração aos meus amados pais José e Clarice, que sempre acreditaram e acreditam em mim e estarão ao meu lado sempre e são responsáveis por tudo que sou hoje.

Meus queridos irmãos, Karolina minha irmã, amiga e até mãezona fomos à família um do outro durante um bom tempo, obrigado pelo carinho, pela amizade e por muitas vezes cuidar do mocinho aqui, você é muito importante pra mim, meu irmão Diogo, como dizem “parente não se escolhe”, mas amigo sim e você sempre será meu eterno amigo, meu parceiro e mesmo distante vamos estar juntos.

Um agradecimento muito especial a minha orientadora Dra. Dirce Dirce Maria Carraro, por abrir as portas pra um aluno de iniciação científica que mal sabia o que era Biologia Molecular, obrigado pela oportunidade, pela paciência, por acreditar em mim e por tudo que me ensinou e me fez crescer, obrigado pela chance que me proporcionou de conviver com pessoas tão especiais nesses quase 4 anos.

Aos membros do antigo LGEA e atual BIMOCA, ao meu grande amigo Aderbal que sempre esteve comigo desde a graduação, muita amizade, muito companheirismo, obrigado pela disposição a me ouvir a dar risada do meu mau humor e me fazer rir também, sei que posso contar sempre com você. Cristina minha marmotinha, não sabe o quanto é importante pra mim cada abraço, cada sorriso seu que recebia a cada dia que chegava ao laboratório, você é uma amiga que nunca quero ficar longe, tem um coração do tamanho do mundo. Nádia minha amiga queria, quem me ensinou muita coisa do que aprendi no laboratório, muita paciência e um coração de ouro, aprendi muito com você e quero ter você do meu lado sempre. Sabrina (sassá), uma amiga daquelas de tirar o fôlego literalmente, é muito bom ter seu carinho, é muito bom ter sua amizade, sentirei falta de começar o dia com aquele abraço, você é um grande exemplo pra mim. Mariana, sempre paciente, prestativa, uma aluna imprescindível a qualquer laboratório,

obrigado pela amizade. Elen, uma amiga querida, sempre disposta a ajudar, obrigado pelas lâminas enquadradas, pela ajuda na reta final, te adoro muito. Eloisa, minha piriri, uma pessoinha que adoro demais, com um coraçãozinho gigante. Louise, sempre prestativa, me ajudou sempre que precisei, obrigado pela grande ajuda, tenho um grande carinho por você. Mev, minha peruana mais querida de todas. Jane, ex-integrante do LGEA, faz muita falta, foi um prazer ter te conhecido, guardo muitas lembranças boas. E em geral a todos os ex-integrantes e atuais do BIMOCA, Eliza, Paulo (meso), Fábio, Gustavo, Felipe, Vera, Adriana, Laura, Jamila, Reimar, Leandra, Alexandre, Luana, Rafael, Waleska.

Gostaria de agradecer alguns integrantes do departamento de Anatomia patológica. Dr. Fernando Soares, por disponibilizar vários dados para realização do projeto, por ceder todas as amostras parafinadas. Dra. Cynthia, obrigado pela ajuda, pelo carinho desde os tempos da Iniciação científica. Dra. Renata obrigado pela ajuda nas microdissecções a laser, pelo bom humor de sempre, foi muito bom trabalhar com você. Severino, Myuki e Carlinhos, que me ajudaram no corte das lâminas, sempre prestativos.

Gostaria de agradecer a Luciana e a Ana da pós, sempre disponíveis, ajudando, cobrando nos prazos, são muito profissionais.

A biblioteca, pelo excelente atendimento, pela atenção, agradeço a todos os membros, de coração e principalmente a Rose, Francine e Suely.

Agradeço a Dr. Helena Brentani, e o Luiz Paulo pela ajuda com a Bioinformática.

Agradeço ao Dr. Jordão e o Lucas do IME, por todas as análises matemáticas e estatísticas, Lucas muito obrigado pela paciência, sempre prestativo, foi muito bom trabalhar com você.

Agradeço a todos os meus amigos Milena, André, Jorge, Xurupito, Beakman, Priscila, Ixá (minha parcerinha querida), Adriana (grande amiga), Carol, Renato, Danielle, Flávia, vocês estão sempre comigo.

Finalmente agradeço a Deus, obrigado, obrigado de coração.

RESUMO

Saraiva TF. **Identificação de marcadores moleculares em câncer de mama relacionados à progressão e metástase.** São Paulo; 2008. [Dissertação de Mestrado-Fundação Antônio Prudente]

O câncer de mama é uma das neoplasias de maior incidência no mundo, afetando uma em cada oito. Dois tipos histológicos de câncer de mama, ductal e lobular, somam aproximadamente 90% dos casos, O carcinoma ductal é o mais freqüente (80%). O câncer de mama é uma doença heterogênea e pacientes com o mesmo diagnóstico e prognóstico clínico podem apresentar diferentes evoluções da doença. Essa diferença ocorre, pela limitação da atual classificação, baseada principalmente na morfologia e em alguns marcadores moleculares, que agrupam tumores molecularmente distintos em um mesmo grupo. Nesse estudo, para identificar marcadores moleculares relacionados à progressão e metástase em câncer de mama, nós analisamos o perfil de expressão gênica de tumores primários de carcinoma ductal. As amostras foram divididas em dois grupos, amostras com e sem comprometimento de linfonodo. Para se obter uma população celular mais homogênea e um perfil de expressão mais fidedigno, as amostras foram microdissecadas a laser. Doze amostras com e 17 sem comprometimento linfonodal foram microdissecadas a laser, o RNA foi extraído, tratado com DNase e amplificado. Uma plataforma customizada de cDNA *microarray* contendo 1808 genes foi utilizada, sendo desses, trezentos e quarenta e nove genes pertencentes a vias de sinalização como a PI3K, Wnt e processo de transição epitélio-mesênquima (EMT), envolvidas na diferenciação embrionária, invasão e progressão do câncer. Na comparação do perfil de expressão gênica entre os dois grupos, nós identificamos 67 genes diferencialmente expressos ($p < 0,05$), dos quais 8 (*CTNNB1*, *DAAM2*, *SKP1A*, *WNT10B*, *PLCD1*, *PLCB4*, *DAAM1*, *KRT19*) pertencem a uma das três vias de interesse. *CTNNB1*, *DAAM2*, *SKP1A* foram mais expressos e

SKP1A, *WNT10B*, *PLCD1*, *PLCB4*, *DAAM1* foram menos expressos no grupo com comprometimento de linfonodo. Além disso, nós analisamos a correlação entre pares de genes nos dois grupos de amostras e identificamos pares de genes com diferença de correlação significativa. Essa perda de correlação entre os dois grupos pode estar relacionada ao processo de invasão linfonodal. Esse estudo identificou alguns genes candidatos, potencialmente envolvidos no processo de invasão linfonodal no carcinoma ductal de mama e devem ser investigados mais profundamente.

SUMMARY

Saraiva TF. **[Identification of molecular markers in breast cancer related to progression and metastasis]**. São Paulo; 2008. [Dissertação de Mestrado-Fundação Antônio Prudente]

Breast cancer is one of the most incident neoplasia in the world, affecting one in eight women. Two histological types of breast cancer, ductal and lobular, represent approximately 90% of all cases. Ductal carcinoma is the most frequent (80%). Breast cancer is a heterogeneous disease and patients with the same diagnostic and clinical prognostic profile can have markedly different clinical outcomes. This difference is due to the limitation of the current classification based mainly on morphology and use of some molecular markers, which groups molecularly distinct diseases into the same group. In this study, to identify molecular markers related to progression and metastasis in breast tumor, we examined gene expression profile of primary ductal carcinoma samples. The samples were divided in two groups, samples without and with lymph node metastases. Laser microdissection was used to obtain homogenous cell population and therefore a cellular-based gene expression profile. Twelve samples without and 17 with lymph-node metastases were laser microdissected, RNA was extracted, DNase treated and amplified. A customized cDNA microarray platform containing 1808 genes was used. Three hundred forty nine of the genes belonged to cancer-related pathways such as PI3K, WNT signaling pathway and Epithelial-mesenchymal transition (EMT). These processes are involved in embryonic differentiation and also in tumors invasion and progression. Comparison of the gene expression profiles of the 2 groups, identified 67 genes as differentially expressed ($p < 0,05$). Eight of them (*CTNNB1*, *DAAM2*, *SKP1A*, *WNT10B*, *PLCD1*, *PLCB4*, *DAAM1*, *KRT19*), belong to one of the 3 pathways. *CTNNB1*, *DAAM2*, *SKP1A* were up and *SKP1A*, *WNT10B*, *PLCD1*, *PLCB4*, *DAAM1* were down regulated in the group with lymph node

metastases. Moreover, we analyzed the correlation between pairs of genes in the 2 sample groups and identified the gene pairs with significant difference in the correlation values between them. Some gene pairs showed loss of correlation in the 2 groups, which may be related to the lymph node invasion process. This study identified some candidate genes potentially involved in lymph node invasion process in ductal carcinoma of the breast, that have to be deeper investigated.

LISTA DE FIGURAS

Figura 1	Esquema representando a mama normal.....	2
Figura 2	Ilustração da localização do IDC e do DCIS.....	3
Figura 3	Esquema demonstrando como ocorre a metástase.....	9
Figura 4	Esquema demonstrando como é medida a distância entre os pares de genes na métrica de correlação.....	40
Figura 5	Gel de Poliacrilamida 8%. RT-PCR do gene GAPDH. Amplificado de 292 pb. M- Marcador de 100pb, 1-9 amostras tumorais extraídas por LCM, C+: controle positivo, No: controle negativo.....	47
Figura 6	Gel de Agarose 1%. M- Marcador de 1000pb, 1, 2, 3, 5, 8 e 9 aRNA de amostras tumorais microdissecadas a laser...	48
Figura 7	Representação das lâminas 4.8K, mostrando respectivamente a controle, swap e composite.....	49
Figura 8	Matrix Gel (Bionalyzer). L- Ladder, 1-10- RNA total de amostras tumorais microdissecadas manualmente, 11- C+.....	51
Figura 9	Gel de Poliacrilamida 8%. RT-PCR do gene GAPDH. Amplificado de 292 pb. M- Marcador de 100pb, 1-10 amostras tumorais extraídas por MM, C+: controle positivo, No: controle negativo.....	52
Figura 10	Gel de Agarose 1%. M- Marcador de 1000pb, 6, 7, 8, 9 e 10 aRNA de amostras tumorais extraídas por MM.....	52
Figura 11	RNA total analisado no Bioanalyzer. L- Ladder, 1- 6N0, 2- 4N1, 3- 5N1, 4- 14N0, 5- 11N1, 6- 15N0, 7- 1N1, 8- 4N0, 9- 12N1, 10- 25N1 e 11- Referência.....	59
Figura 12	aRNA analisado no Bioanalyzer. 1- 1N0, 2- 4N0, 3- 6N0, 4- 20N0, 5- 14N0, 6- 15N0, 7- 24N0, 8- 19N0, 9- 41N0, 10- 42N0, 11- 45N0, 12- Referência.....	61
Figura 13	Representação da lâmina 2.3 Marcação da amostra com Cy3 (verde), referência com Cy5 e composite.....	63

Figura 14	Dendograma (correlação- controle-swap) após normalização por LOWESS.....	64
Figura 15	Cluster (método SOM) dos genes diferencialmente expressos ($p \leq 0,05$).....	67
Figura 16	Rede de relevância para genes pertencentes ao processo EMT.....	77
Figura 17	Redes de relevância para a via Wnt.....	78
Figura 18	Redes de relevância para a via IPI3K.....	79
Figura 19	Redes de relevância para EMT e vias WNT e IPI3K.....	80
Figura 20	Scatter-plot da correlação entre os pares de genes CDS2 e SPARC e entre os pares de genes GSK3 β e SFRP1 no grupos N0 e N1.....	81

LISTA DE TABELAS

Tabela 1	Características clínico-patológicas das pacientes - Amostras FFPET.....	45
Tabela 2	Resultados obtidos nos experimentos utilizando as 10 amostras Seleccionadas FFPET microdissecadas a laser e amplificadas com o Kit Paradise.....	50
Tabela 3	Resultados obtidos nos experimentos utilizando as 10 amostras Seleccionadas FFPET microdissecadas manualmente e amplificadas com o protocolo baseado em Saraiva et al. (2006).....	54
Tabela 4	Características clínico-patológicas das amostras congeladas das pacientes sem comprometimento de linfonodo.....	57
Tabela 5	Características clínico-patológicas das amostras congeladas das pacientes com comprometimento linfonodal. Meta-amostras que apresentaram metástase à distância.....	58
Tabela 6	Genes diferencialmente expressos ($p \leq 0,05$). Comparação entre o grupo N1 com N0, valores positivos de <i>fold</i> indicam maior expressão do gene no grupo N1.....	66

LISTAS DE ABREVIATURAS

Akt	v-akt murine thymoma viral oncogene homolog
ALDH1A1	aldehyde dehydrogenase 1 family, member A1
APC	adenomatous polyposis coli
APOD	apolipoprotein D
Arna	RNA amplificado
ASNS	asparagine synthetase
AZGP1	alpha-2-glycoprotein 1, zinc-binding
BAD	BCL2-antagonist of cell death
BDNF	brain-derived neurotrophic factor
BMP7	bone morphogenetic protein 7
C1S	complement component 1, s subcomponent
CACYBP	calcyclin binding protein
CALML3	calmodulin-like 3
CAMK2G	calcium/calmodulin-dependent protein kinase (CaM kinase) II gamma
CDC7	cell division cycle 7 homolog
CDH13	cadherin 13, H-cadherin
CDK2	cyclin-dependent kinase 2
CDK4	cyclin-dependent kinase 4
CDS2	CDP-diacylglycerol synthase (phosphatidate cytidyltransferase) 2
C-erbB-2	erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog
CHMP4A	chromatin modifying protein 4A
COL5A1	collagen, type V, alpha 1
CR2	complement component (3d/Epstein Barr virus) receptor 2
CSNK1A1	casein kinase 1, alpha 1
CSNK1G2	casein kinase 1, gamma 2
CTNNB1	catenin (cadherin-associated protein), beta 1

DAAM1	dishevelled associated activator of morphogenesis 1
DAAM2	dishevelled associated activator of morphogenesis 2
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein
DCIS	carcinoma ductal in situ
DERP6	chromosome 17 open reading frame 81
DLL1	delta-like 1
DNAJB1	DnaJ (Hsp40) homolog, subfamily B, member 1
DST	dystonin
DYRK1A	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A
DYRK4	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 4
ECM	Extracellular matrix
EIF2AK3	eukaryotic translation initiation factor 2-alpha kinase 3
EMT	epithelial-mesenchymal transition
EPHB6	EPH receptor B6
ER	receptor de estrógeno
EYA4	eyes absent homolog 4
FFPET	Formalin-fixed paraffin-embedded tissues
FLJ20257	methylphosphate capping enzyme
FLJ21439	spastic paraplegia 11
FN3K	fructosamine 3 kinase
FRY	furry homolog
FZD2	frizzled homolog 2
FZD9	frizzled homolog 9
GAK	cyclin G associated kinase
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GNAS	complex locus
GNB1	guanine nucleotide binding protein (G protein), beta polypeptide 1
GPNMB	glycoprotein (transmembrane) nmb
GSK3B	glycogen synthase kinase 3 beta
GSTM3	glutathione S-transferase M3

HE	hematoxilina-eosina
IDC	carcinoma ductal invasivo
ILC	carcinoma lobular invasivo
ITPR1	inositol 1,4,5-triphosphate receptor, type 1
KEGG	Kyoto Encyclopedia of Genes and Genomes
KLK3	kallikrein-related peptidase 3
KRT19	keratin 19
KTN1	kinectin 1 (kinesin receptor)
LCIS	carcinoma lobular in situ
LCM	laser capture microdissection
LEF1	lymphoid enhancer-binding factor 1
LRP6	low density lipoprotein receptor-related protein 6
LRP6	low density lipoprotein receptor-related protein 6
MAP3K1	mitogen-activated protein kinase kinase kinase 1
MM	microdissecção manual
MOG	myelin oligodendrocyte glycoprotein
mRNA	RNA mensageiro
mTOR	FK506 binding protein 12-rapamycin associated protein
MYC	v-myc myelocytomatosis viral oncogene homolog
NCAM2	neural cell adhesion molecule 2
NFATC3	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 3
NKD1	naked cuticle homolog 1
NKD2	naked cuticle homolog 2
NPNT	nephronectin
P21	cyclin-dependent kinase inhibitor 1A
P53	tumor protein p53
PAX1	paired box 1
PCBP1	poly(rC) binding protein 1
PCDH11Y	protocadherin 11 Y-linked
PCTK2	PCTAIRE protein kinase 2
PHACTR2	phosphatase and actin regulator 2

PHB2	prohibitin 2
PHYHIP	phytanoyl-CoA 2-hydroxylase interacting protein
PI3K	phosphatidylinositol 3-kinase
PIK3R4	phosphoinositide-3-kinase, regulatory subunit 4
PIP	prolactin-induced protein
PIP2	Phosphatidylinositol bisphosphate
PIP3	Phosphatidylinositol trisphosphate
PIP4K2B	phosphatidylinositol-5-phosphate 4-kinase, type II, beta
PLCB4	phospholipase C, beta 4
PLCD1	phospholipase C, delta 1
PLCG1	phospholipase C, gamma 1
PORCN	porcupine homolog
PPP1CA	protein phosphatase 1, catalytic subunit, alpha isoform
PPP2R2B	protein phosphatase 2 (formerly 2A), regulatory subunit B, beta isoform
PR	receptor de progesterona
PRKACB	protein kinase, cAMP-dependent, catalytic, beta
PRKCBP1	zinc finger, MYND-type containing 8
PRKX	protein kinase, X-linked
PSMB5	proteasome (prosome, macropain) subunit, beta type, 5
PSMD13	proteasome (prosome, macropain) 26S subunit, non-ATPase, 13
RAC3	ras-related C3 botulinum toxin substrate 3
RAI17	zinc finger, MIZ-type containing 1
RGS4	regulator of G-protein signaling 4
RIN	RNA Integrity Number
ROCK1	Rho-associated, coiled-coil containing protein kinase 1
RPL32	ribosomal protein L32
RPS9	30S ribosomal protein S9
S100A6	S100 calcium binding protein A6
SALL4	sal-like 4 (Drosophila)
SDC1	syndecan 1

SDHB	succinate dehydrogenase complex, subunit B, iron sulfur (lp)
SFRP1	secreted frizzled-related protein 1
SIP1	survival of motor neuron protein interacting protein 1
SKP1	S-phase kinase-associated protein 1
SLC39A8	solute carrier family 39 (zinc transporter), member 8
SLN	linfonodo sentinela
SMAD2	SMAD family member 2
SNRPB	small nuclear ribonucleoprotein polypeptides B and B1
SOM	self-organization map
SOX17	SRY (sex determining region Y)-box 17
SPARC	secreted protein, acidic, cysteine-rich
SPATA11	alkB, alkylation repair homolog 7
TBL1XR1	transducin (beta)-like 1 X-linked receptor 1
TBL1XR1	transducin (beta)-like 1 X-linked receptor 1
TCF	T cell factor
THRB	thyroid hormone receptor, beta (erythroblastic leukemia viral (v-erb-a) oncogene homolog 2
TLE3	transducin-like enhancer of split 3
TMEM98	transmembrane protein 98
TNRC15	GYF protein 2
TSPAN1	tetraspanin 1
TUBB	tubulin, beta
VWF	von Willebrand factor
WASPIP	WAS/WASL interacting protein family, member 1
WNT10B	wingless-type MMTV integration site family, member 10B
ZBTB4	zinc finger and BTB domain containing 4

ÍNDICE

1	INTRODUÇÃO	1
1.1	Metástase e câncer de mama	8
1.1.1	Vias de sinalização que atuam nos processos de diferenciação e invasão tumoral	9
1.2	Análise global de expressão gênica e sua aplicação em câncer de mama	16
2	JUSTIFICATIVA PARA O ESTUDO E DELINEAMENTO EXPERIMENTAL ADOTADO	20
3	OBJETIVO	24
3.1	Objetivo geral	24
3.2	Objetivo específico	24
4	MATERIAL E MÉTODOS	25
4.1	Seleção das amostras	25
4.2	Microdissecção das amostras	25
4.2.1	Procedimento geral para Microdissecção a Laser (LCM)	25
4.2.2	Microdissecção Manual (MM)	27
4.3	Extração de RNA total das amostras	28
4.3.1	Extração de RNA total de Amostras Parafinadas (LCM)	28
4.3.2	Extração de RNA total de amostras congeladas (LCM)	28
4.3.3	Extração de RNA total de amostras parafinadas (MM)	28
4.4	Avaliação da integridade do RNA total	28
4.4.1	Avaliação por RT-PCR dos genes <i>NOTCH2</i> e <i>GAPDH</i>	28
4.4.2	Avaliação pelo 2100 Bioanalyzer (Agilent)	29
4.5	Amplificação do RNA mensageiro (mRNA)	30
4.5.1	Amplificação do RNA Mensageiro de amostras FFPET Microdissecadas a laser (LCM)	30

4.5.2	Amplificação do mRNA de amostras FFPET Microdissecadas Manualmente e da amostra Referência (linhagem HB4A)	30
4.5.3	Amplificação do RNA mensageiro de amostras congeladas (LCM)	33
4.6	Avaliação da qualidade do RNA amplificado (aRNA)	33
4.6.1	Avaliação da Qualidade do RNA Amplificado das Amostras FFPET	33
4.6.2	Avaliação da Qualidade do RNA Amplificado das Amostras Congeladas	34
4.7	Marcação	34
4.8	Hibridização, lavagem e captura de imagem	36
4.9	Plataformas de cDNA <i>microarray</i>	37
4.9.1	Plataforma 4.8K	37
4.9.2	Plataforma 2.3K de vias metabólicas	37
4.10	Análise dos dados	38
4.11	Análises dos Dados	39
4.11.1	Dendograma (método de ligação completa e métrica de correlação)	39
4.11.2	Genes diferencialmente expressos	41
4.11.3	Método de clusterização SOM (<i>self-organization map</i>).	41
4.11.4	Construção e análise de redes de relevância	42
5	RESULTADOS e DISCUSSÃO	43
5.1	Dados referentes às amostras parafinadas do arquivo do Hospital A.C. Camargo a amostras FFPET	43
5.1.1	Seleção das amostras FFPET	43
5.1.2	Extração de RNA total e amplificação das amostras FFPET	46
5.2	Dados referentes às amostras congeladas do banco de tumores do Hospital A.C. Camargo	55
5.2.1	Seleção das amostras congelada	55
5.2.2	Processamento das amostras	59
5.2.3	Marcação e Hibridização	62
5.2.4	Análise matemática e estatística	63

6	CONCLUSÃO	82
7	REFERÊNCIAS BIBLIOGRÁFICAS	84

ANEXOS

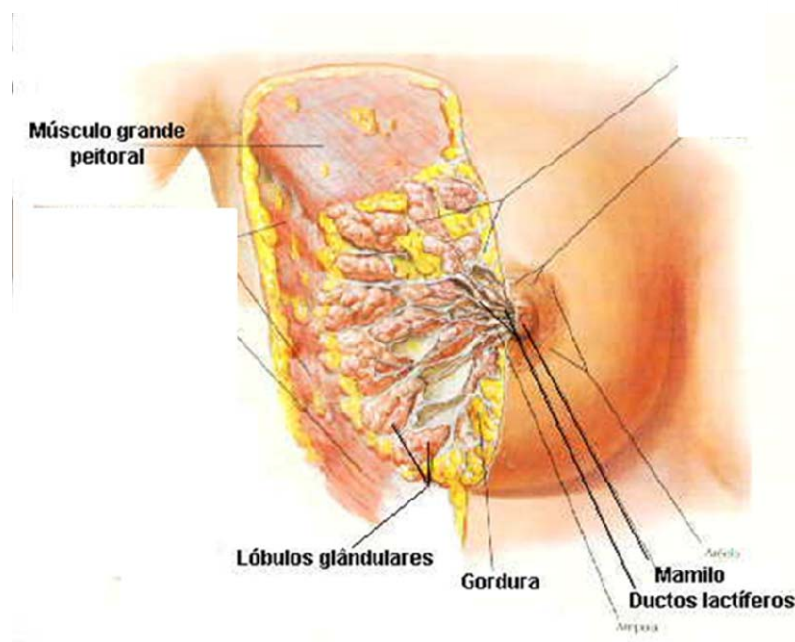
Anexo 1 Saraiva TF, Castro NP, Pineda PHB, et al. Effects of Oligo dT-T7 RNA Primer in RNA Amplification from Paraffin-Embedded Tissue for *Microarray* Experiments. **Appl Cancer Res** 2006; 26:14-20.

Anexo 2 Tabela com os Genes das vias PI3K, WNT e EMT, na lâmina 2.3K.

1 INTRODUÇÃO

Dentre os diversos tipos de câncer, o câncer de mama está entre as neoplasias de maior incidência mundial, afetando em média uma em cada oito mulheres. Devido a sua alta frequência, é o mais temido, sobretudo pelos seus efeitos psicológicos. No Brasil, o câncer de mama é o que mais causa mortes entre as mulheres, dos 466.700 novos casos de câncer com previsão de diagnóstico em 2008, o câncer de mama, será responsável por 49.3000 (Ministério da Saúde 2008).

A mama feminina normal adulta apresenta uma estrutura arborescente e ramificada, sendo composta por elementos epiteliais e estromais. O componente epitelial da mama é organizado em lóbulos que são compostos por grupos de ácinos, estes se conectam em ductos terminais (unidade ducto-lobular terminal) abertura é na papila. Já o estroma é composto de quantidade variável de tecido adiposo e tecido conectivo fibroso, que consiste a maior parte do volume da mama no estado não lactente (Figura 1).



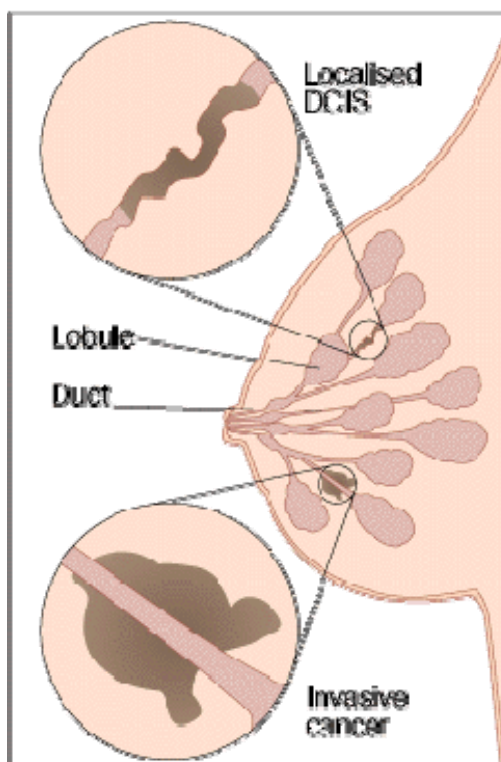
Fonte – <http://members.tripod.com/Reivax/cancerpg12.htm>

Figura 1 - Esquema representando a mama normal.

Quanto ao aspecto histológico, o epitélio ductal normal da mama consiste de camada interna (luminal) de células epiteliais cubóides e camada externa de células mioepiteliais (apoiada na membrana basal do ducto).

Dentre os tipos histológicos de neoplasia de mama, o carcinoma ductal e lobular somam aproximadamente 90% dos casos, sendo o ductal o mais freqüente, somando 80% dos casos. O carcinoma *in situ* (não-invasivo) é o resultado de uma proliferação de células neoplásicas dentro dos ductos ou lóbulos. Há dois tipos de câncer de mama não-invasivo, o carcinoma ductal *in situ* (DCIS) e o lobular *in situ* (LCIS), caracterizado pela divisão anormal das células ductais ou lobulares respectivamente, permanecendo dentro dos ductos e/ou lóbulos. Estes dois tipos de lesões diferem nas suas características clínicas, na morfologia e comportamento biológico. Já o

carcinoma invasivo ocorre quando as células se espalham além dos ductos (IDC) e lóbulos (ILC), com ruptura da membrana basal, apresentando invasão de células tumorais em tecidos adjacentes, sangue e linfa. Uma representação do DCIS e IDC pode ser visualizada na Figura 2.



Fonte: www.e-radiography.net/radpath/d/dcis.htm

Figura 2 - Ilustração da localização do IDC e do DCIS.

Antes da utilização da mamografia como método de prevenção, o DCIS era incomum, contando com aproximadamente 2-3 % de todos os cânceres de mama (ERNSTER et al. 1996). Com o incremento de sua utilização a taxa de detecção aumentou 300% e o DCIS conta atualmente com cerca de 20-30% dos casos de câncer de mama (ERNSTER e BARCLAY 1997).

A remoção do tumor seguido de radioterapia e hormonioterapia é o tratamento empregado em DCIS, sendo curativo em 98% dos casos e apresentando progressão em 2% (NAKHLIS e MORROW 2003). Já para as mulheres que apresentam o tipo histológico IDC (invasivo), o que é levado em consideração na escolha do tratamento, são os fatores prognósticos (idade da paciente, tamanho do tumor, status dos linfonodos axilares, grau patológico e status hormonal) juntamente com outros marcadores moleculares recentes que foram incluídos na rotina clínica (NAKHLIS e MORROW 2003; LEONARD e SWAIN 2004).

O *status* linfonodal certamente representa o mais importante fator prognóstico que guia a seleção de pacientes elegíveis à terapia adjuvante, embora pacientes com câncer de mama linfonodo negativo com pelo menos um fator de risco adicional (por exemplo, T>3, *status* ER negativo ou alto grau) são indicados à terapia adjuvante (CARTER et al. 1989, GOLDBIRSCH et al. 1995; ZUJEWSKI e LIU 1998; GOLDBIRSCH et al. 2001). Assim, a quimioterapia sistêmica adjuvante fica reservada para pacientes linfonodo-negativos de alto risco e todos os pacientes linfonodo-positivos.

A biópsia de linfonodo sentinela (SLNB) é um procedimento cirúrgico comum (MAGGARD et al. 2005) e uma técnica altamente acurada (VERONESI et al. 2003) e segura para o estadiamento do câncer de mama. É um procedimento minimamente invasivo, e utilizado para identificar o primeiro linfonodo que drena e está em direta comunicação com o tumor primário, permitindo uma linfadenectomia seletiva. A presença de

metástases no linfonodo sentinela (SLN) é um dos mais importantes fatores prognósticos para a doença, uma vez que indica que células tumorais intravazaram os vasos linfáticos e poderão migrar através deles para colonizar sítios metastáticos distantes.

O Status hormonal também é um fator prognóstico muito importante, o estrógeno atua no crescimento e desenvolvimento das glândulas mamárias e tem grande contribuição na carcinogênese do câncer de mama (HENDERSON et al. 1988). O estrógeno atua na proliferação celular mediada por receptores intracelular (receptor de estrógeno - ER), e quando ativado, atua como fator de transcrição, regulando a expressão de vários genes. Em geral tumores ER negativos estão associados a uma menor sobrevida, quando comparado com ER positivo. Portanto ER é utilizado como marcador prognóstico e na predição de resposta à hormonioterapia (BLOOM et al. 1980; OSBORNE et al. 1980; BEZWODA et al. 1991). Pacientes ER negativo, raramente respondem à terapia endócrina, o que mostra a correlação entre o nível de expressão de ER e a resposta a terapia (BEZWODA et al. 1991).

Outros marcadores moleculares também são utilizados como fatores prognósticos em câncer de mama, como o receptor de progesterona (PR). Os estrógenos são conhecidos por estimular não apenas a proliferação celular como também pela indução de síntese de proteínas específicas, tais como receptor de progesterona. Portanto, a determinação do receptor de progesterona aumenta a sensibilidade do prognóstico e na previsão de resposta a hormonioterapia (PAYNE et al. 2008).

Outro marcador molecular utilizado na clínica é o oncogene HER-2/neu (c-erbB-2), membro da família do receptor de fator de crescimento epidérmico e encontra-se amplificado em aproximadamente 25% dos casos de câncer de mama, inúmeros estudos da literatura indicam o c-erbB-2 como fator prognóstico independente e preditivo na resposta à quimioterapia adjuvante baseada no uso da adriamicina. Além disso, com a introdução do anticorpo monoclonal humanizado Herceptin®, como nova modalidade terapêutica em câncer de mama, o status do c-erbB-2 tornou-se muito importante na avaliação terapêutica global dessa doença (PAYNE et al. 2008).

Outra proteína que apresenta relação com o prognóstico é a p53, codificada pelo gene *P53* responsável pelo controle da proliferação celular. O descontrole de sua função fisiológica, através de mutações, pode levar à perda do controle de proliferação. Mutações de p53 podem levar ao surgimento de proteínas com alterações conformacionais funcionalmente defeituosas, que têm uma meia vida maior em relação à proteína p53 selvagem (YAN et al. 1996). A proteína p53 selvagem acumulada na célula, promove a parada do ciclo celular na fase G1. Esta é uma pausa reversível e permite que a célula repare o DNA. Se, por algum motivo, este mecanismo falha, a p53 leva a célula à apoptose, para que células mutadas não se dividam (COX e WALKER 1996). A perda de p53 faz com que muitas células repliquem o DNA danificado, incorporando essas mutações ao genoma (COTRAN et al. 1994). Embora uma mutação apenas não seja suficiente para transformar uma célula, a perda da p53 predispõe as células a

mutações adicionais e finalmente à transformação maligna. Alterações estruturais acumulam-se no núcleo (LAVIEILLE et al. 1998) levando à reatividade imuno-histoquímica da proteína p53 (WONG et al. 1998). Aproximadamente 30-40% dos cânceres primários da mama expressam proteína p53, sendo esses casos, segundo alguns estudos, de comportamento clínico mais agressivo (KUMAR et al. 2007). No entanto esse marcador não é utilizado para avaliação clínica de rotina.

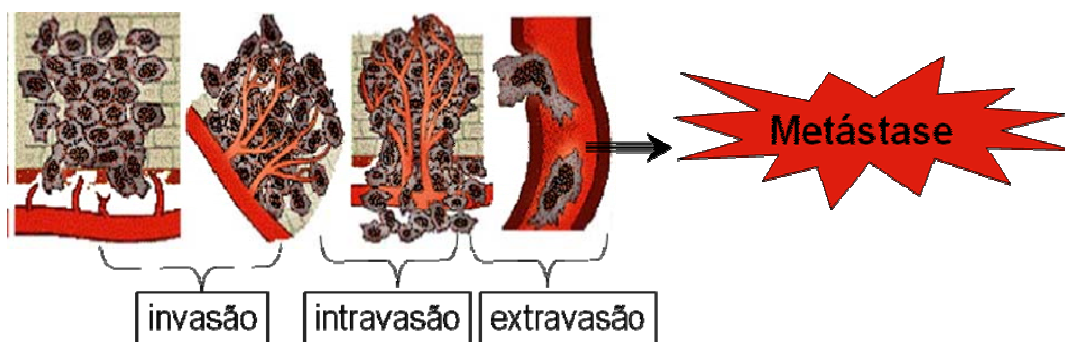
Além desses, tem sido sugeridos dois marcadores: uPA (*Urokinase plasminogen activator*) e PAI-1 (*plasminogen activator inhibitor1*) para seleção de pacientes com câncer de mama linfonodo-negativo que não requerem quimioterapia adjuvante (MOLINA et al. 2005). uPA é uma protease serina implicada na proliferação celular do câncer, invasão e metástase, e PAI-1 é um inibidor endógeno de uPA, porém paradoxalmente está também envolvido na progressão do tumor (ANDREASEN et al. 1997). Esses marcadores ainda não são utilizados na rotina clínica.

Apesar do emprego de marcadores moleculares na rotina clínica, muitas pacientes com mesmo diagnóstico e prognóstico clínico podem apresentar resultados muito diferentes tanto em nível de resposta a tratamento, como recidiva e surgimento de metástase. Isso ocorre, porque a classificação atual baseada nos marcadores utilizados é limitada e classifica pacientes com doenças molecularmente heterogêneas em uma única entidade (NAKHLIS e MORROW 2003), levando pacientes com perfis moleculares distintos terem a mesma classificação por falta de outros indicadores (SORLIE et al. 2001; SOTIRIOU et al. 2003). Assim, uma das

áreas mais importantes na pesquisa contra o câncer de mama é a identificação de novos marcadores moleculares tumorais que possibilitem o desenvolvimento de métodos diagnósticos para detecção precoce, prognósticos mais precisos e também para tratamentos mais individuais, específicos e eficientes.

1.1 METÁSTASE E CÂNCER DE MAMA

A metástase é a principal causa de morte em câncer de mama. Alguns estudos sugerem que ocorre uma seleção clonal das células tumorais e a célula que adquire capacidade invasora é responsável pelo estabelecimento do processo metastático (FIDLER e HARTE 1982). O processo metastático é dividido basicamente em 4 etapas (FIDLER 2003). No primeiro passo denominado invasão as células epiteliais neoplásicas perdem a adesão célula-célula e ganham motilidade o que permite que invadam tecidos adjacentes. Durante o segundo passo, intravasão, as células tumorais penetram através do endotélio sangüíneo ou vasos linfáticos entrando na circulação sistêmica. Apenas algumas células tumorais circulantes são capazes de sobreviver à passagem através da circulação. Algumas destas sobreviventes se dirigem para completarem o próximo passo, extravasão, quando elas extravasam através do endotélio capilar. Finalmente, em um novo ambiente, um pequeno grupo de células metastizantes obtém sucesso na proliferação e crescimento, concluindo a micrometástase, resultando em tumores malignos secundários (Figura 3).



Fonte: GIMENEZ (2008)

Figura 3 - Esquema em 4 etapas demonstrando como ocorre a metástase.

1.1.1 Vias de sinalização que atuam nos processos de diferenciação e invasão tumoral

Para que ocorra o acometimento linfonodal por células tumorais, a célula tumoral tem que gerar alterações moleculares suficientes para se tornar apta a concluir as duas primeiras etapas, invasão e intravasão. Para a conclusão dessas duas etapas, várias propriedades celulares estão envolvidas.

Alterações nas propriedades de adesão das células tumorais estão associadas com progressão tumoral e formação de metástases e freqüentemente, coincidem com a perda da adesão celular mediada por E-caderina (*epithelial cadherin*), alteração na morfogênese e arquitetura anormais do tecido epitelial, perda da polaridade celular e inibição por contato, crescimento desregulado e invasão de tecidos adjacentes (CHRISTOFORI et al. 1999; HANDSCHUH et al. 1999). O processo denominado de transição epitélio–mesênquima (EMT, *epithelial-mesenchymal transition*), que envolve a diminuição da expressão das

moléculas de E-caderina tem sido relatado nas primeiras etapas da invasão e metástases de células tumorais (BOYER et al. 2000). Esse processo ocorre normalmente durante o desenvolvimento embrionário, o qual é caracterizado por uma diminuição da função da E-caderina que marca o surgimento de um programa complexo no qual as células epiteliais adotam um fenótipo fibroblasto-*like* e apresentam atividade invasiva, denominado EMT (CONACCII-SORREL et al. 2002; NELSON e NUSSE 2004). Este evento é essencial durante os movimentos de gastrulação e formação da crista neural no desenvolvimento embrionário (SHOOK e KELLER 2003) e recentemente tem sido associado com as mudanças moleculares observadas nas células neoplásicas metastáticas mais agressivas (THIERY 2003).

Esse processo tem sido investigado em vários tipos tumorais, incluindo câncer de mama (ACKLAND et al. 2003; VINCENT-SALOMON e THIERY 2003). Acredita-se que a célula epitelial tumoral altere seu processo de capacidade proliferativa, adquirindo a capacidade de invadir. Além das alterações genéticas que ocorrem nas próprias células do tumor primário, que resultam na capacidade de invasão, outros fatores também estão relacionados a esse processo. Um deles é o microambiente tumoral, composto principalmente pela matriz extracelular (ECM - *Extracellular matrix*) e produzida pelos fibroblastos, tem grande influência no crescimento celular, migração, morfologia, proliferação e diferenciação (NOEL e FOIDART 1998). Interação estroma-epitélio tem um papel importante no desenvolvimento e progressão do câncer.

A β -catenina também está envolvida nas propriedades de adesão celular, ancorando a E-caderina ao citoesqueleto de actina (JAWHARI et al. 1999). Esta interação é fundamental para que ocorra uma adesão célula-célula eficiente (HIROHASHI e KANAI 2003). Além disso, participa da via de sinalização Wnt. O controle da ativação e desativação da via de sinalização Wnt é importante durante a embriogênese e para a manutenção do tecido adulto. Estudos concentram esforços na compreensão do braço canônico da via Wnt (PEIFER e POLAKIS 2000), que está altamente conservado entre as espécies.

Na célula adulta, a β -catenina está localizada na membrana e citoplasma. Essa proteína está envolvida na formação de complexos protéicos responsáveis pela adesão celular, o que impede sua transferência para o núcleo e evita a transcrição dos seus genes alvos, esse controle é realizado via fosforilação da β -catenina. A fosforilação ocorre nos resíduos de serina e treonina na porção amino-terminal dessa proteína pelo complexo formado pela APC e GSK3 β (PEIFER e POLAKIS 2000). Essa fosforilação resulta na ubiquitinação da β -catenina levando-a para a degradação nos proteossomos (ABERLE et al. 1997). Quando a β -catenina não é fosforilada, é transportada para o núcleo onde forma um complexo com a família de ligantes de DNA TCF/LEF, que é responsável pela regulação da transcrição de diversos genes, tais como o c-MYC (HE et al. 1998) e a *ciclina D1* (TETSU e MCCORMICK 1999) e alguns genes envolvidos na transição epitélio-mesênquima (EMT- *epithelial-mesenchymal transition*) atuando na ativação da transcrição de *SLUG* (CONACCI-SORREL et al. 2003), além da

estabilização da proteína Snail (YOOK et al. 2005), sendo que esses genes atuam na inibição da transcrição de E-caderina.

Alterações na regulação dessa via são reconhecidamente importantes no desenvolvimento de inúmeros tumores (MIYOSHI et al. 1998; SPARKS et al. 1998), uma vez que a β -catenina volta a ativar a transcrição de genes que deveriam permanecer silenciados na célula diferenciada. Assim a β -catenina atua nos processos de adesão celular e de transcrição, ambos importantes para o desenvolvimento de um órgão e para a manutenção celular, portanto, uma desregulação da β -catenina pode levar a alteração na regulação da via Wnt e por sua vez levar ao desenvolvimento do câncer (WALTZER e BIENZ 1999; MOON et al. 2002; HARRIS e PEIFER 2005).

As proteínas Wnts são sinalizadoras da ativação ou inibição da via Wnt. Quando as células são expostas a essas proteínas, estas se ligam ao receptor de superfície celular da família Frizzled, ativando a via. No processo de ativação, o complexo formado - WNT/FRIZZLED - antagoniza o APC-Axina bloqueando a fosforilação da β -catenina por mecanismos ainda não conhecidos, mas que requerem a participação das proteínas Dishevelled (PEIFER e POLAKIS 2000). Este é um processo altamente regulado que controla diversas funções durante o desenvolvimento como a sinalização da proliferação, adesão, diferenciação, morfologia celular e apoptose (BRENNAN e BROWN 2004). Mutações nas próprias WNTs são raramente encontradas como responsáveis pela ativação da via durante a carcinogênese, sendo essa ativação mais freqüentemente associada com mutações de APC ou de β -catenina. Esses genes foram associados com a

promoção da proliferação celular, repressão da diferenciação durante a embriogênese e carcinogênese. Estudos moleculares associaram mutações em diversos genes da via à aproximadamente 90% dos cânceres colorretal e menos freqüentemente a outros cânceres como o carcinoma hepatocelular (GILES et al. 2003).

O surgimento do processo EMT é marcado pela diminuição da função da E-caderina, as células epiteliais passam a apresentar um fenótipo fibroblasto-*like* e apresentam atividade invasiva (BOYER et al. 2000), o que tem sido recentemente associado às mudanças moleculares observadas nas células neoplásicas metastáticas mais agressivas (THIERY 2003).

Entre os vários tipos de regulação da E-caderina está a repressão transcricional causada pelos repressores SIP1 (*smad interacting protein 1*) (COMIJN et al. 2001), SNAIL e SLUG (BATLLE et al. 2000; HEMAVATHY et al. 2000; BOLOS et al. 2003), os quais se ligam às seqüências E-box de seu promotor e diminuem sua expressão, afetando a adesão célula-célula e permitindo a invasão dos tecidos. Foi demonstrado que o Snail tem um domínio semelhante ao da β -catenina sendo regulado pela fosforilação via GSK3 β . Portanto, a proteína GSK3 β (YOOK et al. 2005) é considerada o ponto de conexão dos dois processos - via de sinalização Wnt e processo EMT.

PI3K é também umas das vias fundamentais em algumas propriedades celulares, como crescimento, proliferação, sobrevivência. Alterações no gene *PI3K* estão freqüentemente envolvidas com muitos tipos

de câncer em humanos, incluído cólon, mama, cérebro, estômago e pulmão (SAMUELS e VELCULESCU 2004).

As proteínas PI3Ks são heterodímeros compostos por uma subunidade catalítica p110 e uma subunidade regulatória p85. As subunidades catalíticas do gene *PI3KCA* (gene que codifica a subunidade p110) são compostas por vários domínios modulares: o domínio catalítico lipídio-quinase, domínio helicoidal, domínio RAS e um domínio terminal NH2 que interage com a subunidade regulatória.

Um mecanismo que pode induzir a atividade de PI3K, é através da ativação de receptores de membrana que se ligam a subunidade regulatória, recrutando a subunidade catalítica em complexos subcelulares específicos. Uma vez recrutada para a membrana, a subunidade catalítica fosforila PIP2 (*Phosphatidylinositol bisphosphate*) que participa da via de sinalização de PI3K e forma PIP3 (*Phosphatidylinositol trisphosphate*), que por sua vez recruta proteínas com domínios de ligação fosfolipídicos, para membrana plasmática. Em particular, Akt (proteínas fazem parte da família de quinases serina/treonina) é recrutada a membrana e quando ativada, se encaminha para o citoplasma e núcleo onde fosforila seus alvos. Akt tem várias proteínas como alvo, incluindo mTOR, Bad, caspase 9, GSK3 β e fatores de transcrição, que vão regular múltiplas funções celulares incluindo proliferação, sobrevivência e migração (CANTLEY 2002; VIVANCO e SAWYERS 2002).

Recentes estudos sugerem que uma ativação constitutiva da via PI3K, pode ocorrer por mutação da subunidade p110 do gene *PI3KCA*.

Vários estudos mostram que essas mutações são observadas em 18% a 40% em câncer de mama (KARACAS et al. 2006), sugerindo que mutações em *PI3KCA* podem ativar essas vias.

Outros estudos mostraram a relevância da mutação de *PI3KCA* em pacientes com câncer de mama. LI et al. (2005), concluíram que a presença de qualquer mutação nesse gene é um fator independente de menor sobrevida. BARBARESCI et al. (2007) observaram a presença de mutação em 28% dos casos analisados, mas sugerem que a mutação localizada em diferentes éxons (éxons 9 e 20) pode levar a diferentes prognósticos.

Estas evidências mostram a complexidade da regulação dessas vias, destacando a necessidade de um fino controle para manutenção e funcionamento da célula no tecido ou órgão onde ela está inserida. A caracterização de vias sabidamente envolvidas no processo de diferenciação celular, e também invasão e metástase, como a via de sinalização WNT, fosfatidilinositol (PI3K) e EMT é de grande importância para se entender melhor o processo de transformação celular e invasão de tecidos adjacentes. Alterações de expressão de genes pertencentes a essas vias nas células tumorais podem ser de grande importância na compreensão do processo de instalação e progressão tumoral além de apresentar grande potencial para busca de marcadores moleculares.

1.2 ANÁLISE GLOBAL DE EXPRESSÃO GÊNICA E SUA APLICAÇÃO EM CÂNCER DE MAMA

Atualmente várias metodologias que permitem avaliação de expressão gênica em grande escala, auxiliam na busca de marcadores preditivos e de prognóstico em câncer de mama.

Microarray tem sido muito utilizado para análise global de expressão gênica em câncer (SGROI et al. 1999; VAN'T VEER et al. 2002; SOTIRIOU et al. 2003; VAN'T VEER et al. 2003; MEIRELES et al. 2004). O método consiste na hibridização de milhares de genes imobilizados em uma lâmina de vidro ou membrana de *nylon*, por amostras fluorescentes ou radioativas de cDNA, preparadas a partir de RNA de células, tecidos ou outras fontes biológicas. Essas moléculas reagem com suas seqüências complementares ordenadas na lâmina de vidro ou membrana, que emitem sinais de diferentes intensidade a partir de cada *spot* (várias moléculas de cDNA de mesma seqüência arranjadas na lâmina). A intensidade da fluorescência ou radiação emitida por cada *spot* será, então, proporcional ao número de moléculas hibridizadas que por sua vez é proporcional à expressão do gene que corresponde àquela determinada seqüência.

Para se ter uma análise mais fidedigna das células de interesse em experimentos de *microarray* é importante que se utilize amostras homogêneas, entretanto em amostras tumorais geralmente se encontram vários tipos celulares, pois existem células do sistema imunológico infiltrando o tumor, células estromais, como as células endoteliais dos vasos linfáticos e

sanguíneos, e isso pode naturalmente interferir no perfil de expressão gênica. No caso de amostras tumorais de mama, devido à natureza da lesão onde se encontram células não tumorais, é importante que se obtenha RNA de populações celulares homogêneas resultando em uma correlação mais fidedigna da célula tumoral e seu padrão de expressão gênica. A microdissecção a laser (LCM- *laser capture microdissection*) é uma metodologia capaz de aumentar a sensibilidade desses estudos moleculares, uma vez que possibilita a captura de uma população celular homogênea do tumor. Uma problemática em relação a esta técnica é a obtenção de RNA em pouca quantidade, devido ao pequeno número de células capturadas. Portanto, para obtenção de RNA de alta qualidade, cuidados extremos durante a captura, extração do RNA e também durante a metodologia de amplificação do RNA mensageiro (mRNA) são necessários, uma vez que pode ocorrer degradação do RNA durante os processos (BRENTANI et al. 2005). Procedimentos de amplificação de mRNA são requeridos, pois são capazes de gerar quantidades suficientes para garantir a reprodutibilidade e qualidade dos experimentos (WANG et al. 2000; GOMES et. al. 2003). Em síntese os métodos de amplificação são baseados na produção de cDNA dupla fita a partir de RNA total usando para síntese de primeira fita oligonucleotídeos que trazem a sequência do promotor da RNA polimerase do fago T7 e ancoram na cauda poli-A do RNAm. O cDNA dupla fita, produzido, é então submetido a transcrição *in vitro* e permite a produção de RNA anti-senso. Comparações do RNA anti-senso amplificado (aRNA) com RNA totais são realizados e mostram que o RNA total pode ser

substituído por RNA anti-senso amplificado apresentado aceitável reprodutibilidade e alta fidelidade dos dados (WANG et al. 2000; GOMES et. al. 2003).

A técnica de *microarray* tem sido amplamente utilizada na comparação do perfil de expressão gênica em câncer e tem auxiliado na melhor caracterização molecular de várias neoplasias e a partir de correlações com informações clínicas está sendo utilizada também para subclassificar tumores que apresentam característica clínicas distintas quanto a progressão da doença (SORLIE et al. 2001; VAN'T VEER et al. 2002; SOTIRIOU et al. 2003; NISHIDATE et al. 2004) ,diferentes níveis de resposta a quimioterápicos (PEROU et al. 2000) e genes envolvidos na progressão da doença (VAN'T VEER et al. 2002; FENG et al. 2007).

PEROU et al. (2000), SORLIE et al. (2001) e SOTIRIOU et al. (2003), mostraram através da análise de expressão gênica por *microarray* em câncer de mama que o status hormonal é o fator mais importante na diferença de expressão entre as amostras analisadas, ressaltando a importância de se conhecer melhor a diferença entre os subtipos de tumor.

A busca de genes preditivos de metástase através do *microarray* vem sendo muito empregada, VAN'T VEER et al. (2002) classificou 70 genes associados à um maior risco de metástase à distância em pacientes jovens (abaixo de 55 anos) com câncer de mama e sem comprometimento de linfonodo que tiveram evoluções distintas (pacientes que apresentaram metástase e pacientes que não apresentaram). O argumento para esse

estudo, é que pacientes com diagnósticos semelhantes terão diferentes evoluções da doença e tratamentos mais específicos são necessários.

WANG et al. (2005), compararam o perfil de expressão gênica através de *microarray* de amostras de 286 pacientes com câncer de mama sem comprometimento linfonodal, sem distinção de idade, status hormonal e tamanho do tumor e com diferentes evoluções da doença, sendo que nenhuma paciente recebeu tratamento adjuvante. O estudo identificou 76 genes preditores de metástase em 5 anos, que pode ser aplicado independente das características clínico-patológicas das pacientes (idade, status hormonal, tamanho do tumor e grau).

FENG et al. (2007) utilizando uma plataforma contendo 21.329 genes humanos, compararam 26 tumores primários de pacientes com carcinoma ductal invasivo de mama com seus respectivos linfonodos comprometidos. Eles identificaram 79 genes diferencialmente expressos entre o tumor primário e os linfonodos comprometidos, além disso esses genes também foram capazes de classificar tumores primários em dois grupos com diferentes prognósticos, grupo de alto risco e grupo de baixo risco. Pacientes do grupo de alto risco apresentaram 4,65 vezes mais chances de desenvolver metástase a distância dentro de 43 meses, quando comparado ao grupo de baixo risco.

2 JUSTIFICATIVA PARA O ESTUDO E DELINEAMENTO EXPERIMENTAL ADOTADO

Dentre os fatores prognósticos, o acometimento linfonodal é o mais relevante. Mulheres que apresentam invasão linfonodal são submetidas a terapias adjuvantes mais agressivas, aumentando a sobrevida das pacientes. Nesse sentido um problema ocorre com mulheres que apresentam um bom prognóstico por não apresentarem linfonodo comprometido e muitas vezes deixam de receber tratamento adjuvante e acabam por apresentar progressão da doença e metástase. Foi baseado nessa problemática que o objetivo desse estudo foi delineado. A identificação prévia dos casos com maior potencial de progressão da doença e metástase seria de grande importância, pois essas mulheres poderiam se beneficiar de terapias adjuvantes mais severas e conseqüentemente terem maiores chances de cura da doença.

Assim o objetivo desse estudo é identificar marcadores moleculares de predição de metástase através de *microarray* utilizando uma lâmina com genes pertencentes a vias sabidamente envolvidas no processo de invasão e metástase (EMT, PI3K e Wnt).

Inicialmente esse projeto foi delineado, utilizando-se como amostras, carcinoma ductal de mama de pacientes com bom prognóstico, isto é, pacientes sem acometimento linfonodal, que não apresentaram metástase por pelo menos 10 anos e pacientes que apresentaram metástase. Estudos

para identificar estes tipos de marcadores moleculares preditivos são difíceis de serem delineados e requerem um banco de tumores com um grande número de amostras e longo seguimento clínico. Essa dificuldade está presente principalmente na seleção de amostras para o grupo de pacientes com linfonodo negativo e que apresentaram metástase. Apesar da nossa instituição já manter um banco de tumores de amostras congeladas há quase 10 anos, não foi possível encontrar um número razoável de amostras com essas características, para um bom delineamento experimental. Assim, para esse projeto, a única fonte de amostra seria o arquivo de amostras emblocadas em parafina fixadas em formol (FFPET).

Nosso grupo com objetivo de usar esse tipo de amostra para estudos de expressão vinha trabalhando para o estabelecimento da metodologia de extração e amplificação de RNA de amostras FFPET. Vários protocolos de extração de RNA total foram testados e várias adaptações no protocolo de amplificação como inserção de controles para avaliação do RNA amplificado (aRNA) foram realizadas para que gerasse resultados confiáveis de expressão gênica através de *microarray*. Assim várias adaptações e controles foram estabelecidos durante o procedimento de amplificação do RNA, que resultaram em experimentos de *microarray* com qualidade satisfatória (SARAIVA et al. 2006). Esses resultados obtidos com amostras FFPET de 2000 a 2004 nos encorajaram a propor esse projeto. Dessa forma a proposta inicial era a comparação de expressão gênica, de amostra FFPET de carcinoma ductal de mama sem comprometimento linfonodal, divididas em dois grupos: pacientes que não apresentaram metástase em 10

anos e pacientes que apresentaram metástase. Para tal foram verificados em um banco de dados do Hospital A C. Camargo, 1.092 casos de câncer de mama entre os anos de 1990 a 2002. Inicialmente foram selecionados 313 casos (116 casos que apresentaram metástase e 197 casos que não apresentaram metástase em 10 anos). Após uma minuciosa revisão dos dados coletados, foram selecionados 30 casos para esse estudo, que foram separados em dois grupos (10 casos que apresentaram metástase e 20 casos que não apresentaram metástase em 10 anos) sendo todos pré-menopausa e que receberam apenas tratamento de quimioterapia e/ou radioterapia pós-cirurgia, excluindo as pacientes que receberam hormonioterapia (para evitar influências dos diferentes tratamentos sobre a evolução da doença).

Com o intuito de realizar os experimentos com uma população celular mais homogênea possível, essas amostras foram microdissecadas a laser. Para uma melhor recuperação e amplificação do RNA, foi utilizado o kit ParadiseTM Reagente System (Arcturus), específico para amostras FFPE microdissecadas a *laser*. No entanto, a utilização desse kit não apresentou resultados satisfatórios quando empregado em amostras desse arquivo.

Dessa forma, novas tentativas com essas mesmas amostras foram realizadas utilizando-se protocolo estabelecido pelo grupo, no qual é realizada uma microdissecção manual (MM) possibilitando a extração de uma quantidade maior de RNA. Após a introdução dessas modificações, não houve melhora nos resultados, isso nos levou a concluir que a condição de armazenamento dessas amostras FFPE do Hospital A. C. Camargo, não

deve ter sido adequada o suficiente para recuperação de RNA total de qualidade para este tipo de estudo, uma vez que essas amostras são mais antigas quando comparadas com as amostras utilizadas no nosso estudo anterior (SARAIVA et al. 2006).

Partindo das amostras congeladas de carcinoma ductal de mama disponíveis no banco de tumores, foi proposto um novo delineamento experimental na busca de genes candidatos preditivos de metástase linfonodal, utilizando para esta proposta, amostras de tumor primário de carcinoma ductal invasivo de pacientes sem comprometimento de linfonodo e com seguimento mínimo de 3 anos (grupo 1) e de pacientes com linfonodo comprometido (grupo 2). Foram selecionadas 47 amostras de carcinoma ductal de mama (22 sem comprometimento de linfonodo livre de metástase em 3 anos e 25 com comprometimento de linfonodo). Na busca de genes envolvidos com invasão e metástase, foi utilizada uma lâmina com genes de vias sabidamente envolvidas nesses processos (PI3K, EMT e Wnt) e através da análise desses dois grupos, buscamos diferenças no perfil de expressão desses genes de interesse.

Assim, mediante os acontecimentos os resultados desse estudo foram divididos em duas partes: experimentos utilizando-se amostras FFPET, detalhando todos os resultados e problemas encontrados e resultados dos experimentos utilizando-se amostras congeladas.

3 OBJETIVOS

3.1 OBJETIVO GERAL

Identificar genes candidatos a marcadores moleculares associados à progressão e metástase em carcinoma ductal de mama.

3.2 OBJETIVO ESPECÍFICO

- Selecionar as amostras tumorais de mama a partir de revisão dos prontuários clínicos;
- Microdissecção das amostras tumorais;
- Extração e amplificação do RNA;
- Experimentos de *microarray*;
- Análises Matemáticas e Estatísticas para identificação de diferença no padrão de expressão entre os dois grupos.

4 MATERIAL E MÉTODOS

4.1 SELEÇÃO DAS AMOSTRAS

Foram selecionadas 30 amostras FFPET de carcinoma ductal de mama (10 casos que apresentaram metástase e 20 casos que não apresentaram metástase em 10 anos) do arquivo de parafina do Hospital do A. C. Camargo e 47 amostras congeladas de carcinoma ductal de mama (22 sem comprometimento de linfonodo e sem recidiva em 3 anos e 25 com comprometimento de linfonodo) do banco de tumores do Hospital do A. C. Camargo.

4.2 MICRODISSECÇÃO DAS AMOSTRAS

4.2.1 Procedimento geral para Microdissecção a Laser (LCM)

A Preparação das amostras parafinadas para LCM

- Banho com xilol 2 minutos;
- Banho com xilol 2 minutos;
- Banho com etanol 100% 2 minutos;
- Banho com etanol 95% 1 minuto;
- Banho com etanol 75% 1 minuto;
- Banho com água DEPC 30 segundos;
- Coloração com azul de toluidina; (0,5g em 100mL de álcool 50%)

- Banho com etanol 75% 30 segundos;
- Banho com etanol 95% 30 segundos;
- Banho com etanol 100% 1 minuto;
- Xilol 5 minutos.

B Preparação das amostras congeladas para LCM

- Banho de etanol 70% 30 segundos;
- Banho com água DEPC 30 segundos;
- Coloração com HE;
- Banho com água DEPC 30 segundos;
- Banho com etanol 75% 30 segundos;
- Banho com etanol 95% 30 segundos;
- Banho com etanol 100% 30 segundos;
- Banho com xilol 5 minutos

C Microdissecção a laser (LCM)

Dez lâminas para cada amostra (corte de 5µm) foram preparadas para a LCM e rapidamente armazenadas no gelo seco ou diretamente a -80°C, para que não ocorresse degradação do RNA. O sistema para microdissecção a *laser* consiste basicamente de um microscópio invertido ao qual é acoplada uma fonte de raio *laser* infravermelho de baixa potência e uma câmera ligada a um monitor e um computador (*Pixcell Laser Capture Microdissection System- Arcturus Systems for Microgenomics*). A lâmina com o corte histológico é fixada ao *charriot* por um mecanismo de sucção a

vácuo. O feixe de raio *laser* incide sobre um *cap* revestido por um filme termoplástico. O *cap* fica posicionado entre o feixe de *laser* e o corte histológico. Com o disparo do feixe de *laser*, o filme termoplástico expande, se funde e adere à célula escolhida. Quando o pulso cessa, o filme retrai levando consigo a célula aderida à sua superfície. Desta forma, repete-se a operação tantas vezes quantas forem necessárias. O tempo máximo de captura das células é de aproximadamente 15 minutos por lâmina, a fim de se preservar o RNA total o máximo possível. O procedimento de captura das células foi realizado pela Dra. Renata Coudry do Departamento de Anatomia Patológica dessa instituição. Para as amostras FFPET foram em média 10.000 disparos e para as amostras congeladas 1.000 disparos, para diminuir o tempo da microdissecção e preservar melhor o RNA.

4.2.2 Microdissecção Manual (MM)

A Preparação das amostras parafinadas para a microdissecção manual (MM)

Para cada amostra FFPET foram realizados 4 cortes com 5µm cada e uma microdissecção com bisturi foi feita quando o tumor estava presente em menos de 70% da amostra. Os cortes foram colocados nas lâminas e com auxílio de uma lâmina corada com hematoxilina-eosina (HE), a área tumoral foi separada com bisturi e colocada em tubo de 2ml. Esses cortes foram desparafinizados com 1ml de xilol e em seguida acrescentado um tampão com Proteinase K por 18 horas.

4.3 EXTRAÇÃO DE RNA TOTAL DAS AMOSTRAS

4.3.1 Extração de RNA total de Amostras Parafinadas (LCM)

Ao final da operação de LCM, retirou-se o *cap* com as células de interesse aderidas e seguiu-se com o procedimento de extração do RNA total, conforme recomendado pelo fabricante ParadiseTM Reagente System (*Arcturus*).

4.3.2 Extração de RNA total de amostras congeladas (LCM)

Para as amostras congeladas o procedimento de extração foi realizado através do kit PicoPure RNA Isolation Kit –(*Arcturus*) conforme recomendações do fabricante.

4.3.3 Extração de RNA total de amostras parafinadas (MM)

Para amostras parafinadas que sofreram MM o RNA total foi extraído utilizando-se Rneasy Mini Kit- QIAGEN conforme recomendações do fabricante.

4.4 AVALIAÇÃO DA INTEGRIDADE DO RNA TOTAL

4.4.1 Avaliação por RT-PCR dos genes *NOTCH2* e *GAPDH*

Na ausência de equipamentos para avaliar a integridade do RNA a partir de poucas nanogramas de amostra, essa avaliação foi realizada através de síntese de fita dupla de cDNA seguida de uma reação de PCR

para o gene *NOTCH2* que possui 11.413 pb. Os produtos foram analisados em gel de poliacrilamida 8 %, corados por nitrato de prata (SANGUINETTI et al. 1994). A reação de PCR do gene *NOTCH2* consiste na utilização de um par de *primer*, localizado a 6.336 pb da região 3' que amplifica um fragmento de 231 pb. Outro controle utilizado, foi a PCR do gene *GAPDH*, que possui 1.283pb, e o par de *primers* se anela a 463 pb da extremidade 3' e amplifica um fragmento de 292 pb.

A extensão dos fragmentos de tamanho esperado reflete a integridade do RNA total e a eficiência da síntese do cDNA. A extensão de fragmento do gene *NOTCH2* e *GAPDH* reflete indiretamente a integridade do RNA total e a qualidade do cDNA. A extensão do fragmento *NOTCH2*, reflete um cDNA de qualidade extremamente alta e a do fragmento do *GAPDH* é considerado condição necessária para seguimento dos experimentos.

Primers

NOTCH2-2F – GGTGAACAAGAACAGGAG

NOTCH2-2R – GATGACAACAGCAACAGC

GAPDH-F- CTGCACCACCAAGTGCTTA

GAPDH-R-CATGACGGCAGGTCAGGTC

4.4.2 Avaliação pelo 2100 Bioanalyzer (Agilent)

Esse aparelho permite a análise de integridade de quantidade ínfima de RNA total (RNA 6000 Pico Chip- sensibilidade 0.2 - 5ng de RNA total) e

gera um número de integridade do RNA (RIN, *RNA Integrity Number*). Esta ferramenta é baseada em uma rede neuronal que determina o número RIN. O algoritmo permite a classificação do RNA total em um sistema de numeração de 1 a 10, com 1 sendo o perfil mais degradado e 10 o mais intacto e se baseia no padrão das bandas 18S e 28S do RNA.

4.5 AMPLIFICAÇÃO DO RNA MENSAGEIRO (MRNA)

4.5.1 Amplificação do RNA Mensageiro de amostras FFPET Microdissecadas a laser (LCM)

Os mRNAs foram amplificados em dois *rounds* com a utilização do Kit Paradise™ Reagente System (Arcturus).

4.5.2 Amplificação do mRNA de amostras FFPET Microdissecadas Manualmente e da amostra Referência (linhagem HB4A)

A metodologia de amplificação utilizada no experimento é baseada no protocolo descrito por GOMES et al. (2003) com algumas modificações para amostras FFPET estabelecido em nosso grupo (SARAIVA et. al. 2006). O método de amplificação consiste na reação de síntese de primeira fita de cDNA utilizando RNA total e Oligo dT(24)-T7 primer (5'AAA CGA CGG CCA GTG AAT TGT AAT ACG ACT CAC TAT AGG CGC T(24)3')[57-mer] (relação de RNA total e primer de 1:0,1 para amostras FFPET e relação de 1:1 para linhagem HB4), em um volume total de 7,0µl. Esse *primer* carrega o promotor do fago T7 bacteriano, que será utilizado para a transcrição *in vitro*.

O RNA total com o *primer* Oligo dT(24)-T7 foi desnaturado a 70°C por 15 minutos e esfriado em gelo e então adicionaram-se 4,0µl do tampão primeira fita 5X (Invitrogen), 1,0µl RNaseIN (Promega), 2,0µl dNTP 10mM (Pharmacia), 2,0µl DTT 0,1M (Invitrogen), 2,0µl de água DEPC e 2,0µl Superscript II (Invitrogen). A síntese do cDNA foi completada após 2h a 42°C.

A segunda fita do cDNA foi feita pela adição de 53,0µl água DEPC, 20,0µl de *second strand* (Invitrogen), 1,0µl dNTP 10mM (Pharmacia), 1,0µl RNase H (Invitrogen), 1,0µl de *E. coli* DNA ligase 10U/µl (Invitrogen), 4,0µl *E. coli* DNA polimerase I 10U/µl (Invitrogen), a reação foi incubada por duas horas a 16°C, ao fim desse período foi adicionado 2,0µl de T4 DNA polimerase I 50U/µl (Invitrogen) e novamente incubado a 16°C por 5 minutos. A reação foi finalizada pela adição de 10,0µl de 0,5M EDTA.

A purificação do cDNA dupla fita foi feita pela adição de 1000,0 µl de fenol: clorofórmio: álcool isoamílico (25:24:1) pH 8,0. Essa mistura foi centrifugada a 14000 rpm por 5 minutos / TA. A fase aquosa foi transferida para um novo tubo contendo 100,0 µl de clorofórmio e a mistura foi centrifugada por mais 5 min a 14000 rpm / TA. A fase aquosa foi transferida para um novo tubo para precipitação do cDNA.

O cDNA foi precipitado pela adição de 3,0X o volume recuperado da fase aquosa (V_r) de etanol 100% (MERCK) a -20°C, 40µl de acetato de amônio 7,5M e 2,0µl de acrilamida 200 ng/µl (Ambion, Inc. – Applied Biosystems), incubação por 30 min a -80°C e centrifugação a 14000 rpm

durante 30 min / 4°C. O precipitado foi lavado 3 vezes com etanol 70% e ressuspendido em 10,0 µl de H₂O-DEPC.

Para a transcrição *in vitro*, o ds-cDNA (cDNA dupla fita) foi utilizado o *RiboMAX™ Large Scale RNA Production Systems – T7 - Promega*. O cDNA dupla fita foi ressuspendido em 11,0µl de água DEPC, sendo 1,0µl utilizado para avaliação da integridade do RNA por PCR do gene *NOTCH2* e *GAPDH*, e os 10,0µl restantes acrescidos de 7,5µl rNTP 25mM, 5,0µl tampão de reação 5X, 2,5µl da solução de misturas de enzimas (RNase IN e fago T7 polimerase - Promega) e incubado a 37°C por 6h. A purificação do a RNA foi feita pelo método de Trizol conforme indicações do fabricante (Invitrogen) e precipitado com isopropanol (volume/volume), acetato de sódio (3M) e 0,5µl de glicogênio (20mg/ml).

Para o segundo ciclo de amplificação, ressuspendeu-se o aRNA em 6,0µl de água DEPC, utilizou-se 3,0µl de primer dN6 a 3µg/µl e foi incubado a 70°C/5min, depois foram acrescido de 4,0µl de tampão primeira fita 5X (Invitrogen), 2,0µl DTT 0,1M (Invitrogen), 1,0µl RNaseIN (Promega), 2,0µl dNTP 10mM (*Pharmacia*) e 2,0µl Superscript II (Invitrogen). A síntese do cDNA foi completada após 2h a 42°C. A segunda fita do cDNA foi feita pela adição de 4,0µl Oligo dT(24)-T7 primer (0,1µg/µl), 61,3µl água DEPC, 10,0µl tampão Advantage PCR (Clontech) , 2,0µl dNTP 10mM (*Pharmacia*), 0,7µl RNase H (Invitrogen) e 2,0µl Advantage Polimerase (Clontech) seguindo os seguintes ciclos de incubações: 37°C/10min para digestão do RNA; 94°C/3min para denaturação; 65°C/5min para a ligação dos primers e 75°C/30min para extensão. A parada da reação e a degradação ribossômica

foi realizada pela adição de 6,5µl de NaOH 1N/EDTA pH 8,0 por 15 min a 70°C.

A partir desse ponto repetiu-se exatamente os passos do primeiro ciclo de amplificação, resultando no aRNA segundo *round*. A concentração do RNA amplificado foi verificada em espectrofotômetro utilizando o comprimento de onda equivalente a 260nm (1OD a 260nm equivale a 40ng/µl de RNA) e a pureza através da relação entre as leituras OD 260 e 280.

4.5.3 Amplificação do RNA mensageiro de amostras congeladas (LCM).

Os mRNAs foram amplificados em dois *rounds* com a utilização do Kit RiboAmp (Arcturus) segundo recomendações do fabricante. A concentração do RNA amplificado foi verificada utilizando-se o aparelho NanoDrop (ND-1000 Spectrophotometer).

4.6 AVALIAÇÃO DA QUALIDADE DO RNA AMPLIFICADO (ARNA)

4.6.1 Avaliação da Qualidade do RNA Amplificado das Amostras FFPET

Os produtos dos RNAs amplificados foram analisados por eletroforese em gel de agarose 1%, corados com brometo de etídio na presença de marcador de peso molecular. O critério de avaliação dos produtos

amplificados utilizados no laboratório é uma maior concentração de produto amplificado maior que 200pb.

4.6.2 Avaliação da Qualidade do RNA Amplificado das Amostras Congeladas

Os produtos dos RNAs amplificados foram analisados no Bioanalyzer 2100 (Agilent). O critério de avaliação dos produtos amplificados utilizados no laboratório é uma maior concentração de produto amplificado maior que 200pb.

4.7 MARCAÇÃO

Para a marcação foi utilizado 6µg de aRNA, e adicionado um controle (1,2 µL de Q-gene) na reação de hibridização que hibridiza com um *spot* específico presente na plataforma. Em seguida, foi realizada a síntese do cDNA adicionando 3,0µl de primer dN6 a 3µg/µl ao aRNA, a reação foi submetida à temperatura de 70 °C por 5 minutos. Em seguida adicionou-se quantidade igual (0,6µL dNTP) de adenina, timina e guanina (25mM cada) e citosina (10mM). Foi acrescentado na reação 2µL de Cy3 ou de Cy5 (25mM), que consiste numa solução de citosinas ligadas a um fluoróforo, 6,0µl de tampão primeira fita 5X, 0,5µl RNaseIN (Promega) e 2,0µL da enzima *Super Script Reverse Transcriptase*. A reação permaneceu a 42°C por duas horas. Para a degradação do aRNA foi acrescentado 1,5µL de

EDTA 0,5M e 1,5µL de NaOH 1M, e a reação foi submetida a 70°C por 20 minutos e finalmente neutralizada com µL de HCl 1M.

Para a eliminação de não-incorporados e resíduos de reação, passa-se a solução numa coluna de purificação, denominada AutoSeq G50 (Armesthan Pharmacia), conforme recomendações do fabricante. Após purificação a amostra e sua referência são misturadas e acrescidas de poliA DNA (4µg - Armesthan Pharmacia), Cot 1 (4µg) e BSA (20 µg – BSA fração 5 - Sigma). Ajuste do volume para 12,35 µL e em seguida, adiciona-se os reagentes na concentração final de: tampão de hibridização 1X (previamente aquecido a 70 °C), formamida 25%, Denhardt's 5X e 10µg de DNA de esperma de salmão para um volume final de reação de 95 µL. A solução final é aquecida a 95 °C por 5 minutos para total desnaturação das fitas e aplicadas na estação de hibridização, cujas lâminas são previamente hibridizadas a 42 °C (de 6-16h) em tampão SSC 5X, SDS 0,2%, BSA 1%, solução de Denhardt's 5X (filtrada em membrana de poro 0,22 ou 0,45 µm) e lavadas em água destilada, sendo então montada na estação GeneTac 2000 (Genomic Solutions).

Cada amostra teste e referência foram marcadas em duplicata, utilizando o sistema inverso de incorporação dos corantes (*Dye swap*) com os fluoróforos Cy3 e Cy5. Essas amostras são misturadas e hibridizadas em duplicata.

4.8 HIBRIDIZAÇÃO, LAVAGEM E CAPTURA DE IMAGEM

O processo de hibridização é realizado a 42 °C por no mínimo 16 h e no máximo 22h. Após hibridização, as lâminas passam por um processo de lavagem, no qual os tampões de lavagem são pré-aquecidos a 42° C. As lâminas são lavadas nas seguintes condições: uma vez em solução SSC 2X, duas vezes em solução de SSC 0,1% e SDS 1X e duas vezes em solução de SSC 0,1%, todos os passos são realizados sob agitação por 10 minutos, sendo as lâminas centrífugas em tubo Falcon 50µL a 1.000rpm por 2 minutos para secagem.

As lâminas são submetidas a uma varredura por laser utilizando *scanner a laser confocal ScanArray Express* (Perkin Elmer Life Sciences, Inc., EUA) sendo que as moléculas de Cy3 e Cy5 quando estimuladas pelo laser, emitem fluorescência em tamanhos de ondas específicos capazes de serem detectadas pelo *scanner*. Para as amostras testes foi utilizado PMT de 60% e para a amostra referência PMT de 50%.

Os sinais gerados pela hibridização competitiva entre moléculas das amostras marcadas com os Cy3 e Cy5, são então capturados e quantificados. Este programa permite a localização automática dos pontos de cDNA hibridizados na lâmina baseando-se em marcadores de posição, necessitando checagem e ajuste manual. O programa realiza o cálculo da intensidade de cada ponto de hibridização para cada uma das amostras.

Todas as tabelas geradas após a quantificação dos dados foram submetidas ao Laboratório de Biologia Computacional do Hospital A. C.

Camargo coordenado pela Dra. Helena Brentani para análises de qualidade. Para cada amostra foi possível obter informações sobre as correlações entre as réplicas.

4.9 PLATAFORMAS DE CDNA MICROARRAY

4.9.1 Plataforma 4.8K

A lâmina contém 4.800 (4.8K) seqüências imobilizadas sendo que 4.608 correspondem a clones de ORESTES que representam genes humanos (BRENTANI et al. 2005). Essa plataforma foi utilizada apenas para os testes iniciais de hibridização das amostras FFPET, pois nesse período a construção da plataforma 2.3K (lâmina vias) estava sendo finalizada.

4.9.2 Plataforma 2.3K de vias metabólicas

A plataforma customizada 2.3K contém 1808 fragmentos de cDNA correspondentes a genes humanos (Projeto FAPESP 06-00054). Os genes pertencentes à via de sinalização Wnt e PI3K foram selecionados do banco de dados *Kyoto Encyclopedia of Genes and Genomes*-KEGG. Os genes pertencentes ao processo EMT foram selecionados através de literatura e em outros estudos em desenvolvimento no nosso grupo. Do total de **1808** genes, 349 correspondem a genes selecionados que representam as vias de sinalização de interesse e que foram avaliados nesse estudo. Desses 349, 174 representam genes da via de sinalização Wnt, 97 da via PI3K, 78 do processo EMT (Tabela em Anexo 2) e mais 8 genes de expressão

constitutiva. Outros genes das vias do ácido retinóico, diferenciação neuronal e outros cDNA correspondente a genes humanos oriundos de outros projetos do nosso grupo e colaborações, também foram representados nessa lâmina, somando aproximadamente 1808 seqüências de cDNA, além de 48 controles positivos e 384 controles negativos. Os fragmentos foram organizados em grupos de 48 *subarrays*, dispostos em 4 colunas com 12 fileiras cada um, e cada *subarray* com 49 *spots*, dispostos em 7 colunas com 7 fileiras cada um.

4.10 NORMALIZAÇÃO DOS DADOS

A fim de se ajustar variações sistemáticas lineares e não-lineares, ambas do tipo intensidade-dependentes, principalmente para elementos de baixa intensidade, foi empregado o método de regressão linear localmente ponderada LOWESS – *Locally Weighted Scatter-plot Smoothing* (QUACKENBUSH 2001). Seguindo a abordagem de comparação indireta, empregamos uma referência comum (RNA amplificado de linhagens de mama HB4A). Desta forma, para cada gene foi determinada a razão de expressão, em $\log_2$, entre as intensidades da amostra e referência (CHURCHILL 2002; QUACKENBUSH 2002; YANG e SPEED 2002) realizado pelo grupo da Dra. Helena Brentani.

4.11 ANÁLISES DOS DADOS

As análises matemáticas e estatísticas foram realizadas pelo grupo do Dr. Eduardo Jordão, IME-USP. Resumidamente, foi empregada uma linguagem computacional e gráfica R (IHAKA e GENTLEMAN 1996) do *Contributed R Archive Network* - CRAN (2007), que promove análises estatísticas e gráficas, incluindo modelos lineares e não-lineares, *clusters*, previsões, análises *resampling survival* e *time-series*. Além disso, foram aplicadas ferramentas do projeto bioconductor (<http://www.bioconductor.org>) que permitem análise de dados de genômica e de expressão gênica (QUACKENBUSH 2001; BALL et al. 2003; DUDOIT et al. 2003).

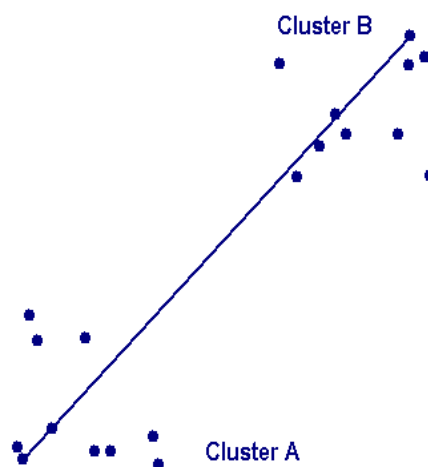
4.11.1 Dendograma (método de ligação completa e métrica de correlação)

O dendograma foi construído utilizando-se método de ligação completa e métrica de correlação, para isso foi medida a distância entre as amostras definida entre o par de genes mais distante, um par de cada amostra (controle e *swap*).

$$D(r,s) = \text{Max} \{ d(i,j) : \text{onde o gene } i \text{ está no cluster } r \text{ e o gene } j \text{ no cluster } s \}$$

A distância entre todos os possíveis pares de genes (i,j) foi computada, onde o gene i está no cluster r e o gene j está no cluster s e o valor máximo dessa distância foi a distância entre esses dois clusters. Ou

seja a distância entre dois clusters, r e s , foi dada pela maior distância entre os genes i e j pertencentes a cada um deles respectivamente. (Figura 4) (XLMINER 2007). Nesse projeto, para uma lâmina ser considerada de boa qualidade, essa correlação foi igual ou superior a 80% após a normalização.



Fonte: XLMINER (2007)

Figura 4 - Esquema demonstrando como é medida a distância entre os pares de genes na métrica de correlação.

4.11.2 Genes diferencialmente expressos

Para se determinar a significância da diferença de expressão foi aplicado o teste paramétrico de *T-student* para todos os genes do conjunto de dados.

4.11.3 Método de clusterização SOM (*self-organization map*).

Para se verificar a similaridade entre os dois grupos de amostras, foi construído um cluster pelo método SOM, é um método de visualização de dados multidimensionais. O objetivo é diminuir a dimensão dos dados, tipicamente para uma ou duas dimensões, de modo que a representação de baixa dimensão seja coerente com o espaço multidimensional.

O método funciona da seguinte forma: uma topologia é definida para grade, que pode ser retangular ou hexagonal e o número de dimensões em x e y que irão determinar o número de nós da mesma. A cada ponto dessa malha escolhe-se aleatoriamente um vetor de dados. Em seguida, repete-se esse processo para os outros pontos de modo que, por competição, os dados mais similares, medidos de acordo com uma métrica pré-selecionada, vão se deslocando e agregando em regiões próximas.

Dessa forma, ao final do processo, além de uma representação de baixa dimensão dos dados, obtêm-se os dados agrupados de acordo com sua similaridade (<http://www.cis.hut.fi/projects/somtoolbox/theory/somalgorithm>).

4.11.4 Construção e análise de redes de relevância

Essa análise foi realizada para se identificar estes pares de genes com correlação linear alterada entre os grupos, foi feito uma comparação entre pares de genes, dentro de cada grupo, e a seguir, foram determinados seus coeficientes de correlação. A partir destes valores, foi construída uma rede, por representação gráfica, atribuindo coloração verde para correlação positiva e, vermelha para correlação negativa. A idéia original proposta por BUTTE et al. (2000) foi ligeiramente modificada para a busca por pares de genes com valores de correlação linear significativamente alterado entre determinados grupos amostrais.

Para comparar as alterações entre os coeficientes de correlação, entre os dois grupos (com e sem comprometimento de linfonodo), foi aplicada primeiramente a transformação Z de Fisher, onde a diferença entre os valores de correlação para cada par de genes, nos dois tipos biológicos estudados, foi transformada em uma nova variável Z, que apresenta distribuição normal aproximada. Esta informação para aferir a significância da diferença de correlação entre os grupos ao aplicarmos o teste de hipótese para esta variável Z.

5 RESULTADOS E DISCUSSÃO

5.1 DADOS REFERENTES ÀS AMOSTRAS PARAFINADAS DO ARQUIVO DO HOSPITAL A. C. CAMARGO

5.1.1 Seleção das amostras FFPET

Inicialmente foram selecionados 313 casos, todos sem linfonodo acometido (116 casos que apresentaram metástase e 197 casos que não apresentaram metástase em 10 anos) a partir de informações de 1.092 casos de câncer de mama entre os anos de 1990 a 2002, operados nessa instituição. Esse banco de dados foi disponibilizados pelo Dr. Fernando Soares.

Os casos de interesse foram selecionados e foi realizada uma revisão de prontuários, a fim de selecionar apenas os casos de carcinoma ductal sem comprometimento de linfonodo e sem tratamento prévio (quimioterapia, radioterapia e tratamento por outro tipo de câncer anterior). Após essa revisão detalhada de 118 casos, foram selecionados (44 casos que apresentaram metástases e 74 casos que não apresentaram metástase em 10 anos). Com o objetivo de comparar grupos mais homogêneos possíveis foi verificado o tipo de tratamento que cada paciente foi submetida após a cirurgia (quimioterapia, radioterapia e hormonioterapia), além do status hormonal (pré ou pós-menopausa). Essa análise resultou na seleção de 30 amostras, todas pré-menopausa, sem tratamento prévio e que receberam

apenas tratamento de quimioterapia e/ou radioterapia (10 casos que apresentaram metástase e 20 casos que não apresentaram em 10 anos). Para se trabalhar com um número equivalente de amostras em cada grupo, inicialmente foram selecionadas 10 casos de cada grupo (Tabela 1).

A idade das pacientes variou de 30 a 52 anos (média de 41,4 anos) e a média de número de linfonodos examinados foi de 19,2 para o grupo com metástase e de 23,4 para o grupo que não apresentou metástase em 10 anos.

Tabela 1 - Características clínico-patológicas das pacientes - Amostras FFPET.

Amostra- Ano da cirurgia	Idade	pTNM	Linfonodos comprometidos /Linfonodos examinados	Tratamento	RE	RP	Metástase	Local
1- 1993	39	T1cN0M0	0/15	RT	-	-	Sim	Osso
3- 1994	43	T1N0M0	0/19	RT	-	+	Sim	Osso
5-1995	44	T2N0M0	0/28	QT	-	-	Sim	Osso
7-1993	43	T2N0M0	0/21	RT/QT	-	-	Sim	Pulmão, SNC
9-1990	38	T2N0M0	0/23	RT/QT	+	+	Sim	Osso, Mama e Pulmão
11-1992	37	T2N0M0	0/25	RT/QT	-	+	Sim	Osso
13-1991	33	T1N0M0	0/17	RT/QT	-	-	Sim	Mama, SNC, Pulmão, Osso
15-1994	42	T2N0M0	0/26	RT/QT	+	+	Sim	Osso, Pleura
17-1996	41	T2N0M0	0/8	RT/QT	-	-	Sim	Pulmão
19-1992	52	T2N0M0	0/10	RT/QT	-	-	Sim	Linfonodo axilar
2- 1993	42	T2N0M0	0/22	RT/QT	-	-	Não	
4-1992	46	T2N0M0	0/28	RT	+	+	Não	
6-1993	46	T2N0M0	0/18	RT/QT	-	-	Não	
8-1993	46	T1N0M0	0/25	RT/QT	-	+	Não	
10-1994	43	T1N0M0	0/17	RT/QT	+	-	Não	
12-1995	39	T1cN0M0	0/28	RT	-	-	Não	
14-1994	45	T1N0M0	0/26	QT	-	-	Não	
16-1992	30	T1CN0M0	0/21	RT/QT	-	+	Não	
18-1994	46	T2N0M0	0/32	RT/QT	+	+	Não	
20-1995	33	T2N0M0	0/17	RT	-	-	Não	

RT- radioterapia, QT- quimioterapia

5.1.2 Extração de RNA total e amplificação das amostras FFPET

Dessas 20 amostras, 10 amostras (cinco de cada grupo, amostras: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10) foram selecionadas para os experimentos iniciais de extração de RNA total e amplificação. Para cada amostra foi preparada uma lâmina corada com HE para confirmação do diagnóstico anatomopatológico e também quatro lâminas sem coloração com 4µm de espessura preparadas para a microdissecção a laser.

O RNA total foi extraído e o controle de integridade foi realizado através da PCR dos genes *NOTCH2* e *GAPDH*, utilizando como molde 1,0µl de cDNA separado durante o primeiro *round* da amplificação.

Para o gene *GAPDH* houve extensão do fragmento no tamanho esperado para 6 das 10 amostras (nove delas apresentadas na Figura 5). Das 10 amostras analisadas para o gene *NOTCH2*, nenhuma apresentou bandas do fragmento amplificado em gel de poliacrilamida 8%, mostrando indiretamente que o RNA apresenta certo nível de degradação o que é esperado para esse tipo de amostra FFPET.



Figura 5 - Gel de Poliacrilamida 8%. RT-PCR do gene GAPDH. Amplificado de 292 pb. M- Marcador de 100pb, 1-9 amostras tumorais extraídas por LCM, C+: controle positivo, No: controle negativo.

Para esse tipo de amostra (FFPET), sabidamente de qualidade inferior a amostras congeladas, o controle de qualidade através da PCR do gene *GAPDH* apresenta melhor resultado, pois o par de *primer* se anela bem próximo à região 3 do gene (463pb). Assim, seguiram-se com o experimento de amplificação somente as amostras que apresentaram extensão de fragmento do gene *GAPDH* (1, 2, 3, 5, 8 e 9).

Após a amplificação do RNA das 6 amostras utilizando-se o kit Paradise (Arcturus), a qualidade do aRNA foi analisada em gel de agarose 1% (Figura 6). Nesse período nosso grupo ainda não tinha experiência em julgar a qualidade do aRNA com esse novo Kit, através da avaliação do *smear* em gel de agarose. No entanto, os resultados obtidos através do *smear*, foram classificados como satisfatórios, aqueles que apresentam concentração de moléculas de aRNA a partir de 200pb.

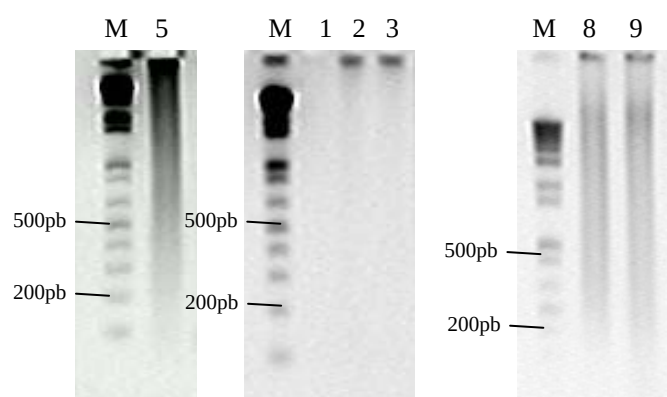


Figura 6 - Gel de Agarose 1%. M- Marcador de 1000pb, 1, 2, 3, 5, 8 e 9 aRNA de amostras tumorais microdissecadas a laser.

Das 6 amostras amplificadas, apenas 3 amostras (5, 8 e 9) tiveram quantidades suficientes e qualidade aparentemente satisfatória de aRNA para experimentos de *microarray* (mais de 6µg). Essas amostras foram marcadas (6µg) e hibridizadas na lâmina 4.8K para avaliar a qualidade de hibridização. O resultado da hibridização foi de baixa qualidade, verificado na Figura 7, que mostra a lâmina marcada com a amostra 5 em verde e a referência em vermelho e o *composite* da lâmina (sobreposição das fluorescências emitidas pelo Cy3 e Cy5 da amostra e da referência). Pode-se notar, que apenas o cDNA da referência (vermelho) apresentou intensidade de sinal detectado na representação da lâmina, e a amostra marcada com Cy3 (verde) apresentou baixo ou nenhum sinal. Esse sinal obtido é proporcional ao número de moléculas hibridizadas com o cDNA correspondente imobilizado na plataforma, sugerindo uma baixa hibridização dessa das moléculas dessa amostra.

Essas lâminas não passaram por pré-análise de qualidade devido à baixa qualidade da hibridização verificada visualmente.

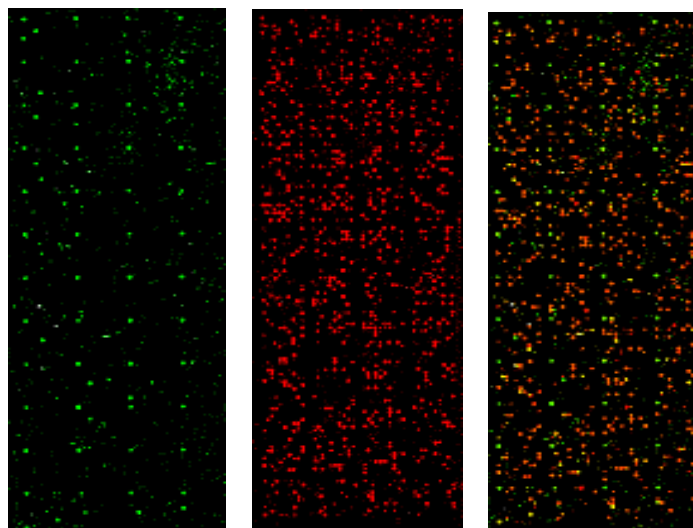


Figura 7 - Representação das lâminas 4.8K, mostrando respectivamente a controle, *swap* e composite.

As amostras 8 e 9, apresentaram resultados de hibridizações semelhantes ao da amostra 5. Esse resultado mostra que apesar dessas amostras terem apresentado fragmentos de aRNA a partir de 200pb, não apresentou qualidade nos experimentos de hibridização, provavelmente está relacionado à pior conservação do RNA de amostras FFPET. Além disso, elas foram microdissecadas a laser, que diminui a quantidade de RNA total extraído, prejudicando experimentos de amplificação.

Os resultados de todos os experimentos realizados com as amostras FFPET microdissecadas a laser se encontram na Tabela 2 a seguir.

Tabela 2 - Resultados obtidos nos experimentos utilizando as 10 amostras Seleccionadas FFPET microdissecadas a laser e amplificadas com o Kit Paradise.

Amostra	Amplificação do gene NOTCH2	Amplificação do gene GAPDH	Massa total de RNA após amplificação μg	Marcação e Hibridização
1	Não	Sim	0,5	Não
2	Não	Sim	1,8	Não
3	Não	Sim	3,5	Não
4	Não	Não	-	Não
5	Não	Sim	23,8	Sim
6	Não	Não	-	Não
7	Não	Não	-	Não
8	Não	Sim	19,5	Sim
9	Não	Sim	7,3	Sim
10	Não	Não	-	Não

Novos testes com essas mesmas amostras foram realizados utilizando o protocolo estabelecido por nosso grupo utilizando amostras FFPET microdissecadas manualmente (SARAIVA et al. 2006) para as quais é possível obter maiores quantidade de RNA total na ordem de microgramas, o que favorece o sucesso no procedimento de amplificação. Assim, com o objetivo de testar nosso protocolo nessas amostras, elas foram então microdissecadas manualmente para obter maiores quantidades de RNA total.

A qualidade do RNA total extraído foi analisada através do aparelho 2100 Bioanalyzer (Agilent) (Figura 8), que nessa época já estava disponível na nossa instituição.

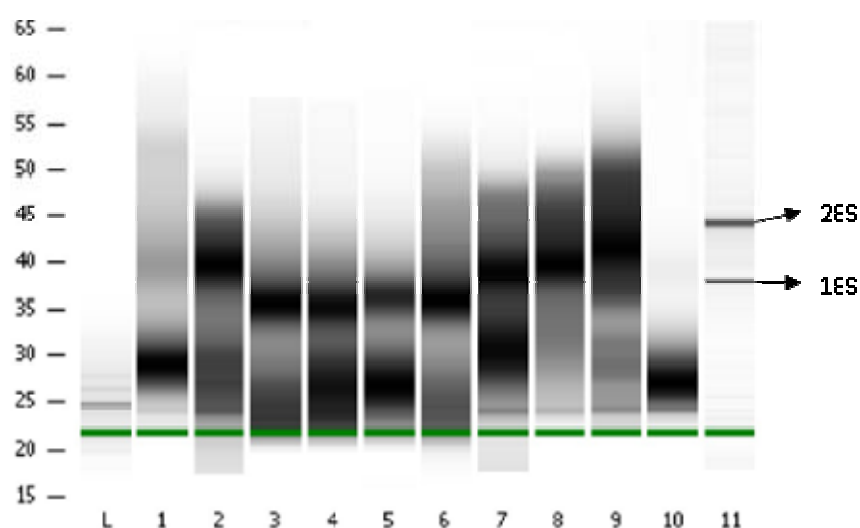


Figura 8 - Matriz Gel (Bioanalyzer). L- ladder, 1-10- RNA total de amostras tumorais microdissecadas manualmente, 11- C+.

As amostras não apresentaram uma boa qualidade (bandas 28s e 18s do RNA ribossômico), quando comparadas ao RNA controle (amostra 11- RNA total de linhagem HB4). No entanto, esse resultado já é esperado, uma vez que RNAs totais extraídos de amostras emblocadas em parafina apresentam degradação (FOSS et al. 1994). Assim nós avaliamos a qualidade da síntese de cDNA através da RT_PCR do gene *NOTCH* e *GAPDH* (Figura 9).

Para 5 (6, 7, 8, 9, 10) das 10 amostras avaliadas houve amplificação do fragmento de 292pb do gene *GAPDH*, as quais foram, utilizadas para os experimentos seguintes. Para o gene *NOTCH2*, não houve amplificação do fragmento para nenhuma amostra.

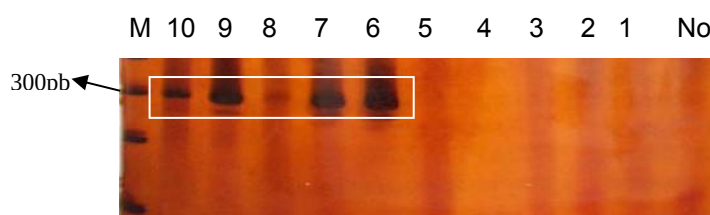


Figura 9 - Gel de Poliacrilamida 8%. RT-PCR do gene GAPDH. Amplificado de 292 pb. M- Marcador de 100pb, 1-10 amostras tumorais extraídas por MM e No: controle negativo.

Dessa forma, o RNA mensageiro dessas 5 amostras foram amplificados, segundo protocolo estabelecido no nosso grupo utilizando a relação de RNA total e *primer* oligo dT-T7 de 1:0,1 para síntese de primeira fita de DNA (SARAIVA et al. 2006). A qualidade do aRNA foi analisada em gel de agarose 1% (Figura 10).

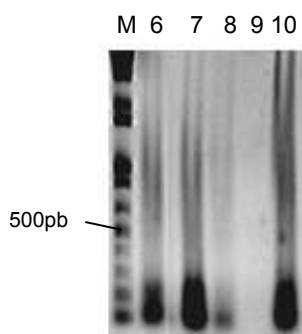


Figura 10 - Gel de Agarose 1%. M- Marcador de 1000pb, 6, 7, 8, 9 e 10 aRNA de amostras tumorais extraídas por MM.

Das 5 amostras avaliadas as amostras 6, 7 e 10 foram as que apresentaram melhor qualidade do aRNA, verificada pelo tamanho das moléculas de RNA geradas. Essas amostras apresentaram aRNA com fragmentos a partir de 100pb, com maior concentração de 100 a 200pb e outra concentração menor entre 500 a 850pb. Essa maior concentração dos fragmentos amplificados nessa faixa, ocorre devido a formação de dímeros de *primers* (oligo dT-T7) durante a síntese de primeira fita. Quando são utilizados em excesso, e a qualidade do RNA inicial é ruim, o *primer* oligo dT-T7 deixa de se anelar à molécula de RNA, e passa a formar esses dímeros que são amplificados juntamente com o RNA.

Apenas a amostra 10 apresentou quantidade suficiente de aRNA (6µg), quantidade necessária para a realização do experimento de *microarray*.

O aRNA da amostra 10 (6µg) foi então utilizado na síntese do cDNA marcado com fluoróforo, e então hibridizado na lâmina 4.8K, mas os resultados de hibridização foram de baixa qualidade, semelhantes as hibridização das amostras LCM (dados não mostrado).

O resultado da hibridização utilizando FFPET microdissecadas manualmente foi bem inferior aos resultados prévios alcançados pelo grupo. O motivo para essa grande diferença, provavelmente está relacionado ao maior tempo de armazenamento dessas amostras, que tem como consequência uma maior degradação do RNA. As amostras utilizadas no estudo prévio eram de 2000-2004 e as amostras testadas nesse estudo foram de 1990-1994.

Os resultados de todos os experimentos realizados com as amostras FFPET microdissecadas manualmente se encontram na Tabela 3 a seguir.

Tabela 3 - Resultados obtidos nos experimentos utilizando as 10 amostras selecionadas FFPET microdissecadas manualmente e amplificadas com o protocolo baseado em Saraiva et al. (2006).

Amostra	Amplificação do gene NOTCH2	Amplificação do gene GAPDH	Massa total de RNA após amplificação μg	Marcação e Hibridização
1	Não	Não	-	Não
2	Não	Não	-	Não
3	Não	Não	-	Não
4	Não	Não	-	Não
5	Não	Não	-	Não
6	Não	Sim	2,4	Não
7	Não	Sim	4,1	Não
8	Não	Sim	2,1	Não
9	Não	Sim	1,3	Não
10	Não	Sim	15,2	Sim

A qualidade do RNA total das amostras é de grande importância para análise de expressão gênica, tanto para permitir a execução dos experimentos, como para a confiabilidade dos dados gerados, pois como foi visto no estudo piloto, RNA com baixa qualidade pode gerar artefatos durante o processo de amplificação, isso acontece devido ao excesso de *primer* (oligo dT-T7) utilizado, se a qualidade do RNA é baixa e há degradação da calda poliA, o *primer* (oligo dT-T7) não se anela as moléculas de RNA, e como consequência, formam dímeros que são

amplificados e quantificados como aRNA, interferindo na qualidade e confiabilidade dos experimentos de *microarray*.

Com esses resultados preliminares, ficou inviável a continuação do projeto utilizando amostras FFPE para estudo de expressão gênica com *microarray*. Porém, essas amostras poderão ser utilizadas para validação de expressão gênica através de RT-PCR em tempo real. Nesse tipo de experimento, cada amostra apresenta um gene normalizador interno que corrige prováveis “vieses” introduzidos pelas diferenças na qualidade das amplificações, diferente dos experimentos de *microarray*, no qual a referência é única para todas as amostras e não há um controle interno específico.

5.2 DADOS REFERENTES ÀS AMOSTRAS CONGELADAS DO BANCO DE TUMORES DO HOSPITAL A. C. CAMARGO

5.2.1 Seleção das amostras congelada

Com o propósito de manter o objetivo inicial do trabalho, na identificação de genes candidatos a marcadores moleculares associados à progressão e metástase em carcinoma ductal de mama, foi realizada uma busca no banco de tumores de amostras congeladas do Hospital A.C. Camargo, para selecionar amostras de carcinoma ductal sem comprometimento de linfonodo e sem tratamento prévio. Foram encontrados 57 casos, no entanto somente três pacientes apresentaram metástase à distância após o tratamento, inviabilizando a manutenção desse desenho

experimental para identificar genes preditivos de metástase, pois o número de casos de pacientes sem comprometimento de linfonodo e que apresentaram metástase foi muito baixo.

Partindo das amostras congeladas de carcinoma ductal de mama disponíveis no banco de tumores, foi proposto um novo delineamento experimental, utilizando amostras de tumor primário de carcinoma ductal invasivo de pacientes sem comprometimento de linfonodo e com seguimento mínimo de 3 anos (grupo 1) e de pacientes com linfonodo comprometido (grupo 2). Foram selecionadas 47 amostras de carcinoma ductal de mama (22 sem comprometimento de linfonodo livre de metástase em 3 anos e 25 com comprometimento de linfonodo, todas sem tratamento prévio e que apresentaram receptor de estrógeno positivo (Tabelas 4 e 5).

No grupo de pacientes que não apresentaram metástase linfonodal, a idade variou de 24 a 80 anos (média de 50,4 anos) e a média de número de linfonodos examinados foi de 19,2 e 81% apresentaram receptor de progesterona positivo para o grupo de pacientes sem comprometimento linfonodal. Nas pacientes com comprometimento de linfonodo a idade variou de 28 a 78 (média de 60,2 anos), o número de linfonodos examinados foi em média 25,7 e 76% apresentaram receptor de progesterona positivo.

Tabela 4 - Características clínico-patológicas das amostras congeladas das pacientes sem comprometimento de linfonodo.

Amostras	Idade	pTNM	Linfonodos comprometidos / Linfonodos examinados	RE	RP
1-N0	34	T2N0M0	0/26	+	-
2-N0	39	T2N0M0	0/33	+	+
4-N0	34	T2N0M0	0/11	+	+
6-N0	34	T2N0M0	0/11	+	+
7-N0	43	T1N0M0	0/30	+	+
8-N0	48	T2N0M0	0/48	+	+
10-N0	51	T2N0M0	0/14	+	+
11-N0	49	T1N0M0	0/2	+	+
12-N0	80	T2N0N0	0/16	+	+
14-N0	44	T3N0M0	0/32	+	+
15-N0	72	T2N0M0	0/21	+	+
16-N0	59	T2N0M0	0/30	+	+
19-N0	55	T1N0M0	0/22	+	+
20-N0	80	T1N0M0	0/11	+	+
23-N0	59	T1CN0M0	0/14	+	+
24-N0	76	T1CN0M0	0/3	+	-
36-N0	24	T2N0M0	0/29	+	+
41-N0	29	T2N0M0	0/23	+	-
42-N0	29	T2N0M0	0/22	+	-
45-N0	48	T1CN0M0	0/2	+	+
46-N0	46	T1N0M0	0/10	+	+
50-N0	76	T2N0M0	0/14	+	+

Tabela 5 - Características clínico-patológicas das amostras congeladas das pacientes com comprometimento linfonodal. Meta- amostras que apresentaram metástase à distância.

Amostras	Idade	pTNM	Linfonodos comprometidos / Linfonodos examinados	RE	RP
1-N1	36	T2N1M0	2/25	+	+
4-N1	40	T2N3M0	15/33	+	+
5-N1	61	T1CN1M0	4/25	+	+
6-N1	50	T2N1M0	3/17	+	-
7-N1	78	T2N1M0	1/20	+	+
8-N1	70	T2N2M0	24/44	+	+
9-N1	67	T4N2M0	36/36	+	+
11-N1	78	T2N1M0	4/30	+	+
12-N1	68	T4BN1M0	2/30	+	+
13-N1	74	T2N1M0	3/30	+	-
14-N1	58	T2N1M0	10/20	+	+
16-N1	70	T2N1M0	8/25	+	+
17-N1	62	T2N1M0	1/13	+	+
18-N1	61	T1N2M0	5/15	+	+
19-N1	53	T2N1M0	1/11	+	+
20-N1	69	T4N1M0	1/28	+	+
22-N1-meta	61	T2N3MX	18/30	+	-
23-N1-meta	73	T4BN3MX	11/14	+	+
24-N1	76	T2N1M0	1/13	+	+
25-N1	58	T2N1M0	1/32	+	-
28-N1	72	T2N1M0	57/57	+	+
32-N1	30	T2N1M0	5/15	+	-
33-N1	28	T2N1M0	1/36	+	+
36-N1-meta	45	T2N1MX	2/30	+	+
38-N1	67	T1CN1M0	1/14	+	-

5.2.2 Processamento das amostras

As 47 amostras foram cortadas para confecção de lâminas de HE e verificadas pela patologista Dra. Renata Coudry. O procedimento para cada lâmina de LCM foi realizado em um tempo máximo de 15 minutos e o número médio de disparos de laser foi de 1500. O RNA foi extraído utilizando o PicoPure RNA Isolation Kit –(Arcturus), sendo que a integridade do RNA total de todas as amostras microdissecadas a laser foi verificada no 2100 Bioanalyzer (Agilent), aparelho com sensibilidade para analisar quantidades de RNA da ordem de picogramas (Figura 11).

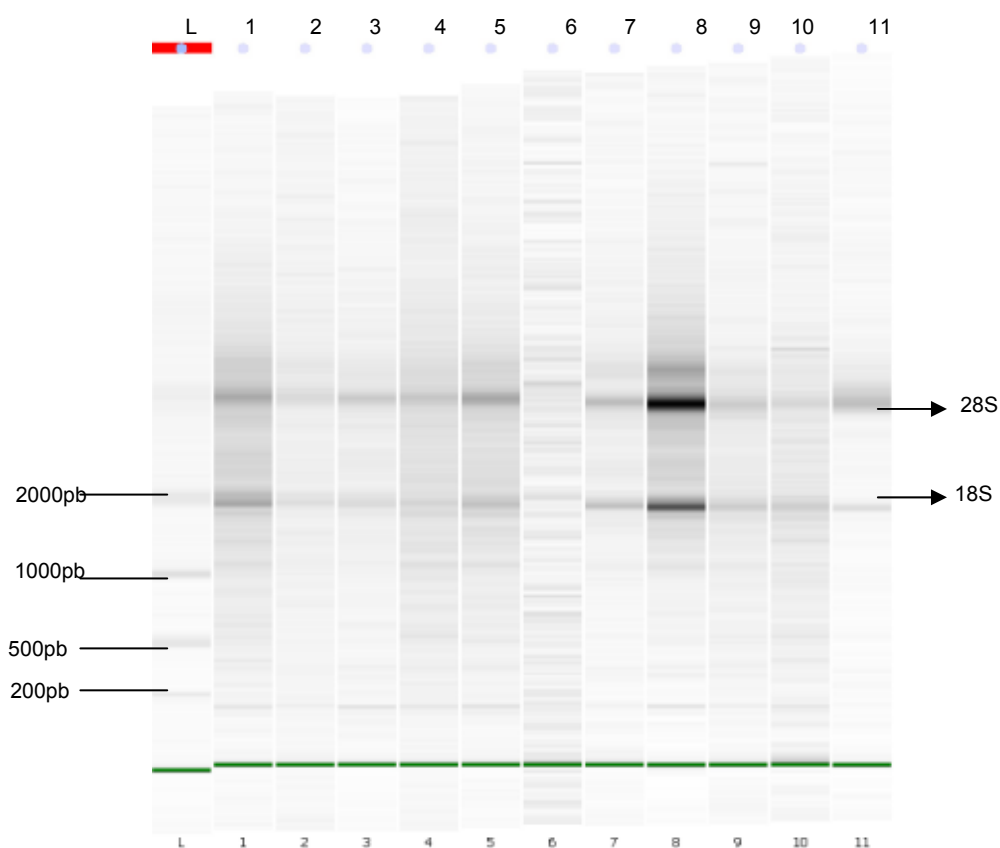


Figura 11 - RNA total analisado no Bioanalyzer. L- Ladder, 1- 6N0, 2- 4N1, 3- 5N1, 4- 14N0, 5- 11N1, 6- 15N0, 7- 1N1, 8- 4N0, 9- 12N1, 10- 25N1 e 11- Referência.

As 47 amostras selecionadas apresentaram RNA total de boa qualidade avaliado pelo RIN, apresentando valores >5 (valor estipulado através de experimentos realizados pelo grupo, como sendo suficiente para gerar RNA amplificado de boa qualidade). Para as amostras que apresentaram RNA de baixa qualidade o procedimento de LCM foi repetido de modo que todas as amostras apresentassem qualidade suficiente para os experimentos seguintes.

Todas as 47 amostras foram amplificadas, utilizando-se o RiboAmp kit (Arcturus), que se baseia na síntese de primeira fita utilizando o iniciador oligo dT-T7 que carrega um sítio de reconhecimento da RNA polimerase permitindo a transcrição *in vitro*, essa amplificação é realizada em dois *rounds* para um maior rendimento final em massa de RNA. Esse Kit foi utilizado a fim de se manter uma menor diferença possível no procedimento de amplificação entre as amostras, pois há uma menor manipulação e, além disso, o tempo do procedimento de amplificação utilizando o kit é bem menor que o tempo necessário para amplificação do protocolo estabelecido no nosso laboratório para esse tipo de amostra.

A qualidade dos aRNAs foi analisada utilizando-se o 2100 Bioanalyzer (Agilent) (Figura 12). Essa análise é realizada para identificar amplificações com qualidade suficiente para experimentos de *microarray*. Normalmente amplificações que resultam em fragmentos muito pequenos (100-200pb) apresentam hibridizações de baixa qualidade, pois representam dímeros de *primers* oligo dT-T7 amplificados.

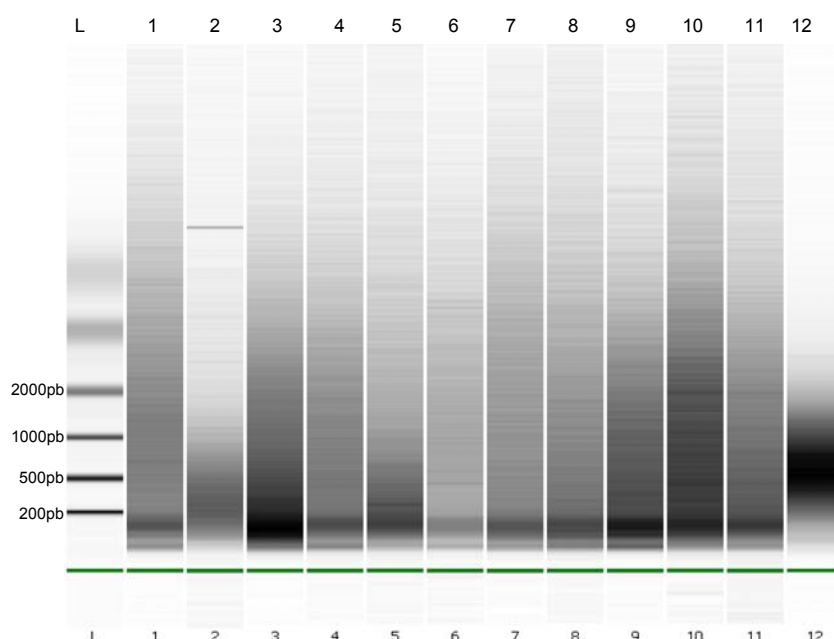


Figura 12 - aRNA analisado no Bioanalyzer. 1- 1N0, 2- 4N0, 3- 6N0, 4- 20N0, 5- 14N0, 6- 15N0, 7- 24N0, 8- 19N0, 9- 41N0, 10- 42N0, 11- 45N0, 12- Referência.

Os resultados de amplificação obtidos por nós utilizando esse kit, apresentaram aRNA com tamanho variado, não apresentando o padrão esperado quando utilizamos o nosso protocolo de amplificação. Para comparação foi utilizado aRNA da linhagem HB4A, sabidamente de boa qualidade, previamente verificada (Figura 12 – linha 12). No perfil de amplificação resultante quando utilizamos o kit apresentou uma alta concentração abaixo de 200pb, que poderia ser sinal de baixa eficiência da amplificação ou uma característica do próprio kit. No entanto, foi observado aRNA de tamanho maior a 200pb, o que sugeriu que esses aRNAs poderiam

ser utilizados para marcação, já que foram as primeiras amplificações utilizando-se esse kit e não tínhamos experiências prévias.

Das 47 amplificações realizadas, 35 apresentaram massa acima de 12µg, quantidade necessária pra realizar a marcação em sistema *dye swap*. Dessas 35 amostras amplificadas, 16 e 19 pertencem ao grupo de amostras sem e com comprometimento linfonodal, respectivamente.

5.2.3 Marcação e Hibridização

Seis microgramas de cada uma das 35 amostras testes e 4µg da amostra referência foram utilizadas para síntese de cDNA marcado com os fluoróforos Cy3 e Cy5 e posteriormente hibridizadas na lâmina vias 2.3K. Na Figura 13 está representado o resultado da hibridização de uma das amostras na plataforma utilizada.

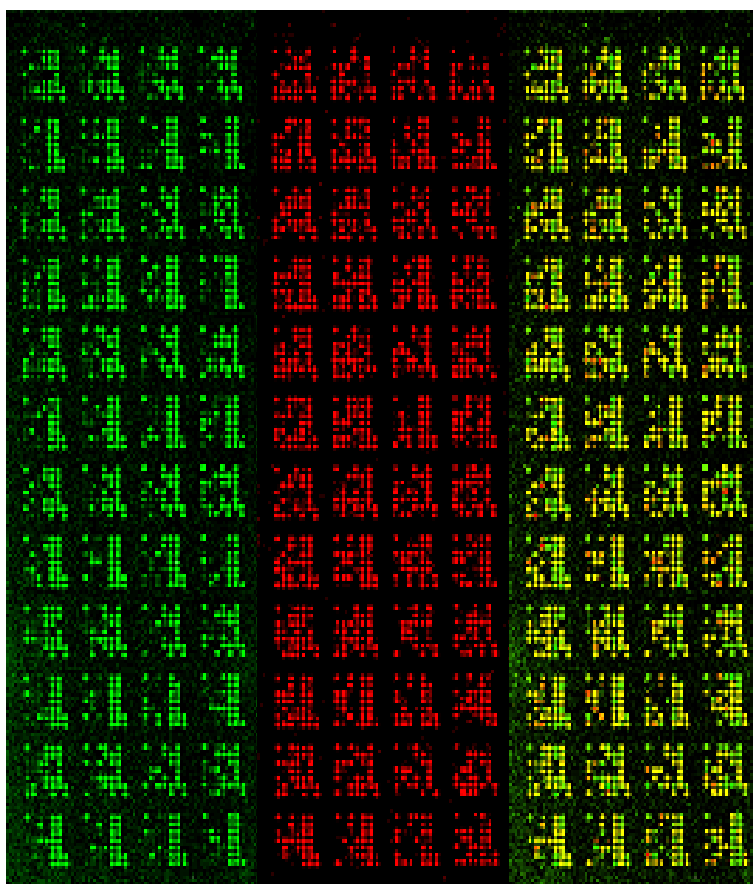


Figura 13 - Representação da lâmina 2.3K Marcação da amostra com Cy3 (verde), referência com Cy5 e composite.

5.2.4 Análise matemática e estatística

Das 35 amostras, 29 amostras tiveram lâminas de qualidade satisfatória apresentando correlação de Pearson entre as lâminas controle e $swap \geq 0.8$ após normalização e foram selecionadas para as análises seguintes. Para essas amostras foi construído um dendograma de correlação onde se verifica a similaridade entre as amostras controle e swap (Figura 14).

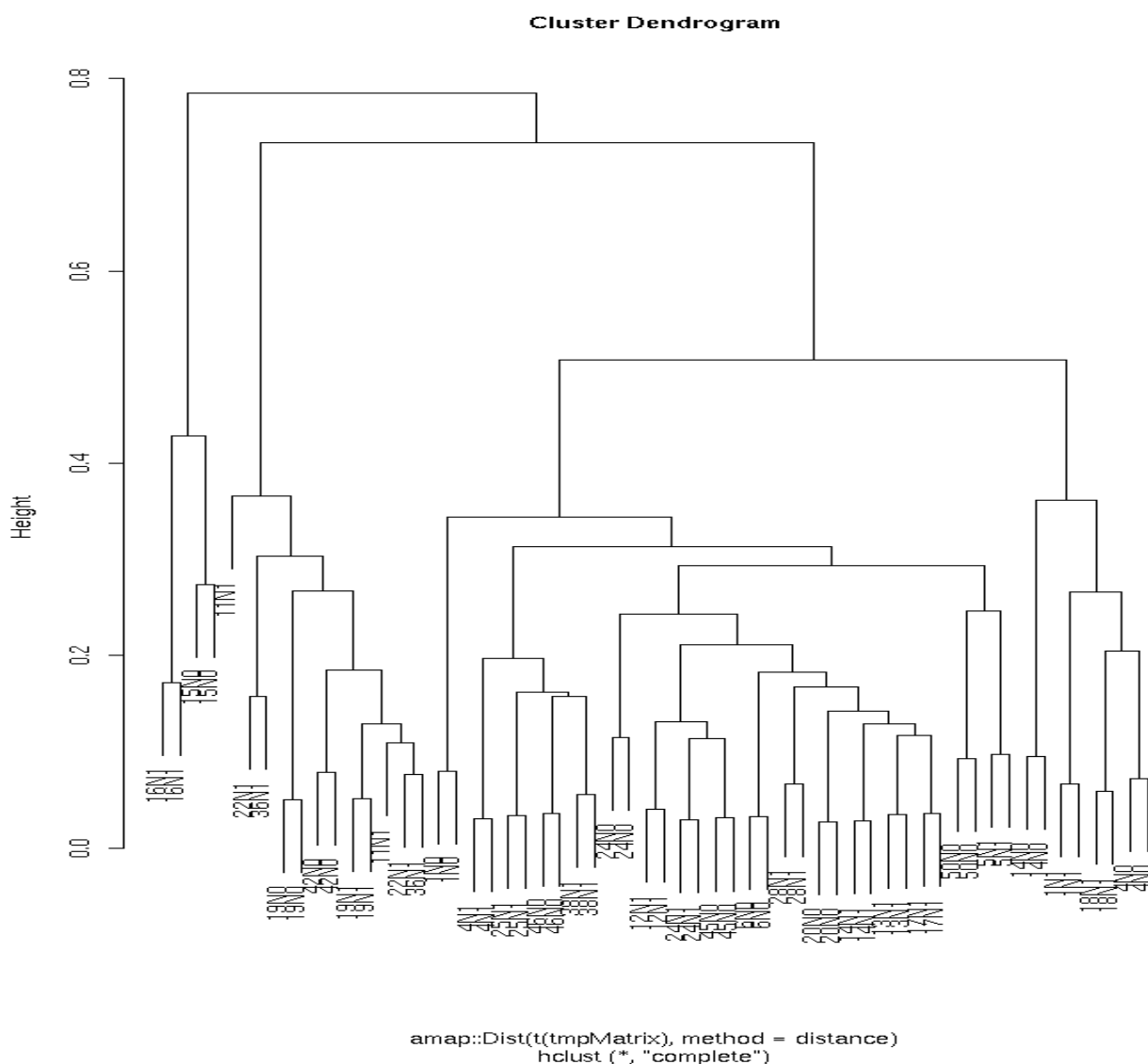


Figura 14 - Dendrograma (correlação- controle-swap) após normalização por LOWESS.

Das 29 amostras selecionadas que apresentaram correlação controle-swap > 0.8 , 26 (89%) apresentaram maior similaridade entre as replicatas da mesma amostra, do que qualquer outra lâmina do estudo, revelando alta reprodutibilidade experimental.

Para a identificação de genes diferencialmente expressos entre os 2 grupos de amostras de nosso estudo foram utilizadas 12 amostras do grupo de linfonodo negativo (N0) (1N0, 4N0, 6N0, 14N0, 15N0, 19N0, 20N0, 24N0, 42N0, 45N0, 46N0, 50N0) e 17 amostras do grupo com linfonodo acometido (N1) foram utilizadas (1N1, 4N1, 5N1, 11N1, 12N1, 13N1, 14N1, 16N1, 17N1, 18N1, 19N1, 22N1, 24N1, 25N1, 28N1, 36N1, 38N1).

Na comparação entre as amostras dos dois grupos, foram identificados 67 genes diferencialmente expressos entre os grupos, que foram selecionados pelo critério estabelecido de $p\text{ value} \leq 0,05$ (Tabela 6), sendo que 8 deles pertencem a as vias de interesse; WNT (*WNT10B*, *SKP1A*, *DAAM1*, *DAAM2* e *PLCB4*), EMT (*CTNNB1* e *KRT19*) e PI3K(*PLCD1*) e 5 (*ALDH1A1*, *BDNF*, *SALL4*, *THRB*, *TUBB*) pertencentes a via do ácido de retinóico também presente na lâmina. O *fold* máximo encontrado entre o grupo com comprometimento e o grupo sem comprometimento de linfonodo foi de 3,48 e o mínimo de -1,35.

Para os 67 genes diferencialmente expressos foi construído um cluster através do método SOM, que foi capaz de discriminar 70% das amostras do grupo com comprometimento de linfonodo e 75% das amostras sem comprometimento de linfonodo, sugerindo um padrão de expressão baseado nesses genes associado ao acometimento linfonodal (Figura 15).

Tabela 6 - Genes diferencialmente expressos ($p \leq 0,05$). Comparação entre o grupo N1 com N0, valores positivos de fold indicam maior expressão do gene no grupo N1.

Gene Symbol	P.value	Fold	Gene Symbol	P.value	Fold
<i>PIP</i>	0,04	3,48	<i>EPHB6</i>	0,03	-1,03
<i>APOD</i>	0,01	2,37	<i>WNT10B</i>	0,05	-1,03
<i>CN484301</i>	0,02	2,04	<i>BDNF</i>	0,04	-1,04
<i>T09065</i>	0,02	1,77	<i>CR2</i>	0,05	-1,04
<i>SPATA11</i>	0,04	1,74	<i>MOG</i>	0,01	-1,05
<i>GSTM3</i>	0,01	1,74	<i>PLCD1</i>	0,05	-1,05
<i>VPS24</i>	0,00	1,68	<i>ADAM23</i>	0,03	-1,05
<i>LOC648326</i>	0,01	1,64	<i>PLCB4</i>	0,05	-1,05
<i>LOC387930</i>	0,01	1,63	<i>SALL4</i>	0,03	-1,05
<i>VWF</i>	0,00	1,61	<i>DAAM1</i>	0,03	-1,06
<i>NPNT</i>	0,04	1,53	<i>PHACTR2</i>	0,05	-1,06
<i>AZGP1</i>	0,00	1,52	<i>THRB</i>	0,01	-1,07
<i>BX538341</i>	0,00	1,46	<i>KLK3</i>	0,03	-1,08
<i>RPL32</i>	0,02	1,43	<i>PHB2</i>	0,04	-1,12
<i>AB209172</i>	0,03	1,39	<i>ASNS</i>	0,05	-1,14
<i>TSPAN1</i>	0,04	1,37	<i>DNAJB1</i>	0,05	-1,15
<i>C1S</i>	0,02	1,33	<i>KRT19</i>	0,05	-1,15
<i>RAI17</i>	0,03	1,33	<i>TUBB</i>	0,04	-1,15
<i>TMEM98</i>	0,04	1,25	<i>FLJ21439</i>	0,01	-1,16
<i>CHMP4A</i>	0,04	1,24	<i>GPNMB</i>	0,01	-1,16
<i>PCBP1</i>	0,03	1,21	<i>PSMB5</i>	0,00	-1,19
<i>CTNNB1</i>	0,04	1,21	<i>PPP1CA</i>	0,05	-1,21
<i>DST</i>	0,04	1,18	<i>SNRPB</i>	0,00	-1,23
<i>MAP3K1</i>	0,02	1,18	<i>KTN1</i>	0,00	-1,24
<i>RPS9</i>	0,04	1,18	<i>FLJ20257</i>	0,01	-1,26
<i>HSP90AA1</i>	0,03	1,16	<i>DERP6</i>	0,00	-1,26
<i>SDHB</i>	0,03	1,15	<i>S100A6</i>	0,01	-1,27
<i>ZBTB4</i>	0,02	1,15	<i>PSMD13</i>	0,00	-1,27
<i>FRY</i>	0,04	1,15	<i>GNAS</i>	0,03	-1,29
<i>SLC39A8</i>	0,02	1,11	<i>TNRC15</i>	0,02	-1,29
<i>PRKCBP1</i>	0,04	1,11	<i>GNB1</i>	0,00	-1,35
<i>DAAM2</i>	0,03	1,09			
<i>PHYHIP</i>	0,02	1,08			
<i>ALDH1A1(3)</i>	0,02	1,07			
<i>SKP1A</i>	0,05	1,05			

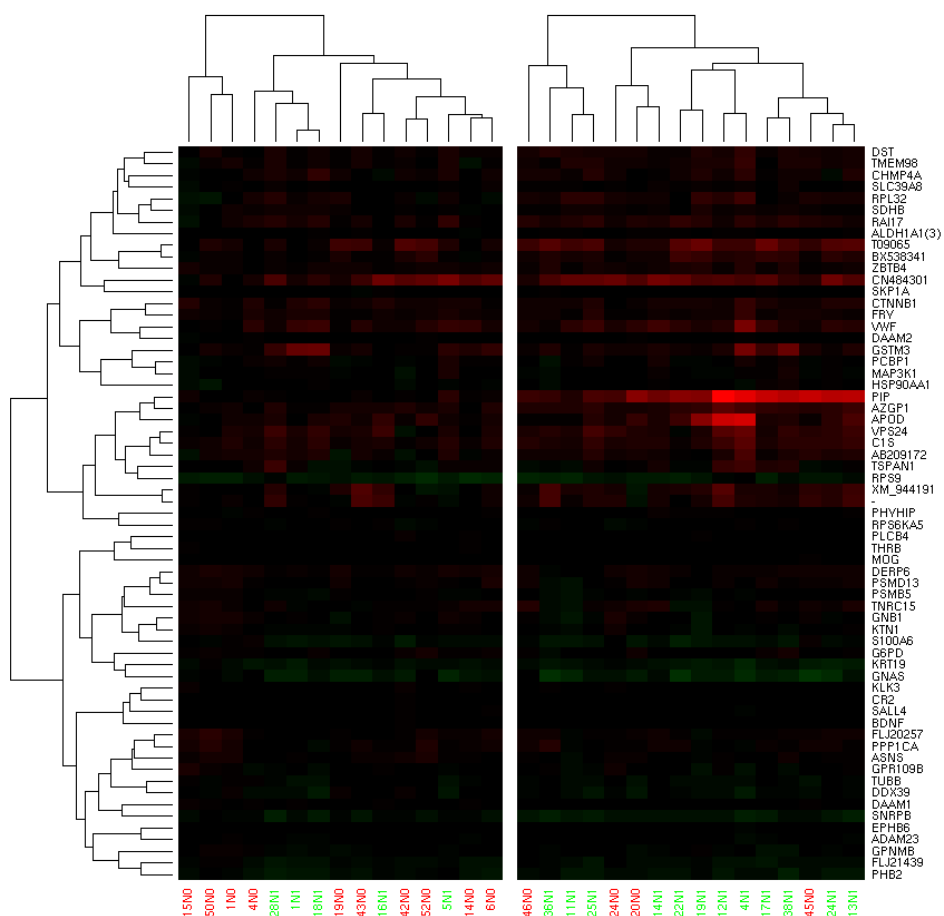


Figura 15 - Cluster (método SOM) dos genes diferencialmente expressos ($p \leq 0,05$).

Entre os genes diferencialmente expressos ($p \leq 0,05$) pertencentes às vias de interesse desse estudo, cinco se apresentaram mais expresso no grupo N0 (*KRT19*, *WNT10B*, *PLCD1*, *PLCB4*, *DAAM1*) e 3 se apresentaram mais expressos no grupo N1 (*CTNNB1*, *DAAM2*, *SKP1A*).

Alguns desses genes apresentam envolvimento no câncer previamente relatado na literatura, além dos outros, pouco descritos, e que

foram relatados nesse estudo, sendo novos candidatos a estarem envolvidos com a progressão do carcinoma ductal de mama.

Gene - *CTNNB1- catenin (cadherin-associated protein), beta 1* (β -catenina).

Esse gene apresentou super expressão no grupo de amostras com comprometimento linfonodal.

β -catenina está presente na junção aderente, é responsável pelo estabelecimento e manutenção da camada epitelial, pois ancora a E-caderina ao citoesqueleto de actina, impedindo que em células normais adultas, a β -catenina migre para o núcleo e regule fatores de transcrição (TCF/LEF), responsáveis pelo crescimento e sobrevivência celular. Apresenta grande influência no desenvolvimento embrionário, além disso, está fortemente relacionada ao ganho de potencial metastático, devido à perda do contato célula-célula e migração para o núcleo onde atua sobre fatores de transcrição e conseqüente ganho de motilidade (<http://www.ncbi.nlm.nih.gov>).

Alguns trabalhos vêm demonstrando o aumento da expressão de β -catenina em câncer de mama. PARK et al. (2007) avaliaram a expressão dessa proteína em 196 casos de IDC primário com seus respectivos linfonodos axilares acometidos, e verificou uma maior expressão de β -catenina nesses linfonodos em relação ao tumor primário e *downregulation* no tumor primário e linfonodo em relação ao tecido normal. BUKHOLM et al. (2000), também encontraram resultados semelhantes usando esse mesmo tipo de comparação mostrando uma *downregulation* de β -catenina em

tumores primários, quando comparados ao tecido normal, e uma re-expressão dessa proteína em fases mais avançadas do tumor (metástase).

Nesse sentido nossos dados corroboram com esses estudos, pois houve um aumento da expressão desse gene no grupo de pacientes com linfonodo acometido. A re-expressão de moléculas de adesão (E-caderina, β -catenina, α -catenina) tem sido verificada em estágios mais avançados do tumor de mama (BUKHOLM et al. 2000; PARK et al. 2007) o mecanismo que atua para o aumento dessa expressão ainda não está elucidado, mas um argumento para essa re-expressão é que essas proteínas vão capacitar as células tumorais a colonizar outros meios. Falha na função de alguns genes pode estimular a expressão de cateninas, como por exemplo, mutação na E-caderina. A alta expressão pode desequilibrar a fosforilação da β -catenina por GSK3 β , essa fosforilação direciona a β -catenina ao proteossomo para ser degradada, ocorrendo um desequilíbrio dessa atividade por excesso de β -catenina, ela deixa de ser degradada e se direciona ao núcleo, onde vai agir sobre fator de transcrição (PARK et al. 2007).

Gene: *DAAM2* e *DAAM1*-- *dishevelled associated activator of morphogenesis*.

O gene *DAAM2* apresentou super expressão no grupo de amostras com comprometimento linfonodal e o gene *DAAM1* foi menos expresso no grupo de amostras com comprometimento linfonodal

As proteínas Damm1 e Daam2 pertencem a família das proteínas Formin homology (FH), e estão implicados na polaridade celular em

humanos. A ativação da via Wnt, via ativação do complexo Wnt/Fz - Rho (via Wnt) requer proteínas citoplasmáticas Dishevelled (Dvl) além da proteína Daam que se liga a Dvl e Rho, e faz a mediação do complexo Wnt-Dvl-Rho. A inibição ou perda de Daam evita a ativação de Rho por Wnt/Fz. Rho por sua vez vai atuar na reorganização do citoesqueleto de actina, Assim o aumento de Daam pode sinalizar a ativação da via Wnt via complexo Wnt/Fz – Rho, levando a ativação aberrante dessa via de sinalização em câncer, resultando em fenótipos mais agressivos, como polaridade anormal, invasão e metástase (KATOH 2005).

DAAM1 mostrou resultado de expressão contraditório a *DAAM2*, que foi super expresso no grupo com comprometimento de linfonodo, o que não era esperado, pois até o momento parece que as duas proteínas exercem a mesma função. No entanto esse dado deve ser investigado mais profundamente.

Gene - *PLCD1- phospholipase C, delta 1*

Esse gene foi menos expresso no grupo de amostras com comprometimento linfonodal.

A proteína codificada por esse gene está associada a parada do ciclo celular na fase G1 - S (*checkpoint*), por regular positivamente p21. Esta transição é particularmente importante no estudo de células tumorais, pois é na fase S que se dá a replicação do DNA, uma das fases onde existe maior probabilidade de ocorrerem mutações. A proteína p21 se liga às ciclinas D e E, inibindo as quinases dependentes destas ciclinas (cdk2 e cdk4)

necessárias para a progressão do ciclo celular. Devido à sua função e intercooperação com a proteína p53, a p21 é considerada uma proteína supressora de tumores. Além disso, *PLCD1* regula negativamente Akt, que está envolvido em múltiplas funções celulares incluindo proliferação, sobrevivência e migração (CANTLEY 2002; VIVANCO e SAWYERS 2002) (<http://www.genecards.org/>).

Assim, nossos dados sugerem uma maior proliferação celular nos tumores primários dos pacientes que apresentaram linfonodo acometido, ocasionado pela menor expressão desse gene que controla negativamente o ciclo celular.

Além desses genes com possível envolvimento na progressão tumoral, outros genes das vias (PI3K, Wnt e EMT) também foram diferencialmente expressos ($p \leq 0,05$).

Gene - *SKP1*- S-phase kinase-associated protein 1.

Esse gene apresentou super expressão no grupo de amostras com comprometimento linfonodal.

Este gene codifica uma proteína F-box, que funciona como um substrato de reconhecimento para a ubiquitina. Essa proteína participa da rede de proteínas que controlam o nível de β -catenina e afeta o nível de atividade dos fatores de transcrição dependentes de β -catenina (<http://www.ncbi.nlm.nih.gov>). Em uma avaliação simples, esse dado de certa forma contraria o resultado encontrado no nível de expressão aumentado de β -catenina no grupo com comprometimento linfonodal, pois o

gene *SKP1* tem ação indireta na degradação β -catenina. No entanto, como essa via tem vários fatores reguladores, esse gene por si só não deve ser capaz de diminuir a expressão de β -catenina. Além disso, o aumento de expressão desse gene pode ser uma estratégia para controlar a expressão de β -catenina como resultado da busca de um equilíbrio celular. Essas colocações, são apenas especulações frente a nossos dados obtidos, confirmações através de experimentos funcionais são necessárias.

Gene symbol- *WNT10B*- wingless-type MMTV integration site family, member 10B.

Esse gene foi menos expresso no grupo de amostras com comprometimento linfonodal.

A família de genes *WNT* codifica proteínas de sinalização. Essas proteínas têm sido implicadas em processos de desenvolvimento e na oncogênese. *WNT10B* está envolvida em câncer de mama por estar relacionada ao controle da adipogênese.

Era esperado um aumento da expressão desse gene no grupo com linfonodo comprometido, já que é uma proteína de ativação da via Wnt, porém outras proteínas Wnt também exercem essa atividade e podem estar ativando essa via (<http://www.ncbi.nlm.nih.gov>).

Além dos genes envolvidos nas vias de interesses, outros genes com expressão alterada entre os grupos foram encontrados e não deixam de ser importantes na busca de marcadores de invasão e metástase.

Gene - *PIP*- *prolactin-induced protein*

Esse gene apresentou a maior super expressão de todos os genes da plataforma ($p \leq 0,05$) no grupo de amostras com comprometimento linfonodal (*fold* 3,14).

PIP é um gene expresso em glândulas exócrinas, em condições patológicas, é expresso em câncer de mama, apresentando característica apócrina. Está localizado no braço longo do cromossomo 7, uma região freqüentemente alterada em tumor de mama. AUTIERO et al. (2002) analisaram a estrutura do gene *PIP* em câncer de mama e notaram sua amplificação e formação de DNA extracromossômico em 21,4% dos casos analisados, sugerindo um aumento da instabilidade desse gene.

Nossos dados indicam que esse gene pode estar expresso em estágios mais avançados do câncer de mama, devido à maior expressão (*fold* 3,14) das amostras com linfonodo comprometido em relação às amostras sem linfonodo comprometido, ainda não se tem muitos trabalhos relacionando esse gene a progressão tumoral. Nesse estudo, este gene apresentou-se como candidato a estar envolvido na progressão do câncer de mama.

Gene - *ADAM23*- ADAM metallopeptidase domain 23

Esse gene foi menos expresso no grupo de amostras com comprometimento linfonodal.

Esse gene codifica uma proteína membro da família ADAM e está envolvido na adesão célula-célula e interação célula-matriz e alterações de adesão celular é um processo importante na progressão tumoral.

COSTA et al. (2004) utilizaram linhagens celulares de câncer de mama e tumor primário e analisaram a presença de metilação desse gene. Foi visto que a menor expressão desse gene estava relacionada com hipermetilação do promotor, sendo que 66,7% das linhagens celulares tumorais e 69,2% das amostras de câncer de mama, apresentaram hipermetilação no promotor desse gene. Também foi visto que tumores com grau mais avançado apresentaram maior nível de metilação, sugerindo que a molécula de adesão ADAM23 deve estar implicada na progressão do câncer de mama. Esses dados corroboram com nossos resultados, indicando uma menor expressão desse gene em estágios mais avançados da doença, nesse caso, pacientes com comprometimento de linfonodo.

Além desses genes citados, outros genes que apresentaram maiores *folds* entre os grupos, se mostram bastante interessante para uma possível validação, mesmo que até o momento não apresente nenhuma função relacionada a progressão tumoral, são candidatos que devem ser avaliados.

Comparando nosso set de genes presentes na lâmina vias 2.3K com os 70 genes identificados no trabalho da VAN'T VEER et al. (2002), classificados como indicadores de pior prognóstico em pacientes sem comprometimento linfonodal, identificamos 6 genes em comum. Esses 6 genes foram utilizados para a construção de um cluster (método SOM) para analisar o comportamento das amostras desse estudo. Nosso resultado

mostrou que não ocorreu uma separação eficiente entre os dois grupos (58% do grupo N0, e 66% do grupo N1). Entre os 6 genes em comum, apenas 1 (*GSTM3*) se apresentou diferencialmente expresso ($p \leq 0,05$) na comparação entre os grupos N1 e N0, no entanto nosso resultado foi diferente ao encontrado por VAN'T VEER et al. (2002). No nosso estudo o gene *GSTM3* se mostrou mais expresso no grupo N1, e no trabalho da VAN'T VEER et al. (2002) esse gene foi mais expresso no grupo N0 que não apresentou metástase. Essa diferença encontrada pode ser devido a heterogeneidade molecular dos sets das amostras utilizadas nos dois estudos. Nesse trabalho foram utilizadas apenas amostras ER positivo, sendo que VAN'T VEER et al. (2002) utilizaram amostras ER positivo e negativo. Trabalhos atuais mostram que o *status* hormonal é a principal característica clínico-patológica que se associa a padrões de expressão gênica em câncer de mama (SORLIE et al. 2001; VAN'T VEER et al. 2002; SOTIRIOU et al. 2003) e pode estar influenciando o nível de expressão desse gene. Além disso, nesse estudo todas as amostras foram microdissecadas a *laser*, o que permite uma análise de expressão mais fidedigna das células tumorais, pois apenas essas células são capturadas, No estudo da VAN'T VEER et al. (2002) foi realizada microdissecção manual (mínimo de 50% de tumor na amostra) que resulta na análise de expressão gênica não apenas das células tumorais, mas de todas as células presentes no microambiente tumoral.

Além disso, nesse trabalho, foi empregado o conceito de redes de relevância, na qual se busca pares de genes cujo padrão de expressão

gênica apresenta-se linearmente correlacionado, procuramos quebra de correlação entre os pares de genes nos dois grupos de amostras, para isso identificamos as correlações entre pares de genes em cada grupo de amostras e comparamos o valor de correlação entre eles.

Assim, após essa comparação identificamos pares de genes cujos coeficientes de correlação foram significativamente alterados ($p \leq 0,05$) entre os grupos de amostras, sem e com comprometimento linfonodal.

As análises das redes de relevância foram aplicadas apenas entre genes pertencentes às vias de interesse nesse estudo (Wnt, PI3K e EMT). Os valor de correlação dos níveis de expressão dos pares de genes pertencentes a cada via isoladamente, foram comparados entre os grupos de amostra, além de uma rede de relevância com os genes de uma via contra os genes das 2 vias restantes, com o objetivo de identificar associações ainda não descritas. Para isso foi obtido os coeficientes de correlação entre a expressão de 2 genes utilizando os valores de expressão das amostras dentro de cada grupo. A partir destes valores, foi construída uma rede de relevância por representação gráfica, atribuindo coloração verde para correlação positiva e, vermelha para correlação negativa.

Para comparar as alterações entre os coeficientes de correlação, entre os dois grupos (com e sem comprometimento de linfonodo), foi aplicado teste Z de Fisher, onde a diferença entre os valores de correlação para cada par de genes, nos dois tipos biológicos estudados, foi transformada em uma nova variável Z, que apresenta distribuição normal aproximada. E usamos esta informação para aferir a significância da

diferença de correlação entre os grupos ao aplicarmos o teste de hipótese para esta variável Z.

Assim, os pares de genes que apresentaram diferença de correlação estatisticamente significativa estão apresentados nas Figuras 16 a 19.

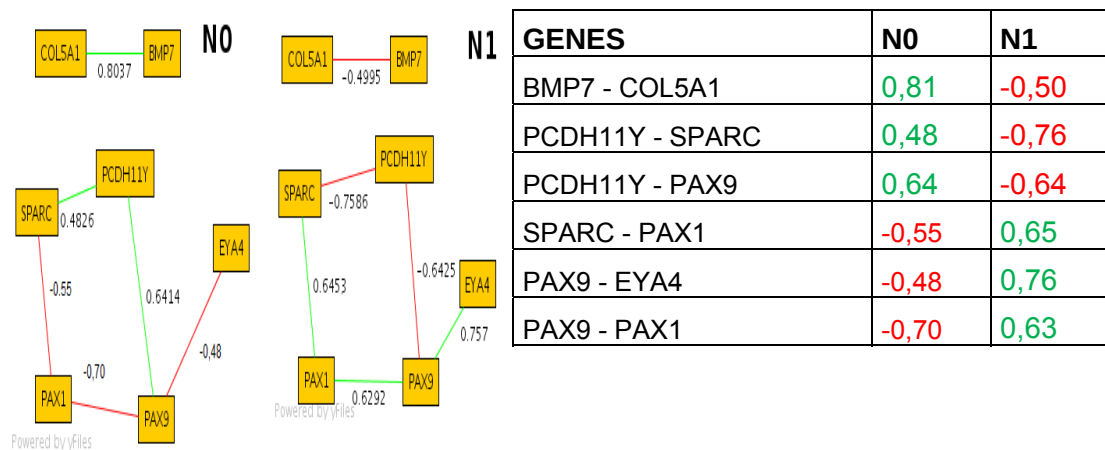
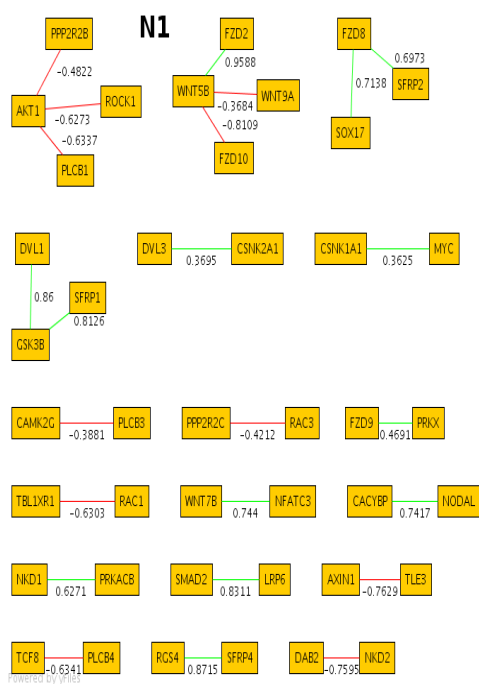
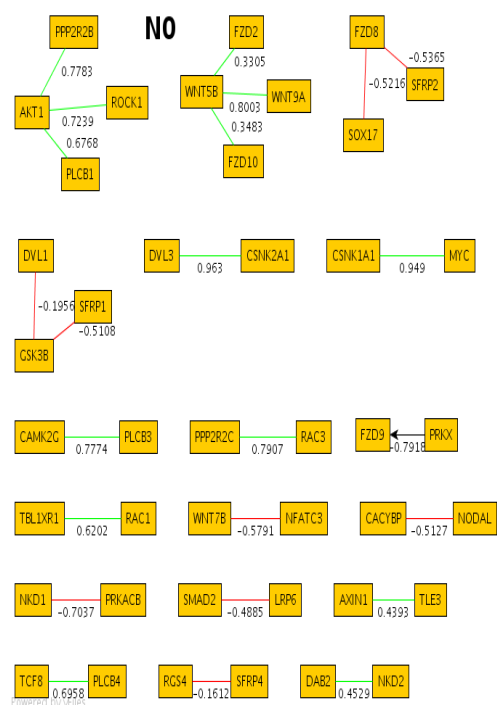


Figura 16 - Rede de relevância para genes pertencentes ao processo EMT. N0-grupo sem comprometimento de linfonodo, N1- grupo com acometimento de linfonodo.



Genes	N0	N1
PLCB4 - TCF8	0,70	-0,63
MYC - CSNK1A1	0,95	0,63
NFATC3 - WNT7B	-0,58	0,74
PPP2R2B - AKT1	0,78	-0,48
DVL1 - GSK3B	-0,20	0,86
NKD2 - DAB2	0,45	-0,76
NODAL - CACYBP	-0,51	0,74
SFRP1 - GSK3B	-0,51	0,81
TLE3 - AXIN1	0,44	-0,76
RAC3 - PPP2R2C	0,79	-0,42
ROCK1 - AKT1	0,72	-0,63
AKT1 - PLCB1	0,68	-0,64
PRKX - FZD9	-0,79	0,47
CSNK2A1 - DVL3	0,96	0,37
LRP6 - SMAD2	-0,49	0,83
FZD2 - WNT5B	0,33	0,96
SFRP4 - RGS4	-0,16	0,87
RAC1 - TBL1XR1	0,62	-0,63
FZD10 - WNT5B	0,35	-0,81
PRKACB - NKD1	-0,70	0,63
FZD8 - SOX17	-0,52	0,71
FZD8 - SFRP2	-0,54	0,70
PLCB3 - CAMK2G	0,78	-0,39
WNT9A - WNT5B	0,80	-0,37

Figura 17 - Redes de relevância para a via Wnt. N0-grupo sem comprometimento de linfonodo, N1- grupo com acometimento de linfonodo.

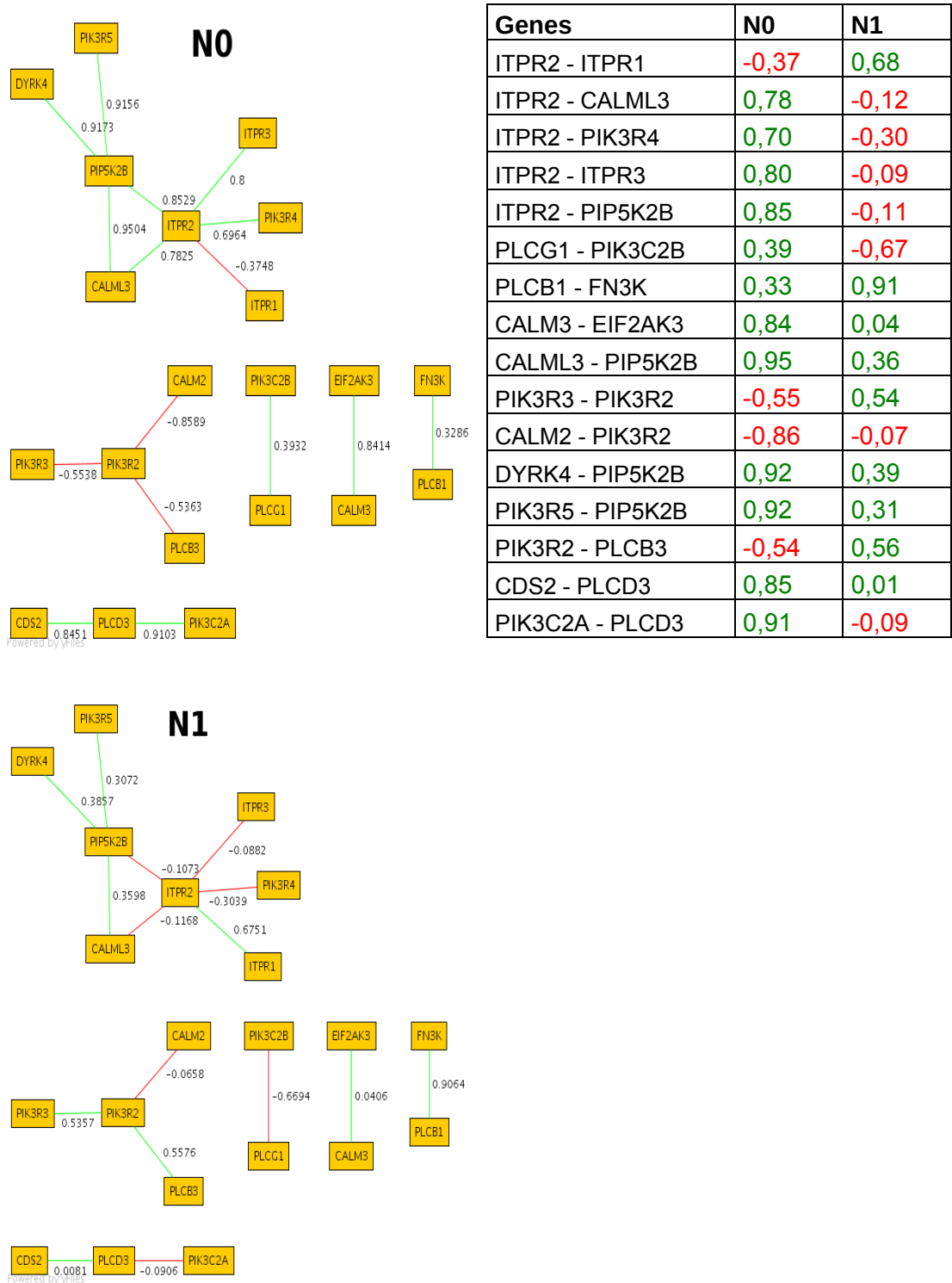
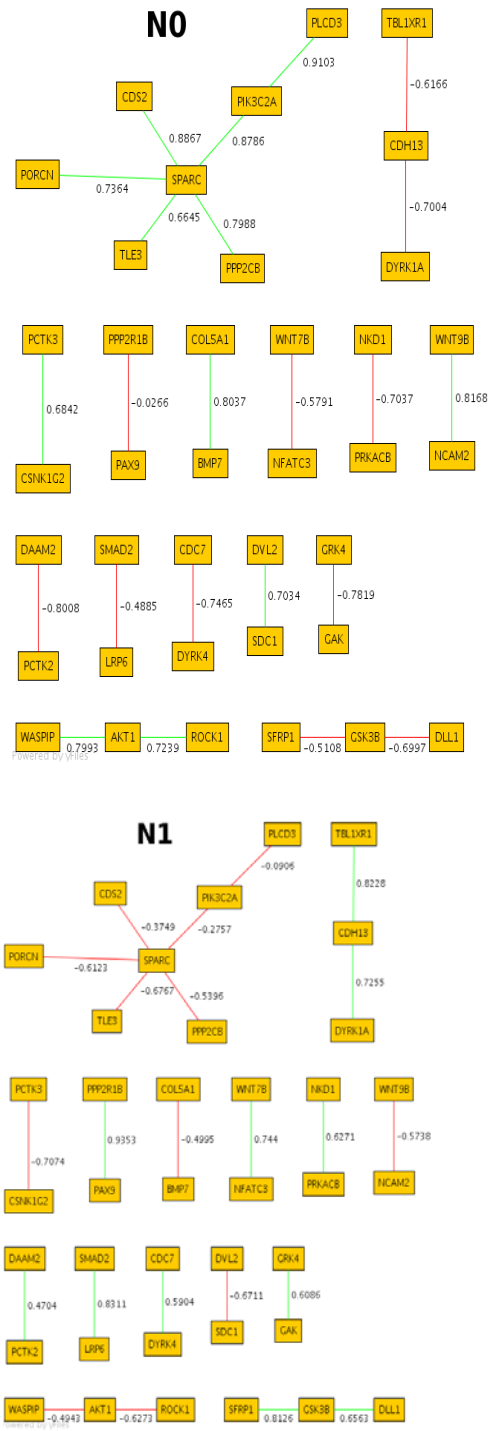


Figura 18 - Redes de relevância para a via PI3K. N0-grupo sem comprometimento de linfonodo, N1- grupo com acometimento de linfonodo.



Genes	N0	N1
NCAM2 - WNT9B	0,82	-0,57
GAK - GRK4	-0,79	0,61
WASPIP - AKT1	0,80	-0,49
BMP7 - COL5A1	0,80	-0,50
NFATC3 - WNT7B	-0,58	0,74
CDH13 - TBL1XR1	-0,62	0,82
CDH13 - DYRK1A	-0,70	0,73
PCTK2 - DAAM2	-0,80	0,47
PORCN - SPARC	0,74	-0,61
SFRP1 - GSK3B	-0,51	0,81
TLE3 - SPARC	0,66	-0,68
ROCK1 - AKT1	0,72	-0,63
SPARC - CDS2	0,89	-0,37
SPARC - PIK3C2A	0,88	-0,28
SPARC - PPP2CB	0,80	-0,54
LRP6 - SMAD2	-0,49	0,83
DYRK4 - CDC7	-0,75	0,59
PAX9 - PPP2R1B	-0,03	0,94
SDC1 - DVL2	0,70	-0,67
DLL1 - GSK3B	-0,70	0,66
PRKACB - NKD1	-0,70	0,63
CSNK1G2 - PCTK3	0,68	-0,71
PIK3C2A - PLCD3	0,92	-0,09

Figura 19 - Redes de relevância para EMT e vias WNT e PI3K. N0-grupo sem comprometimento de linfonodo, N1 - grupo com acometimento de linfonodo.

A partir da obtenção desses gráficos foi avaliado o *scatter-plot* de cada par de genes cujo valor da diferença de correlação entre os genes dos 2 grupos tenha sido $>0,8$ para selecionar as melhores duplas de genes com expressão linearmente correlacionadas. Nenhum deles sugeriu associação no nível de expressão do par de genes analisado entre os grupos de amostras (N0 e N1). Provavelmente o fator limitante para analisar essas diferenças de correlações entre os genes dos dois grupos, foi o pequeno número de amostras, além de provavelmente existir diferenças mais tênues entre os dois grupos, uma vez são tumores de mama que se diferem apenas pelo *status* linfonodal. Ao analisar o *scatter-plot*, o que se nota, é que a correlação linear entre os pares de genes dos dois grupos é extremamente débil e não permite sugerir associação de expressão que foi perdida entre os pares de genes a algum fenômeno biológico (invasão e metástase). Exemplos desses gráficos estão apresentados na Figura 20.

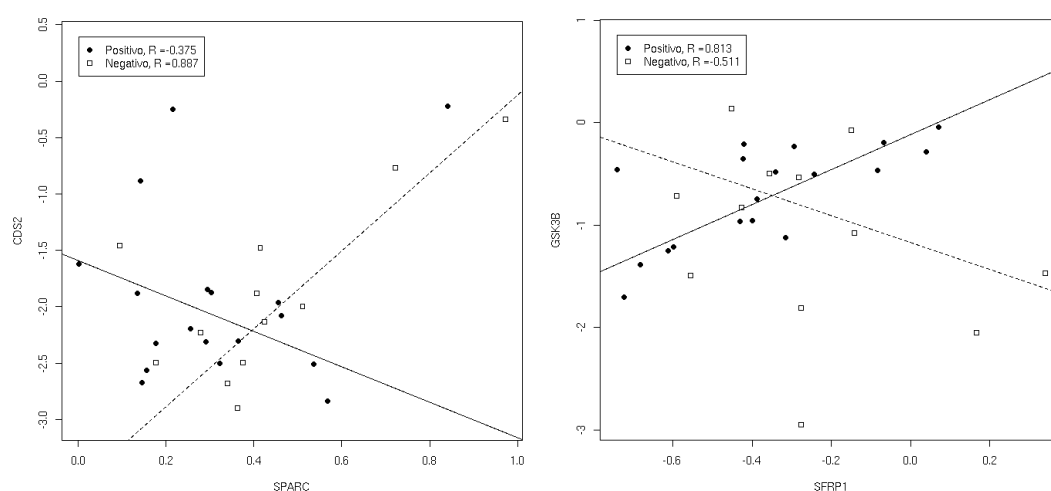


Figura 20 - Scatter-plot da correlação entre os pares de genes CDS2 e SPARC e entre os pares de genes GSK3β e SFRP1 nos grupos N0 e N1.

6 CONCLUSÃO

Nossos dados mostraram que as amostras emblocadas em parafina, não apresentaram RNA de qualidade suficiente para avaliação de expressão gênica através de experimentos de *microarray* nas condições testadas, porém para avaliação de expressão gênica através de RT-PCR esse tipo de amostra se mostra bastante promissor.

Amostras congeladas de mama se mostraram adequadas para microdissecção a laser, para amplificação de RNA e utilização na plataforma customizada 2.3K.

Na comparação entre o grupo com comprometimento de linfonodo e o grupo sem comprometimento de linfonodo, foram encontrados 67 genes diferencialmente expressos com *P* valor estatisticamente significativo.

Utilizando-se os 67 genes diferencialmente expressos encontrados, foi possível discriminar razoavelmente bem os dois grupos de amostra (70% grupo N1 e 75% grupo N0). Sugerindo que o padrão de expressão está associado com o comprometimento linfonodal.

Foram construídas redes de relevância, e encontramos significância na diferença de correlação entre vários pares de genes dos dois grupos. No entanto, ao se analisar o *scatter-plot* desses genes, não foi possível encontrar nenhum par de genes promissor, provavelmente por não termos um número grande de amostras no estudo.

Foram revelados genes candidatos a estarem envolvidos no processo de invasão e metástase, alguns anteriormente relacionados à progressão da doença (*DAAM2*, *DAAM1*, *CTNNB1*, *PLCD1* e *ADAM23*) e grande parte dos outros genes diferencialmente expressos que são candidatos a marcadores moleculares e que ainda não foram relacionados à progressão e metástase.

7 REFERÊNCIAS BIBLIOGRÁFICAS

Aberle H, Bauer A, Stappert J, Kispert A, Kemler R. beta-catenin is a target for the ubiquitin-proteasome pathway. **EMBO J** 1997; 16:3797-804.

Ackland ML, Newgreen DF, Fridman M, et al. Epidermal growth factor-induced epithelio-mesenchymal transition in human breast carcinoma cells. **Lab Invest** 2003; 83:435-48.

Andreasen PA, Kjoller L, Christensen L, Duffy MJ. The urokinase-type plasminogen activator system in cancer metastasis: a review. **Int J Cancer** 1997; 72:1-22.

Autiero M, Camarca A, Ciullo M, et al. Intragenic amplification and formation of extrachromosomal small circular DNA molecules from the PIP gene on chromosome 7 in primary breast carcinomas. **Int J Cancer** 2002; 99:370-7.

Ball CA, Chen Y, Panavally S. An introduction to *microarrays* bioinformatics. In: Bowtell D, Sambrook J, editors. **DNA microarrays: a molecular cloning manual**. New York: Cold Spring Harbor Laboratory Press; 2003. p.509-94.

Barbareschi M, Buttitta F, Felicioni L, et al. Different prognostic roles of mutations in the helical and kinase domains of the PIK3CA gene in breast carcinomas. **Clin Cancer Res** 2007; 13:6064-9.

Batlle E, Sancho E, Franci C, et al. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. **Nat Cell Biol** 2000; 2:84-9.

Bezwodna WR, Esser JD, Dansey R, Kessel I, Lange M. The value of estrogen and progesterone receptor determinations in advanced breast cancer. Estrogen receptor level but not progesterone receptor level correlates with response to tamoxifen. **Cancer** 1991; 68:867-72.

Bloom ND, Tobin EH, Schreiber B, Degenshein GA. The role of progesterone receptors in the management of advanced breast cancer. **Cancer** 1980; 45:2992-7.

Bolos V, Peinado H, Perez-Moreno MA, Fraga MF, Esteller M, Cano A. The transcription factor Slug represses E-cadherin expression and induces epithelial to mesenchymal transitions: a comparison with Snail and E47 repressors. **J Cell Sci** 2003; 116:499-511.

Boyer B, Valles AM, Edme N. Induction and regulation of epithelial-mesenchymal transitions. **Biochem Pharmacol** 2000; 60:1091-9.

Brennan KR, Brown AM. Wnt proteins in mammary development and cancer. **J Mammary Gland Biol Neoplasia** 2004; 9:119-31.

Brentani RR, Carraro DM, Verjovski-Almeida S, et al. Gene expression arrays in cancer research: methods and applications. **Crit Rev Oncol Hematol** 2005; 54:95-105.

Bukholm IK, Nesland JM, Borresen-Dale AL. Re-expression of E-cadherin, alpha-catenin and beta-catenin, but not of gamma-catenin, in metastatic tissue from breast cancer patients [see comments] **J Pathol** 2000; 190:15-9.

Butte AJ, Tamayo P, Slonim D, Golub TR, Kohane IS. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. **Proc Natl Acad Sci U S A** 2000; 97:12182-6.

Cantley LC. The phosphoinositide 3-kinase pathway. **Science** 2002; 296: 1655-7.

Carter CL, Allen C, Henson DE. Relation of tumor size, lymph node status, and survival in 24,740 breast cancer cases. **Cancer** 1989; 63:181-7.

Christofori G, Semb H.. The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene. **Trends Biochem Sci** 1999; 24:73-6.

Churchill GA. Fundamentals of experimental design for cDNA microarrays. **Nature Genet Suppl** 2002; 32:490-4.

Comijn J, Berx G, Vermassen P, et al. The two-handed E box binding zinc finger protein SIP1 down regulates E-cadherin and induces invasion. **Mol Cell** 2001; 7:1267-78.

Conacci-Sorrell M, Simcha I, Ben-Yedidia T, Blechman J, Savagner P, Ben-Ze'ev A. Autoregulation of E-cadherin expression by cadherin-cadherin interactions: the roles of beta-catenin signaling, Slug, and MAPK. **J Cell Biol** 2003; 163:847-57.

Conacci-Sorrell M, Zhurinsky J, Ben-Ze'ev A. The cadherin-catenin adhesion system in signaling and câncer. **J Clin Invest** 2002; 109:987-91.

Costa FF, Verbisck NV, Salim AC, et al. Epigenetic silencing of the adhesion molecule ADAM23 is highly frequent in breast tumors. **Oncogene** 2004; 23:1481-8.

Cotran RS, Robbins SL, Kumar V. **Robbins pathologic basis of disease.** 5th ed. Philadelphia: W B Saunders; 1999. Neoplasia: tumor suppressor gene; p.298-306.

Cox SC, Walker DM. Epithelial growth fraction and expression of p53 tumour suppressor gene in oral submucous fibrosis. **Aust Dent J** 1996; 41:91-6.

DCIS. **Ductal carcinoma in situ**. Available from: <URL:<http://www.e-radiography.net/radpath/d/dcis.htm>> [2008 abr 18]

Dudoit S, Gentleman RC, Quackenbush J. Open source software for the analysis of microarray data. **Biotechniques** 2003; Suppl:45-51.

Ernster VL, Barclay J, Kerlikowske K, Grady D, Henderson C. Incidence of and treatment for ductal carcinoma in situ of the breast. **JAMA** 1996; 275:913-8.

Ernster VL, Barclay J. Increases in ductal carcinoma in situ (DCIS) of the breast in relation to mammography: a dilemma. **J Natl Cancer Inst Monogr** 1997; 22:151-6.

Feng Y, Sun B, Li X, et al. Differentially expressed genes between primary cancer and paired lymph node metastases predict clinical outcome of node-positive breast cancer patients. **Breast Cancer Res Treat** 2007; 103:319-29.

Fidler IJ, Hart IR. Biological diversity in metastatic neoplasms: origins and implications. **Science** 1982; 217:998-1003.

Fidler IJ. The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited. **Nat Rev Cancer** 2003; 3:453-8.

Foss RD, Guha-Thakurta N, Conran RM, Gutman P. Effects of fixative and fixation time on the extraction and polymerase chain reaction amplification of RNA from paraffin-embedded tissue. Comparison of two housekeeping gene mRNA controls. **Diagn Mol Pathol** 1994; 3:148-55.

Giles RH, van Es JH, Clevers H. Caught up in a Wnt storm: Wnt signaling in cancer. **Biochim Biophys Acta** 2003; 1653:1-24.

Gimenez V. **A gênese do câncer.** Disponível em: <URL:<http://ctc.fmrp.usp.br/.../biocancer/valeria3.html>> [2008 jan 12]

Goldhirsch A, Glick JH, Gelber RD, Coates AS, Senn HJ. Meeting highlights: International Consensus Panel on the Treatment of Primary Breast Cancer. Seventh International Conference on Adjuvant Therapy of Primary Breast Cancer. **J Clin Oncol** 2001; 19:3817-27.

Goldhirsch A, Wood WC, Senn HJ, Glick JH, Gelber RD. Meeting highlights: international consensus panel on the treatment of primary breast cancer. **J Natl Cancer Inst** 1995; 87:1441-5.

Gomes LI, Silva RL, Stolf BS. Comparative analysis of amplified and nonamplified RNA for hybridization in cDNA *microarray*. **Anal Biochem** 2003; 321:244-51.

Handschuh G, Candidus S, Lubert B. Tumour-associated E-cadherin mutations alter cellular morphology, decrease cellular adhesion and increase cellular motility. **Oncogene** 1999; 18:4301-12.

Harris TJ, Peifer M. Decisions, decisions: beta-catenin chooses between adhesion and transcription. **Trends Cell Biol** 2005; 15:234-7.

He TC, Sparks AB, Rago C, et al. Identification of c-MYC as a target of the APC pathway. **Science** 1998; 281:1509-12.

Hemavathy K, Ashraf SI, Ip YT. Snail/slug family of repressors: slowly going into the fast lane of development and cancer. **Gene** 2000; 257:1-12.

Henderson BE, Ross R, Bernstein L. Estrogens as a cause of human cancer: the Richard and Hinda Rosenthal Foundation award lecture. **Cancer Res** 1988; 48:246-53.

Hirohashi S, Kanai Y. Cell adhesion system and human cancer morphogenesis. **Cancer Sci** 2003; 94:575-81.

Ihaka R, Gentleman R. R: a language for data analysis and graphics. **J ComputGraph Statist** 1996; 5:299-314.

Jawhari AU, Farthing MJ, Pignatelli M. The E-cadherin/epidermal growth factor receptor interaction: a hypothesis of reciprocal and reversible control of intercellular adhesion and cell proliferation. **J Pathol** 1999; 187:155-7.

Karacas B, Bachman KE, Park BH. Mutations of the PIK3CA oncogene in human cancer. **Br J Cancer** 2006; 94:455-9.

Kato M. WNT/PCP signaling pathway and human cancer (review). **Oncol Rep** 2005; 14:1583-8.

Kumar S, Walia V, Ray M, Elble RC. p53 in breast cancer: mutation and countermeasures. **Front Biosci** 2007; 12:4168-78.

Lavieille JP, Righini C, Reyt E, Brambilla C, Riva C. Implications of p53 alterations and anti-p53 antibody response head and neck squamous cell carcinomas. **Oral Oncol** 1988; 34:84-92.

Leonard GD, Swain SM. Ductal carcinoma in situ, complexities and challenges. **J Natl Cancer Inst** 2004; 96:906-20.

Li SY, Rong M, Grieco F, Iacopetta B. PIK3CA mutations in breast cancer are associated with poor outcome. **Breast Cancer Res Treat** 2006; 96:91-5.

Maggard MA, Lane KE, O'Connell JB, Nanyakkara DD, Ko CY. Beyond the clinical trials: how often is sentinel lymph node dissection performed for breast cancer? **Ann Surg Oncol** 2005; 12:41-7.

Meireles SI, Cristo EB, Carvalho AF. Molecular classifiers for gastric cancer and nonmalignant diseases of the gastric mucosa. **Cancer Res** 2004; 64:1255-65.

Ministério da Saúde. **Estimativa/2008 Incidência de câncer no Brasil**. Rio de Janeiro: INCA; 2007.

Miyoshi Y, Iwao K, Nawa G, Yoshikawa H, Ochi T, Nakamura Y. Frequent mutations in the beta-catenin gene in desmoid tumors from patients without familial adenomatous polyposis. **Oncol Res** 1998; 10:591-4.

Molina R, Barak V, van Dalen A, et al. Tumor markers in breast cancer-European Group on Tumor Markers recommendations. **Tumour Biol** 2005; 26:281-93.

Moon RT, Bowerman B, Boutros M, Perrimon N. The promise and perils of Wnt signaling through beta-catenin. **Science** 2002; 296:1644-6.

Nakhliis F, Morrow M. Ductal carcinoma in situ. **Surg Clin North Am** 2003; 83:821-39.

Nelson WJ, Nusse R. Convergence of Wnt, β -catenin, and cadherin pathways. **Science** 2004; 303:1483-7.

Nishidate T, Katagiri T, Lin ML, et al. Genome-wide gene-expression profiles of breast-cancer cells purified with laser microbeam microdissection: identification of genes associated with progression and metastasis. **Int J Oncol** 2004; 25:797-819.

Noel A, Foidart JM. The role of stroma in breast carcinoma growth in vivo. **J Mammary Gland Biol Neoplasia** 1998; 3:215-25.

Osborne CK, Yochmowitz MG, Knight WA 3rd, McGuire WL. The value of estrogen and progesterone receptors in the treatment of breast cancer. **Cancer** 1980; 46(12 Suppl):2884-8.

Park D, Karesen R, Axcrone U, Noren T, Sauer T. Expression pattern of adhesion molecules (E-cadherin, alpha-, beta-, gamma-catenin and claudin-7), their influence on survival in primary breast carcinoma, and their corresponding axillary lymph node metastasis. **APMIS** 2007; 115:52-65.

Payne SJ, Bowen RL, Jones JL, Wells CA. Predictive markers in breast cancer--the present. **Histopathology** 2008; 52:82-90.

Peifer M, Polakis P. Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus. **Science** 2000; 287:1606-9.

Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumor. **Nature** 2000; 406:747-52.

Quackenbush J. Computational analysis of microarray data. **Nat Rev Genet** 2001; 2:418-27.

Quackenbush J. Microarray data normalization and transformation. **Nature Genet** 2002; Suppl 32:496-501.

Samuels Y, Velculescu VE. Oncogenic mutations of PIK3CA in human cancers. **Cell Cycle** 2004; 3:1221-4.

Sanguinetti CJ, Dias Neto E, Simpson AJ. Rapid silver staining and recovery of PCR products separated on polyacrilamide gels. **Biotechniques** 1994; 17:914-21.

Saraiva TF, Castro NP, Pineda PHB, et al. Effects of Oligo dT-T7 RNA Primer in RNA Amplification from Paraffin-Embedded Tissue for *Microarray* Experiments. **Appl Cancer Res** 2006; 26:14-20.

Sgroi DC, Teng S, Robinson G, LeVangie R, Hudson JR Jr, Elkahloun AG. In vivo gene expression profile analysis of human breast cancer progression. **Cancer Res** 1999; 59:5656-61.

Shook D, Keller R. Mechanisms, mechanics and function of epithelial-mesenchymal transitions in early development. **Mech Dev** 2003; 120:1351-83.

Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proc Natl Acad Sci U S A** 2001; 98:10869-74.

Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. **Proc Natl Acad Sci U S A** 2003; 100:10393-8.

Sparks AB, Morin PJ, Vogelstein B, Kinzler KW. Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. **Cancer Res** 1998; 58:1130-4.

T & D. Medicina. **Câncer de mama**. Disponível em: <URL:<http://members.tripod.com/Reivax/cancerpg12.htm>> [2008 abr. 18]

Tetsu O, McCormick F. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. **Nature** 1999; 398:422-6.

Thiery JP. Epithelial-mesenchymal transitions in development and pathologies. **Curr Opin Cell Biol** 2003; 15:740-6.

Van 't Veer LJ, Dai H, van de Vijver MJ, et al. Expression profiling predicts outcome in breast cancer. **Breast Cancer Res** 2003; 5:57-8.

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.

Veronesi U, Paganelli G, Viale G, et al. A randomized comparison of sentinel-node biopsy with routine axillary dissection in breast cancer. **N Engl J Med** 2003; 349:546-53.

Vincent-Salomon A, Thiery JP. Host microenvironment in breast cancer development: epithelial-mesenchymal transition in breast cancer development. **Breast Cancer Res** 2003; 5:101-6.

Vivanco I, Sawyers CL. The phosphatidylinositol 3-kinase AKT pathway in human cancer. **Nat Rev Cancer** 2002; 2:489-501.

Waltzer L, Bienz M. The control of beta-catenin and TCF during embryonic development and cancer. **Cancer Metastasis Rev** 1999; 18:231-46.

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

Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. **Lancet** 2005; 365:671-9.

Wong YK, Liu TY, Chang KW, Lin SC, Chao TW, Li PL, Chang CS. p53 alterations in betel quid- and tobacco-associated oral squamous cell carcinomas from Taiwan. **J Oral Pathol Med** 1998; 27:243-8.

XLMINER. Data mining in Excel: XLMiner®. Available from: <URL:<http://www.resample.com/xlminer>> [2008 abr 18].

Yan JJ, Tzeng CC, Jin YT. Overexpression of p53 protein in squamous cell carcinoma of buccal mucosa and tongue in Taiwan: an immunohistochemical and clinicopathological study. **J Oral Pathol Med** 1996; 25:55-9.

Yang YH, Speed T. Design issues for cDNA microarray experiments. **Nature** 2002; 3:579-88.

Yook JI, Li XY, Ota I, Fearon ER, Weiss SJ. Wnt-dependent regulation of the E-cadherin repressor snail. **J Biol Chem** 2005; 280:11740-8.

Zujewski J, Liu ET. The 1998 St. Gallen's Consensus Conference: an assessment. **J Natl Cancer Inst** 1998; 90:1587-9.

ANEXOS

Original Article

Effects of Oligo dT-T7 RNA Primer in RNA Amplification from Paraffin-Embedded Tissue for Microarray Experiments

Thiago Felcar Saraiva,¹ Nádia Pereira de Castro,¹ Paulo Henrique Baldan Pineda,¹ Cynthia A. Bueno de Toledo Osório, M.D.,² Luiz Paulo Camargo,³ Helena Paula Brentani, PhD,³ Dirce Maria Carraro, PhD¹

¹ Molecular Biology Lab – Hospital A. C. Camargo, São Paulo, Brazil

² Department of Pathological Anatomy – Hospital A. C. Camargo, São Paulo, Brazil

³ Department of Bioinformatics– Hospital A. C. Camargo, São Paulo, Brazil

Abstract

Introduction: Formalin-Fixed Paraffin-Embedded Tissue samples (FFPET) represent a valuable source for studies of gene expression comparisons, since a great number of these samples is available in archive and presents a long time of clinical follow-up. However, the quality of total RNA of these samples is known to be inferior to frozen samples, being many times inadequate for studies of gene expression using conventional methodologies. **Objective:** This study aims to establish a protocol for amplification of messenger RNA (mRNA) derived from FFPET samples for using in microarray experiments. **Material and Methods:** 4 tumoral samples of invasive ductal breast carcinoma FFPET-buffered 10% were used. Total RNA was extracted and the mRNA was linearly amplified in two rounds based on T7 RNA polymerase methodology using different concentrations of oligo dT-T7 Primer for first strand cDNA (1st-cDNA) synthesis. Amplified antisense RNA (aRNA) was labeled with cyanine-Cy3 through reverse transcription in the presence of random primers and co-hybridized with reference RNA (HB4a) labeled with cyanine-Cy5 in a customized platform containing 4,608 cDNAs corresponding to human genes. **Results:** The amplified RNA quality was influenced by the relative amount of oligo dT-T7, showing better results for ratio of 1:0.1 (total RNA : oligo dT-T7). Hybridizations showed value of intensity signals for the most of cDNAs immobilized in the platform. **Conclusion:** This study showed that the control of the relative amounts of RNA derived from FFPET material and oligo dT-T7 is extremely important to obtain high-quality amplified RNA, allowing its use in microarray experiments.

Keywords: Microarray analysis. Nucleic Acid Amplification Techniques. Tissue Fixation.

Introduction

With the development of new technologies for comparative analysis of gene expression, such as microarray, a great number of tumoral samples are required. This methodology allows simultaneously analysis of the expression profile of thousands of genes and it has been extensively used in the characterization of the biological events that lead to tumors development.¹ Moreover, this technology has also been used to searching for diagnostic, prognostic,²⁻⁵ and predictive markers,⁶⁻⁹ and also for therapeutic target. Therefore, in

clinical-oriented studies, the demand for appropriate tumoral samples has increased in the last years. An important requirement for this type of gene expression analysis is the quality of total RNA. Frozen tumor samples are normally used, since they present a better RNA preservation. Several groups have made efforts for adequately

Correspondence

Dirce Maria Carraro

Rua Professor Antônio Prudente, 211, 1st floor
01509-900 São Paulo, Brazil

Phone 55 11 33883262

E-mail dcarraro@ludwig.org.br

Original Article

Effects of Oligo dT-T7 RNA Primer in RNA Amplification from Paraffin-Embedded Tissue for Microarray Experiments

Thiago Felcar Saraiva,¹ Nádia Pereira de Castro,¹ Paulo Henrique Baldan Pineda,¹ Cynthia A. Bueno de Toledo Osório, M.D.,² Luiz Paulo Camargo,³ Helena Paula Brentani, PhD,³ Dirce Maria Carraro, PhD¹

¹ Molecular Biology Lab – Hospital A. C. Camargo, São Paulo, Brazil

² Department of Pathological Anatomy – Hospital A. C. Camargo, São Paulo, Brazil

³ Department of Bioinformatics– Hospital A. C. Camargo, São Paulo, Brazil

Abstract

Introduction: Formalin-Fixed Paraffin-Embedded Tissue samples (FFPET) represent a valuable source for studies of gene expression comparisons, since a great number of these samples is available in archive and presents a long time of clinical follow-up. However, the quality of total RNA of these samples is known to be inferior to frozen samples, being many times inadequate for studies of gene expression using conventional methodologies. **Objective:** This study aims to establish a protocol for amplification of messenger RNA (mRNA) derived from FFPET samples for using in microarray experiments. **Material and Methods:** 4 tumoral samples of invasive ductal breast carcinoma FFPET-buffered 10% were used. Total RNA was extracted and the mRNA was linearly amplified in two rounds based on T7 RNA polymerase methodology using different concentrations of oligo dT-T7 Primer for first strand cDNA (1st-cDNA) synthesis. Amplified antisense RNA (aRNA) was labeled with cyanine-Cy3 through reverse transcription in the presence of random primers and co-hybridized with reference RNA (HB4a) labeled with cyanine-Cy5 in a customized platform containing 4,608 cDNAs corresponding to human genes. **Results:** The amplified RNA quality was influenced by the relative amount of oligo dT-T7, showing better results for ratio of 1:0.1 (total RNA : oligo dT-T7). Hybridizations showed value of intensity signals for the most of cDNAs immobilized in the platform. **Conclusion:** This study showed that the control of the relative amounts of RNA derived from FFPET material and oligo dT-T7 is extremely important to obtain high-quality amplified RNA, allowing its use in microarray experiments.

Keywords: Microarray analysis. Nucleic Acid Amplification Techniques. Tissue Fixation.

Introduction

With the development of new technologies for comparative analysis of gene expression, such as microarray, a great number of tumoral samples are required. This methodology allows simultaneously analysis of the expression profile of thousands of genes and it has been extensively used in the characterization of the biological events that lead to tumors development.¹ Moreover, this technology has also been used to searching for diagnostic, prognostic,²⁻⁵ and predictive markers,⁶⁻⁹ and also for therapeutic target. Therefore, in

clinical-oriented studies, the demand for appropriate tumoral samples has increased in the last years. An important requirement for this type of gene expression analysis is the quality of total RNA. Frozen tumor samples are normally used, since they present a better RNA preservation. Several groups have made efforts for adequately

Correspondence

Dirce Maria Carraro

Rua Professor Antônio Prudente, 211, 1st floor
01509-900 São Paulo, Brazil

Phone 55 11 33883262

E-mail dcarraro@ludwig.org.br

storing surgical material in biorepositories.¹⁰ The advances of diagnostic methods have allowed early tumor detection significantly increasing the possibilities for cure. In these cases, collected material is totally embedded in paraffin and used for anatomopathological examinations. Fixation and preservation of this material aim especially to maintain its morphology. Currently, great attention has been given to the RNA preservation using more adequate fixative and also fixation time¹¹⁻¹³ with the aim of gene expression studies. Real time RT-PCR methodology^{14,15} has been preferentially applied for gene expression studies using FFPET samples stored in archives, because this methodology requires small amounts of total RNA (not necessarily of excellent quality) that can be converted into cDNA using random primers as triggers. These primers anneal to any mRNA region, and do not require intact poly-A tail. However, mRNA amplification is required for analyzing gene expression by microarray technology due to insufficient amount of mRNA normally recovered from frozen or FFPET. Several linear amplification protocols are available in the literature showing acceptable reproducibility and high data fidelity when compared to total RNA^{16,17} from the same source. These protocols are based on *in vitro* transcription using oligonucleotides that carry RNA polymerase T7 phage promoter. As this primer anneals in mRNA's poly-A tail for synthesizing 1st-cDNA, the presence of a significant amount of mRNA molecules with intact polyA tail is necessary. In the present study, in order to establish an amplification protocol for total RNA obtained from FFPET samples, we compared the quality of amplified RNA after two rounds of amplification using different oligo dT-T7 concentrations for 1st-cDNA synthesis from total RNA. The quality of the two-round amplified RNA (aRNA-2nd) was evaluated by agarose gel. Our results showed that relative amount of total RNA and oligo dT-T7 primer influences the quality of aRNA-2nd. Thus, through the control of this ratio we were able to obtain amplified RNA of satisfactory quality from FFPET samples appropriated for microarray experiments.

Materials and Methods

Samples description and extraction of total RNA

Four FFPET (buffered 10%) samples of breast invasive ductal carcinoma from years 2000, 2001, 2003 and 2004, were obtained from the archive of Hospital A. C. Camargo. Four 5µm sections were cut from each paraffin block. The cuts were deparaffinized in 1mL of xylol and centrifuged at 14000rpm for 10min. After removing xylol, the pellet obtained was washed in 70% ethanol. The proteins and membranes digestion was carried out with proteinase K (Invitrogen) for 18 hours at 52°C.¹⁸ For extraction of total RNA, RNEasy Mini Kit (Qiagen) was used according to the manufacturer recommendations. Total RNA extraction from Cell line HB4a was performed with TRIzol® reagent (Invitrogen Corp., Carlsbad, CA), according to the manufacturer protocol. The RNA integrity was evaluated by 1% agarose gel and ethidium bromide staining.

Double-strand DNA synthesis

For 1ststrand cDNA synthesis, 1mg total RNA was used with different amounts of oligodT-T7 (5'5' GGC CAG TGA ATT GTA ATA CGA CTC ACT ATA GGG AGG CGG TTT TTT TTT TTT TTT TTT TTT TTT (24)3 ') [57-pb]

Total RNA and oligo dT-T7 were denatured at 70°C for 15 minutes and snap cooled on ice. Four microliters of first strand buffer 5X, 0.01M DTT (Invitrogen), 40U rRNasin (Promega, Madison,WI), 1mM dNTP (Amersham Biosciences, NJ), 400U Superscript II (Invitrogen) and 4µl of DEPC (diethylpyrocarbonate) treated water were added to the reaction in a 20µl final volume. The reaction was incubated at 42°C for 2 hours. For 2nd strand cDNA synthesis (double strand cDNA), 53µl of DEPC-treated water, 20ml of 2nd strand buffer (Invitrogen), 1mM dNTP (Amersham Biosciences, NJ), 1.4U RNase H (Invitrogen), 10U *E. Coli* DNA ligase, 40U *E. Coli* DNA polymerase I were added to the reaction in 100µl final volume. The reaction was incubated at 16°C for two hours. Next, 10U *T4* DNA polymerase I was added and incubated at 16°C for

5 minutes. The reaction was stopped by adding 0.05M EDTA. Double strand DNA (ds-cDNA) purification was carried out by phenol:chloroform:isoamyl alcohol (25:24:1) (pH 8.0) and precipitated in 300µl absolute ethanol, 40µl sodium acetate (3M) and 1µl of acrylamide (5mg/ml). The pellet was washed 5 times with 1ml of ethanol 70%, and resuspended in 11µl of DEPC treated water. For indirect evaluation of RNA integrity recovered from FFPET samples RT-PCR reaction of the GAPDH gene (NM_002046.2) using Forward-CTGCACCACCAAGTGCTTA and Reverse-CATGACGGCAGGTCAGGTC primers was performed. This primer pair anneals in exon 7 and exon 8 at 463bp of 3' end amplifying a 292 bp fragment. One microliter ds-DNA of each sample was diluted in 9µl water. Two µl of this dilution was used as template for the RT-PCR reaction adding 1.25mM MgCl₂, 0.2mM dNTPs (Amersham Biosciences, NJ), 0.5mM of each primer, 2 µl reaction buffer 10X, 1U *Taq DNA Polymerase* (Invitrogen), in a final volume of 20µl. The program consisted of 2 minutes initial denaturation at 94°C, 40 cycles of 45 seconds at 94°C, 45 seconds at 61°C and 1 minute at 72°C, followed by an additional 7 minutes at 72°C. The product was visualized by 8% polyacrylamide gel electrophoresis and silver nitrate stained.¹⁹

***In vitro* Transcription**

In vitro transcription was performed with ds-DNA used as template and 7.5mM rNTP, 5.0µl reaction buffer 5X, 2.5µl enzyme solution (RNase inhibitor and T7 RNA polymerase - Promega) in a 10µl final volume. The reaction was incubated at 37°C for 6h. The purification of aRNA was carried out with TRIzol® (Invitrogen Corp., Carlsbad, CA) following the fabricant recommendations and the precipitation was performed with absolute ethanol (3X volume), 60µl sodium acetate (3M) and 0.5µl glycogen (20mg/ml). The pellet was washed 5 times with 1 ml of 70% ethanol prepared in DEPC-treated water. For the second round of amplification, aRNA-1st (aRNA first round) was resuspended in 6µl DEPC-treated water and incubated at 70°C with 9µg of dN6 random primer for 5min. After that, 4µl first strand Buffer 5X, 10mM DTT (Invitrogen), 40U rRNasin (Promega, Madison,

WI), 1mM dNTP (Amersham Biosciences, NJ) and 400U *SuperScriptII* (Invitrogen) were added in 20µl final volume. The reaction was incubated for 2h at 42°C. The 2nd strand cDNA was produced by adding 0.4µg OligodT(24)-T7 primer, 61.3µl of DEPC-treated water, 10µl *Advantage PCR Buffer* (Clontech), 1mM dNTP (Amersham Biosciences, NJ), 1.4U RNase H (Invitrogen) and 5X *Advantage cDNA Polymerase mix* (Clontech, CA) in a final volume of 100µl, following the program 37°C for 10min, 94°C for 3min, 65°C for 5min, 75°C for 30min. For enzymatic inactivation 7.5µl 1M NaOH and 2mM EDTA was used and the reaction was incubated at 65°C for 10 minutes. The ds-cDNA was recovered and the second round of *in vitro* transcription was carried out following the exactly steps of the first round amplification. The RNA amplification from normal breast cell line HB4a 20 was carried out following the same procedure used for FFPET samples.

Evaluation of aRNA-2nd quality

The evaluation of aRNA-2nd quality was carried out by electrophoresis for 50 minutes at 60mV. Five microliters TAE buffer with 7M urea, 3X bromophenol blue *Loading Buffer* and 1µg ethidium bromide were added to 1µg aRNA-2nd and the mixture was applied into the agarose gel.

Labeling and hybridization

Five micrograms of aRNA-2nd from FFPET samples and 4µg of aRNA-2nd from HB4a were labeled with cyanine. To each aRNA-2nd, 1.2µl 4.5X mRNA control (Q-gene) and 9µg random hexamer primer (dN6) were added, and primer annealing was performed at 70°C for 5 minutes. To this mixture 1X first strand buffer, 0.01M DTT (Invitrogen), 20U rRNasin (Promega, Madison, WI), 0.6µL dNTP in equal amounts of adenine, thiamine and guanine (25mM each) and cytosine (10mM), 2µL of dCTP, Cy3 for test sample and Cy5 for reference sample (25mM) and 400U *SuperScript Reverse TranscriptaseII* (Invitrogen) were added in 30µl final volume. The reaction was incubated at 42°C for 42 hours. For RNA degradation 0.025M EDTA and 0.05M NaOH was added and incubated at 70°C for 20minutes, followed by neutralization with 0.05M HCl. For eliminating the non-incorporated dNTPs, the solution was purified by AutoSeq G50 (Armesthan Pharmacia) purification

column according to the manufacturer recommendations. Sample and its reference were mixed and concentrated in speed vac. Four micrograms of poly-A DNA (*Armeshan Pharmacia*), 4µg Cot 1 (*Gibco-BRL*), 20µg BSA (BSA fraction 5 - *Sigma*), 47,5µl hybridization buffer (10X SSC, 0.2% SDS), preheated at 70°C, formamide 25%, 5X Denhardt's solution and 10µg salmon sperm DNA were added to the mixture in 95µl final volume. The final solution was heated at 95°C for 5 minutes for denaturation and applied on cDNA platform 21 containing 4.608 ORESTES clones corresponding to human genes. The platform was previously hybridized at 42°C (6 to 16 hours) in 5X SSC SDS 0.2%, BSA 1%, 5X Denhardt's solution (filtered in pore membrane 0.45µm). The slides were mounted in GeneTac 2000 (GenomicSolutions) hybridization station at 42°C for 20hours. After hybridization, the slides were washed in the following conditions: one time in 2X SSC solution, twice in 0.1% SSC solution and 1X SDS and twice in 0.1% SSC solution. All steps were carried out under agitation for 10 minutes and all solutions were previously heated at 42°C.

Signal capture

Signals generated by competitive hybridization of samples were captured with ScanArrayTM Express (*Packard BioScience*) dualchannel confocal laser scanner using PMT 50% for all slides (*Photomultiplier voltage*). Data were recovered by ScanArray (*Perkin Elmer Life Sciences, Inc., U.S.A.*) using the histogram method.

Mathematical analyses

Saturated points (higher or equal to 63000 approximately 16 BITS) were eliminated and local background subtraction was performed. Only spots with positives values were considered for analysis.

Results and Discussion

Evaluation of the aRNA-2nd quality

It has been reported in the literature that a high initial concentration of oligo dT-T7 used for amplification can influence the quality of amplified RNA,²² by generation of primers dimers that ser-

ves as template in the *in vitro* transcription reaction. These primer dimers yield a great amount of amplified RNA with low molecular weight.²² In this study we first tested the Jensen et al. findings in our conditions. In an attempt to simulate FFPET samples, which normally present low quantity of intact mRNA, we used small amount of a high quality total RNA (50ng) recovered from the HB4a cell line. Three amounts (500, 50 and 5 ng) of oligo dT-T7 were tested in the 1st-cDNA synthesis, resulting in 3 different relative amounts (RNA:primer) 1:10, 1:1, and 1:0.1 listed in Table 1. The same steps of the amplification procedure were performed in the 3 experiments. In experiment 1, with a 1:10 ratio (RNA:primer), aRNA-2nd was predominantly of 200 to 500bp obtaining a total mass of 96.72µg. In experiment 2, with a 1:1 ratio (RNA:primer), aRNA-2nd predominantly of 300 to 800bp obtaining a total mass of 5.84µg; and finally, in experiment 3, with a 1:0.1 (RNA:primer) ratio aRNA-2nd was predominantly of 300 to 800bp, equivalent to the 1:1 ratio, but with total mass of 0.8µg (Figure 1, Table 1). Data provided by our laboratory have showed that aRNA-2nd molecules with size ranging from 300 to 800bp produce excellent quality hybridizations. In contrast, aRNA-2nd molecules with size ranging from 100 to 300bp produce poor quality hybridizations, showing intensity signal of hybridization in a few cDNAs immobilized in our customized platform. Taking this data into consideration we can suggest that the ratios 1:1 and 1:0.1 produced aRNA-2nd adequate for microarray experiments. However, this was not observed in the 1:10 ratio, which produced aRNA-2nd molecules of lower molecular weight. Our findings confirmed Jensen et al.²² data and showed that higher aRNA size range was obtained when lower primer concentrations related to total RNA were used. Although in the

Table 1 - Different ratio of total RNA and oligo dT-T7, final mass and aRNA-2nd

Experiments	RNA:primer ratio	aRNA mass	aRNA-2 nd (pb)
1	1:10	96.72µg	200-500
2	1:1	5.84µg	300-800
3	1:0.1	0.8µg	300-800

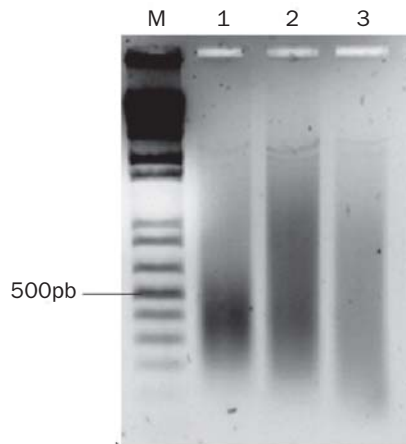


Figure 1 - Agarose Gel- 500ng aRNA-2nd from HB4a using different amounts of oligo dT-T7 in 1st-cDNA synthesis (1, 2 and 3- 500ng, 50ng and 5ng, respectively). M- ladder; 1, 2 and 3- aRNA2nd

experiment 3 (1:0.1) an excellent quality aRNA-2nd was obtained the total mass (0.8µg) produced was not enough to perform microarray experiments, which normally use 4µg. However, in experiment 2 (1:1) the amount of aRNA-2nd produced (5.84µg) was enough for hybridizing one microarray platform.

Evaluation of aRNA-2nd quality from FFPET samples

To test if equivalent results would be observed using RNA of inferior quality, the same experiment using the same relative concentrations of experiments 2 and 3 (1:1 and 1:0.1, respectively) was performed with one FFPET sample (year 2004). One µg of total RNA was evaluated by agarose gel and presented degradation (data not shown). Therefore, its quality was indirectly evaluated by RT-PCR of *GAPDH* gene using ds-DNA as template, cited in material and methods. One hundred nanograms of total RNA were used as starting material for the amplification experiments. In this procedure two quantities of oligo dT-T7 were used, 100 and 10ng, reproducing the 1:1 and 1:0.1 ratios, respectively. In experiment 4 (1:1 ratio), 7µg of aRNA-2nd was obtained, ranging from 200 to 400bp. In the experiment 5 (1:0.1 ratio), 0.5µg of aRNA-2nd was obtained in the range of 300 to 700bp (data not shown), which produces high-quality hybridizations, based on our previous data. These results showed that the

optimized RNA:primer ratio for FFPET sample (low quality total RNA) differed from the optimized ratio for HB4a (high quality total RNA). The alteration of the RNA:primer ratio can be explained by the lower number of intact mRNA molecules in FFPET sample than in HB4a cell line, even when using double quantity of total RNA as starting material. In experiment 5 (1:0.1 ratio) the aRNA-2nd ranged from 300 to 700bp corresponding to high quality hybridization. However, the total amount obtained was insufficient for microarray experiments. Similar to the previous results, these data showed that the amount of oligo dT-T7 affected the quality of amplified RNA (Table 2). To obtain the necessary amount for performing microarray experiments (around 5µg), the next step was to increase the amount of starting material (total RNA) recovered from FFPET samples and to use a 1:0.1 RNA:primer ratio, that showed to be more appropriated for this type of sample. For this, total RNAs were extracted from 4 FFPET samples (years 2000, 2002, 2003 and 2004) and the amplification procedure were initiated with 1µg of total RNA and 100ng of oligo dT-T7 for 1st-cDNA synthesis (RNA:primer ratio: 1:0.1). The RNA integrity of these samples was indirectly evaluated by RT-PCR for *GAPDH* gene (Figure 2).

The samples were amplified, producing aRNA-2nd with a satisfactory quantity (more than 5µg), as shown in Table 3. aRNA-2nd quality was evaluated by agarose gel for 3 out of 4 samples (which presented sufficient amount) (Figure 3). Although the sample 1 was not checked by agarose gel, due to the low quantity obtained (5.2µg), all aRNA-2nd from the FFPET samples were used in microarray experiments.

Microarray experiments

For hybridization, 5µg aRNA-2nd from FFPET samples and 4µg aRNA-2nd from HB4a

Table 2 - Different ratio of total RNA and oligo dT-T7, final mass and aRNA-2nd

Experiments	RNA:primer ratio	aRNA mass	aRNA-2 nd (pb)
4	1:1	7µg	200-400
5	1:0.1	0.5µg	300-700

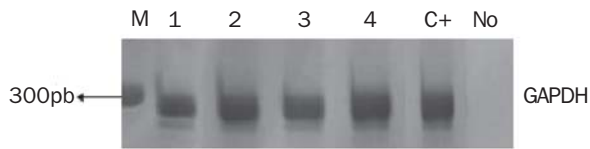


Figure 2 - Polyacrylamide Gel. GAPDH gene RT_PCR. M- ladder; 1, 2, 3 and 4- FFPET samples; C+ - positive control and No- negative control

Table 3 - aRNA-2nd data of FFPET samples using 1:0.1 RNA:primer ratio

Samples	RNA: primer ratio	aRNA mass	aRNA-2 nd (pb)
2000	1:0.1	5.2µg	-
2002	1:0.1	14.5µg	200-800
2003	1:0.1	13.4µg	200-500
2004	1:0.1	14.8µg	200-800

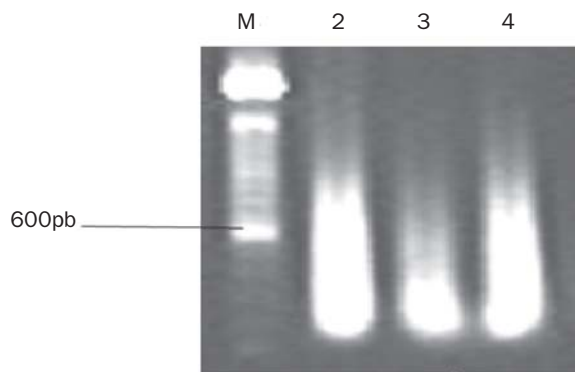


Figure 3 - Agarose Gel- 1µg aRNA-2nd from FFPET samples. M- ladder; 1- FFPET from 2002; 2- FFPET from 2003 and 3- FFPET from 2004.

(used as reference) were labeled with Cy3 and Cy5, respectively, and co-hybridized in the customized 4.8K platform. The hybridizations showed that most cDNAs produced positive intensity signal (Table 4) corresponding to signals higher than the background. Figure 4 presents the images produced by the *ScanArray* program (*Perkin Elmer Life Sciences, Inc., U.S.A.*) showing high quality of hybridization for all slides. The higher number of cDNAs with signal intensity lower than the *background* signals was observed in the slides correspondent to the reference samples. This fact is probably due to two reasons; the two dyes incorporate into samples at different rates, creating “dye bias” 23-25 and a lower

Table 4 - Number of spots with intensity signal lower than background

Year of FFPET fixation	Number of spots with intensity signal lower than background signal	
	FFPET samples	Reference samples
2000	12	73
2002	4	63
2003	5	84
2004	0	10

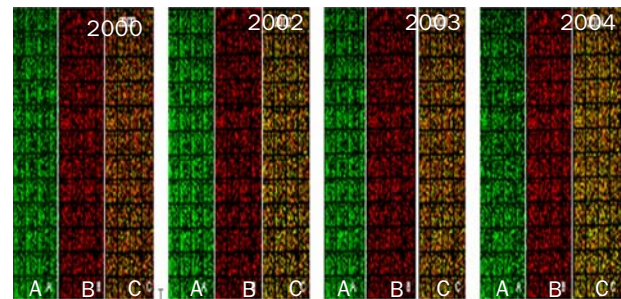


Figure 4 - Representation of the results of hybridization 4.8K microarray platform. A) cDNA hybridization of FFPET from 2000, 2002, 2003 and 2004; B) cDNA hybridization of reference sample and C) Composite (green: FFPET and red: reference).

amount of aRNA-2nd (4µg) used in hybridization. These preliminary results produced in this article showed that starting amount of oligo dTT7 used in the amplification procedure of paraffin embedded samples caused a great influence on the quality and amount of amplified RNA. Moreover, due to the relatively good quality in hybridization experiments provided in this study our data suggested that FFPET samples of the Archive A. C. Camargo Cancer Hospital can be used for gene expression studies. However, further experiments are necessary for evaluating the reproducibility and reliability of the data. One must emphasize the necessity of improving the procedures used to fixation of paraffin-embedded samples for a better RNA preservation, thus allowing its use for gene expression studies.

Financial Support

Conselho Nacional de Desenvolvimento

Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa de São Paulo (FAPESP/CEPID)

Acknowledgements

We are grateful to all participants from Ludwig Institute and Hospital do Câncer A.C. Camargo, specially Carlos Ferreira do Nascimento, and also for the financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa de São Paulo (FAPESP/CEPID).

References

- Provenzani A, Fronza R, Loreni F, Pascale A, Amadio M, Quattrone A. Global alterations in mRNA polysomal recruitment in a cell model of colorectal cancer progression to metastasis. *Carcinogenesis* 2006;27:1323-33.
- Alizadeh AA, Ross DT, Perou CM, van de Rijn M. Towards a novel classification of human malignancies based on gene expression patterns. *J Pathol* 2001;195: 41-52.
- Van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415:530-6.
- Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100:10393-8.
- Seth A, Kitching R, Landberg G, Xu J, Zubovits J, Burger AM. Gene expression profiling of ductal carcinomas in situ and invasive breast tumors. *Anticancer Res* 2003;23:2043-51.
- Wolff AC, Davidson NE. Primary systemic therapy in operable breast cancer. *J Clin Oncol* 2000;18:1558-69.
- Chang JC, Wooten EC, Tsimelzon A, Hilsenbeck SG, Gutierrez MC, Elledge R et al. Gene expression profiling for the prediction of therapeutic response to docetaxel in patients with breast cancer. *Lancet* 2003;362:362-9.
- Shimizu D, Ishikawa T, Ichikawa Y, Togo S, Hayasizaki Y, Okazaki Y et al. Current progress in the prediction of chemosensitivity for breast cancer. *Breast Cancer* 2004;11:42-8.
- Folgueira MA, Carraro DM, Brentani H, Patrao DF, Barbosa EM, Netto MM et al. Gene expression profile associated with response to doxorubicin-based therapy in breast cancer. *Clin Cancer Res* 2005;11:7434-43.
- Auray-Blais C, Patenaude J. A biobank management model applicable to biomedical research. *BMC Med Ethics* 2006;7:E4.
- Foss RD, Guha-Thakurta N, Conran RM, Gutman P. Effects of fixative and fixation time on the extraction and polymerase chain reaction amplification of RNA from paraffin-embedded tissue. Comparison of two housekeeping gene mRNA controls. *Diagn Mol Pathol* 1994;3:148-55.
- Su JM, Perlaky L, Li XN, Leung HC, Antalfi B, Armstrong D et al. Comparison of ethanol versus formalin fixation on preservation of histology and RNA in laser capture microdissected brain tissues. *Brain Pathol* 2004;14:175-82.
- Benčekroun M, DeGraw J, Gao J, Sun L, von Boguslawsky K, Leminen A et al. Impact of fixative on recovery of mRNA from paraffin-embedded tissue. *Diagn Mol Pathol* 2004;13:116-25.
- Godfrey TE, Kim SH, Chavira M, Ruff DW, Warren RS, Gray JW et al. Quantitative mRNA expression analysis from formalin-fixed, paraffin-embedded tissues using 5' nuclease quantitative reverse transcription-polymerase chain reaction. *J Mol Diagn* 2000;2:84-91.
- Specht K, Richter T, Muller U, Walch A, Werner M, Hofler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol* 2001;158:419-29.
- Wang E, Miller LD, Ohnmacht GA, Liu ET, Marincola FM. High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol* 2000;18:457-9.
- Gomes LI, Silva RL, Stolf BS, Cristo EB, Hirata R, Soares FA et al. Comparative analysis of amplified and nonamplified RNA for hybridization in cDNA *microarray*. *Anal Biochem* 2003;321:244-51.
- O'Driscoll L, Kennedy S, McDermott E, Kelehan P, Clynes M. Multiple drug resistance-related messenger RNA expression in archival formalin-fixed paraffin-embedded human breast tumour tissue. *Eur J Cancer* 1996;32A:128-33.
- Sanguinetti CJ, Dias N, Simpson AJG. Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* 1994;17:3-6.
- Harris RA, Eichholtz TJ, Hiles ID, Page MJ, O'Hare MJ. New model of ErbB-2 over-expression in human mammary luminal epithelial cells. *Int J Cancer* 1999;80:477-84.
- Brentani RR, Carraro DM, Verjovski-Almeida S, Reis EM, Neves EJ, de Souza SJ et al. Gene expression arrays in cancer research: methods and applications. *Crit Rev Oncol Hematol* 2005;54:95-105.
- Jenson SD, Robetorye RS, Bohling SD, Schumacher JA, Morgan JW, Lim MS et al. Validation of cDNA microarray gene expression data obtained from linearly amplified RNA. *Mol Pathol* 2003;56:307-12.
- Tseng, G., Oh, M., Rohlin, L., Liao, J. & Wong, W. (2001). Issues in cDNA microarray analysis: Quality filtering, channel normalization, models of variations and assessment of gene effects, *Nucleic Acids Research* 29: 2540-2557.
- Yang, Y. H., Dudoit, S., Luu, P. & Speed, T. P. (2001). Microarrays: Optical Technologies and Informatics, SPIE, San Jose, CA, chapter Normalization of cDNA microarrays.
- Yang, Y., Dudoit, S., Luu, P., Lin, D., Peng, V., Ngai, J. & Speed, T. (2002). Normalization for cDNA microarray data: A robust composite method addressing single and multiple slide systematic variation, *Nucleic Acids Research* 30: e15.

Anexo 2 - Tabela com os Genes das vias PI3K, WNT e EMT, na lâmina 2.3K.

PI3K			EMT		
ALS2CR7	GRK5	PIP5K2A	BCDO1	IGF1R	TWIST1
BCR	GRK6	PIP5K2B	BHLHB2	IGF2	VEGF
CALM1	GRK7	PIP5K2C	BMP7	ITGA2	WASPIP
CALM2	HIPK1	PIP5K3	C1QDC1	ITGA5	
CALM3	HIPK2	PLCB1	CALD1	KRT19	
CALML3	IMPA1	PLCB2	CDC42	LAMA3	
CARKL	IMPA2	PLCB3	CDH1	MEOX1	
CDC2L1	INPP1	PLCB4	CDH13	MET	
CDC2L5	INPP4A	PLCD1	CDH6	MSN	
CDC7	INPP4B	PLCD3	CITED4	MYCN	
CDIPT	INPP5A	PLCD4	COL4A1	NCAM1	
CDS1	INPP5B	PLCE1	COL4A2	NCAM2	
CDS2	INPP5D	PLCG1	COL5A1	NGFR	
CSNK1G1	INPPL1	PLCG2	COL5A2	NID2	
CSNK1G2	ITPKA	PLCZ1	CRABP2	NNAT	
CSNK1G3	ITPKB	PRKAA1	CTNNA1	NOX1	
DCAMKL1	ITPKC	PRKAA2	CTNNA2	OFD1	
DGKA	ITPR1	PRKCA	CTNNB1	PAX1	
DGKB	ITPR2	PRKCB1	CTNND1	PAX2	
DGKD	ITPR3	PRKCG	DDR1	PAX3	
DGKE	OCRL	PTEN	DDR2	PAX4	
DGKG	PCTK1	RIPK4	DLL1	PAX5	
DGKH	PCTK2	SNF1LK	DSG2	PAX6	
DGKI	PCTK3	SYNJ1	EDN1	PAX7	
DGKQ	PIK3C2A	SYNJ2	ETV4	PAX8	
DGKZ	PIK3C2B	TESK1	EYA1	PAX9	
DMPK	PIK3C2G	TESK2	EYA2	PCDH11X	
DYRK1A	PIK3C3	TTK	EYA3	PCDH11Y	
DYRK1B	PIK3CA		EYA4	PRAME	
DYRK2	PIK3CB		FGFR3	RARG	
DYRK3	PIK3CD		FRZB	SDC1	
DYRK4	PIK3CG		GATA3	SIP1	
EEF2K	PIK3R1		GH1	SIX1	
EIF2AK2	PIK3R2		GH2	SPARC	
EIF2AK3	PIK3R3		HDGF	TCF3	
EIF-4E	PIK3R4		HOXA11	TCF8	
FASTK	PIK3R5		ID1	TGFA	
FN3K	PIK4CA		ID2	TGFBR2	
GAK	PIP5K1A		ID3	TIMP1	
GRK1	PIP5K1B		ID4	TIMP3	
GRK4	PIP5K1C		IGF1	TITF1	

WNT				
AKT1	DVL1	MAPK10	PPP3R1	TBL1Y
APC	DVL2	MAPK8	PPP3R2	TCF7
APC2	DVL3	MAPK9	PRICKLE1	TCF7L1
AXIN1	EP300	MMP7	PRICKLE2	TCF7L2
AXIN2	EVI1	MTA1	PRKACA	TCF8
BTRC	FOSL1	MYB	PRKACB	TLE1
CACYBP	FOSL2	MYC	PRKACG	TLE2
CAMK2A	FRAT1	NFAM1	PRKCA	TLE3
CAMK2B	FRAT2	NFAT5	PRKCB1	TLE4
CAMK2D	FZD1	NFATC2	PRKCG	TP53
CAMK2G	FZD10	NFATC3	PRKX	TWIN
CAV1	FZD2	NFATC4	PRKY	VANGL1
CCND1	FZD3	NKD1	PSEN1	VANGL2
CCND2	FZD4	NKD2	RAC1	WIF1
CCND3	FZD5	NLE1	RAC1P4	WNT1
CD44	FZD6	NLK	RAC2	WNT10A
CER1	FZD7	NODAL	RAC3	WNT10B
CHD8	FZD8	NOTCH2NL	RBX1	WNT11
CITED1	FZD9	OPN1LW	RGS4	WNT16
CLDN1	GREM1	PLAUR	RHOA	WNT2
CREBBP	GSK3B	PLCB1	ROCK1	WNT2B
CSNK1A1	ID2	PLCB2	ROCK2	WNT3
CSNK1A1L	JUN	PLCB3	RP1-303F19.1	WNT3A
CSNK1E	KIF2	PLCB4	RUUBL1	WNT4
CSNK2A1	KIF3A	PORCN	SBNO1	WNT5A
CSNK2A2	KIF4A	PPARA	SDC2	WNT5B
CSNK2B	KIF5A	PPARBP	SENP2	WNT6
CTBP1	KIF7	PPARD	SFRP1	WNT7A
CTBP2	KNY	PPARG	SFRP2	WNT7B
CTNNB1	LEF1	PPARGC1A	SFRP4	WNT8A
CTNNBIP1	LGALS3	PPARGC1B	SFRP5	WNT8B
CUL1	LOC442511	PPP2CA	SIAH1	WNT9A
CXXC4	LOC644293	PPP2CB	SKP1A	WNT9B
DAAM1	LOXL2	PPP2R1A	SMAD2	WT1
DAAM2	LRP1	PPP2R1B	SMAD3	
DAB2	LRP2	PPP2R2A	SMAD4	
DALLY	LRP3	PPP2R2B	SOST	
DKK1	LRP4	PPP2R2C	SOX17	
DKK2	LRP5	PPP3CA	SOX9	
DKK4	LRP6	PPP3CB	TBL1X	
DNER	MAP3K7	PPP3CC	TBL1XR1	