

**IDENTIFICAÇÃO DE MARCADORES MOLECULARES
ÚTEIS AO DIAGNÓSTICO E PROGNÓSTICO EM
LEIOMIOSSARCOMAS E SARCOMAS
PLEOMÓRFICOS**

ROLANDO ANDRÉ RIOS VILLACIS

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

Área de concentração: Oncologia

Orientadora: Prof. Dra. Silvia Regina Rogatto

**São Paulo
2011**

FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Rios Villacis, Rolando André

Identificação de marcadores moleculares úteis ao diagnóstico e prognóstico em leiomiossarcomas e sarcomas pleomórficos / Rolando André
Rios Villacis – São Paulo, 2011.

120p.

Dissertação (Mestrado)-Fundação Antônio Prudente.

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

Orientadora: Silvia Regina Rogatto

Descritores: 1. LEIOMIOSSARCOMA. 2. HISTIOCITOMA FIBROSO MALIGNO. 3. PROGNÓSTICO. 4. DIAGNÓSTICO. 5. BIOLOGIA MOLECULAR. 6. MARCADORES BIOLÓGICOS. 7. PERFILAÇÃO DA EXPRESSÃO GÊNICA. 8. SARCOMAS PLEOMÓRFICOS.

DEDICATÓRIA

Aos meus pais,

Rolando e Silvia,

Aos meus irmãos,

Angel e Danielli

DEDICO

À Melina, cujo apoio, carinho,

amor, dedicação e

companheirismo

foram indispensáveis,

OFEREÇO

AGRADECIMENTOS

À pesquisadora e orientadora Prof. Dra. Silvia Regina Rogatto, pela confiança, auxílio e ensinamentos importantes para a realização deste trabalho.

À pesquisadora Dra. Isabela Werneck da Cunha pela colaboração valiosa na revisão histopatológica das amostras que compõem este estudo.

A todos os membros do Laboratório Neogene, pelo companheirismo, amizade e auxílio prestados durante a realização deste projeto. Em especial ao Fábio Marchi, quem realizou as análises de bioinformática, à Priscila Cirilo, quem me ajudou nos experimentos de expressão gênica e à Sara Martorelli, quem compartilhou as mesmas dificuldades e alegrias de trabalhar com sarcomas.

Ao Renato Puga, do LBHC do Hospital AC Camargo, que também contribuiu com as análises de bioinformática.

Aos funcionários do SAME do Hospital AC Camargo pela atenção e paciência durante a pesquisa de prontuários médicos, especialmente ao Luciano, Paulo e Odirlei.

Aos meus pais Rolando e Silvia por acreditarem na minha capacidade e pelo suporte emocional, mesmo à distância. Sem o seu apoio esta empreitada não seria possível.

A minha namorada Melina, pelo companheirismo, amor, incentivo e paciência, compartilhando todos os momentos alegres e difíceis no decorrer da pós graduação.

À Coordenadoria de Aperfeiçoamento de Ensino Superior (CAPES), pela concessão da bolsa de estudos;

Enfim, quero agradecer a todos que colaboraram para a realização e finalização deste trabalho.

RESUMO

Rios Villacis RA. **Identificação de marcadores moleculares úteis ao diagnóstico e prognóstico em leiomiossarcomas e sarcomas pleomórficos.** São Paulo; 2011. [Dissertação de Mestrado-Fundação Antônio Prudente]

Os sarcomas de partes moles (SPM) são neoplasias heterogêneas de origem mesenquimal que podem ser classificados por sua histogênese e comportamento clínico. Neste grupo estão incluídos os sarcomas pleomórficos (SP) e os leiomiossarcomas (LM). Os tumores que não apresentam uma linha de diferenciação celular específica fazem parte do grupo dos sarcomas pleomórficos. Estudos de perfis de expressão gênica em larga escala são utilizados como estratégia metodológica para classificar diferencialmente os sarcomas pleomórficos de outros sarcomas, principalmente de leiomiossarcomas. Até o presente não foi identificado um marcador biológico capaz de diferenciar essas neoplasias. Neste estudo foi avaliado o padrão de expressão gênica de 22 SP e 22 LM com o objetivo de contribuir para a identificação de marcadores moleculares que diferenciem estes tumores. As análises realizadas por testes supervisionados e não supervisionados com FDR=0 não foram capazes de diferenciar estes grupos tumorais. Foi observada similaridade genética quanto ao padrão de expressão gênica entre os SP e LM. Além disso, não foi possível encontrar relação de algum gene ou grupo de genes com características clínico-patológicas. Por outro lado, análises hierárquicas supervisionadas permitiram a identificação de um grupo formado apenas de LM, caracterizados pela expressão aumentada de vários genes associados com o desenvolvimento e a função muscular. Análises utilizando ferramentas para construção de redes gênicas permitiram identificar um grande número de genes correlacionados e responsáveis pelo controle de importantes funções celulares. A análise pareada entre quatro sarcomas e seus respectivos tecidos adjacentes normais revelou 40 genes diferencialmente expressos. Os genes com expressão aumentada estavam relacionados à via do *TP53* e ao ciclo celular, enquanto os genes com expressão diminuída se associavam a diferentes vias metabólicas. O padrão de

expressão dos genes *SRF*, *SRC*, *BAD* e *TOP2A* foram confirmados por ensaios de RT-PCR quantitativa em tempo real. Os resultados apresentados reforçam os obtidos na literatura que sugerem uma linhagem de diferenciação comum aos SP e LM. Além disso, em LM, foram observadas alterações específicas em genes relacionados à diferenciação muscular que podem contribuir para a caracterização diagnóstica mais acurada destes tumores quando comparados aos SP.

SUMMARY

Rios Villacis RA. **[Identification of potential molecular markers for diagnosis and prognosis in leiomyosarcomas and undifferentiated pleomorphic sarcoma]**. São Paulo; 2011. [Dissertação de Mestrado-Fundação Antônio Prudente]

Soft tissue sarcomas (STS) are heterogeneous neoplasms of mesenchymal origin, that can be classified by its histogenesis and clinical behavior. In this group are included undifferentiated pleomorphic sarcoma (UPS) and leiomyosarcoma (LMS). Tumors without a specific line of cell differentiation are included in pleomorphic sarcoma group. Microarray gene expression profile studies are used as a methodological strategy to differentiate the pleomorphic sarcoma from other sarcomas, mainly leiomyosarcomas. Currently, is unknown a biological marker to distinguish these neoplasms. The aim of this study was to evaluate the gene expression profile in 22 UPS and 22 LMS, contributing to identify biological markers able to differentiate these tumors. Supervised and non-supervised tests with FDR=0 were unable to distinguish UPS from LMS. These tumors presented a similar genetic expression profile. Furthermore, there was no relation between a gene or group of genes with clinico-pathological features. On the other hand, supervised hierarchical clustering allowed the identification of a group composed exclusively by LMS, showing overexpression of genes related to the development and muscle function. The network analysis presented several over and downexpressed central genes, including those related with important cellular functions. In a paired analysis between four sarcomas and their respective adjacent normal tissue, it was found 40 genes differentially expressed. In general, genes with downexpression were associated to different metabolic pathways, while genes overexpressed were related to *TP53* and to cell cycle pathways. The expression levels were confirmed for *SRF*, *SRC*, *BAD* and *TOP2A* genes by reverse transcriptase real-time PCR. The results presented here support the findings obtained in literature, which suggest a common differentiation lineage in UPS and LMS. Moreover, they also demonstrated that LMS has specific

alterations in genes related to muscle differentiation, contributing to a more accurate diagnosis of these tumors in comparison with UPS.

LISTA DE FIGURAS

Figura 1	Aspecto morfológico de quatro casos de sarcomas pleomórficos e leiomiiossarcomas.....	35
Figura 2	Agrupamento entre os 44 casos de sarcomas por análise não supervisionada.....	47
Figura 3	Análise supervisionada baseada no tipo tumoral resultou em 390 genes diferencialmente expressos (A) e resultou em dois grupos (B).....	49
Figura 4	Anotações das vias do KEGG (http://www.funnet.info) para os genes diferencialmente expressos entre o <i>cluster</i> de LM e o <i>cluster</i> de SP e LM.....	50
Figura 5	Resultado da análise dos dados usando a base de dados <i>Gene Ontology</i> (http://www.funnet.info) resultando nos processos biológicos (A) e na função molecular (B) a que os genes diferencialmente expressos estão envolvidos.....	51
Figura 6	Análise supervisionada dos 44 casos de sarcomas, considerando a topologia das amostras.....	53
Figura 7	Agrupamento entre os 42 casos de sarcomas por análise não supervisionada (A) e supervisionada (B), após exclusão dos casos com a síndrome de Li-Fraumeni.....	54
Figura 8	Agrupamento entre os 40 casos de sarcomas por análise não supervisionada (A) e supervisionada (B), após exclusão dos casos com metástase ao diagnóstico.....	55

Figura 9	Agrupamento entre os 38 casos de sarcomas por análise não supervisionada (A) e supervisionada (B), após exclusão dos casos com metástase ao diagnóstico e com a síndrome de Li-Fraumeni.....	56
Figura 10	Agrupamento entre os 24 casos de SP e LM mais característicos por análise não supervisionada.....	57
Figura 11	(A) Resultados da análise supervisionada entre os 24 casos de sarcomas demonstrando 127 genes diferencialmente expressos. (B) Agrupamento entre os 24 casos de sarcomas por análise supervisionada.....	58
Figura 12	Resultados da busca por processos biológicos (A) e funções moleculares (B) dos genes diferencialmente expressos usando o banco de dados <i>Gene Ontology</i> (http://www.funnet.info).....	59
Figura 13	(A) Análise de clusterização hierárquica pareada, utilizando as quatro amostras tumorais e seus respectivos tecidos adjacentes normais. (B) Anotações das vias do KEGG para os genes diferencialmente expressos nos grupos pareados, quatro tecidos adjacentes normais e quatro amostras tumorais.....	61
Figura 14	Genes diferencialmente expressos entre os 44 casos de sarcomas com $-1 > fold\ change > 0,7$ associados as doenças e funções celulares e moleculares que apresentaram os maiores scores.....	64
Figura 15	Vias canônicas correlacionando com os genes diferencialmente expressos com $-1 > fold\ change > 0,7$ em sarcomas e que apresentaram os maiores scores.....	65
Figura 16	Representação esquemática de quatro redes baseadas nos genes diferencialmente expressos ($-1 > fold\ change > 0,7$).....	67

Figura 17	Média do coeficiente de estabilidade de expressão (M). Cálculo para determinar o melhor par de genes controles, após a exclusão progressiva dos genes de expressão menos estável.....	72
Figura 18	Expressão relativa dos genes estudados por RT-qPCR e análise de expressão gênica global nas amostras de LM e SP. O gráfico mostra a distribuição de todos os valores entre as barras, com exceção de <i>outliers</i> e valores extremos. Foi utilizado o teste de Mann-Whitney para estas comparações.....	73
Figura 19	Expressão relativa do <i>TOP2A</i> por RT-qPCR e expressão gênica global nas amostras tumorais e normais. O gráfico mostra a distribuição de todos os valores entre as barras, com exceção de <i>outliers</i> e valores extremos. Foi utilizado o teste de Wilcoxon para estas comparações.....	74

LISTA DE QUADROS

Quadro 1	Principais estudos de expressão gênica em larga escala que incluíram sarcomas pleomórficos (SP) e leiomiossarcomas (LM).....	17
Quadro 2	Dados clínicos e patológicos dos pacientes com sarcomas pleomórficos (SP) e leiomiossarcomas (LM).....	30
Quadro 3	Linhagens celulares utilizadas na composição do RNA de referência universal.....	37
Quadro 4	Sequências de bases dos iniciadores utilizados nos experimentos de expressão dos transcritos alvo <i>SRF</i> , <i>SRC</i> , <i>BAD</i> e <i>TOP2A</i> por RT-qPCR.....	40
Quadro 5	Resultados da análise imunoistoquímica utilizando marcadores protéicos relacionados com a função muscular entre as amostras do presente estudo.....	46
Quadro 6	Genes com expressão aumentada relacionados com a estrutura e função muscular observado no grupo dos 16 LM.....	52
Quadro 7	Genes diferencialmente expressos entre o grupo de quatro amostras tumorais e seus respectivos tecidos adjacentes normais...	62
Quadro 8	Relação dos genes diferencialmente expressos nas quatro vias canônicas que apresentaram maiores <i>scores</i>	65
Quadro 9	Descrição dos genes com e sem alteração de expressão gênica presentes em cada uma das quatro redes gênicas com maiores <i>scores</i>	69

Quadro 10	Expressão relativa, fator de normalização (FN) e coeficiente de estabilidade de expressão (M).....	71
Quadro 11	Coeficiente de estabilidade de expressão (M).....	74

LISTA DE SIGLAS E ABREVIATURAS

Lista de siglas em inglês

aCGH	<i>Array comparative genomic hybridization</i>
BAC	<i>Bacterial artificial chromosome</i>
cDNA	<i>Complementary DNA</i>
CGH	<i>Comparative genomic hybridization</i>
cRNA	<i>Complementary RNA</i>
Ct	<i>Cycle Treshold</i>
DNA	<i>Desoxiribonucleic acid</i>
FDR	<i>False Discovery Rate</i>
FISH	<i>Fluorescent in situ hybridization</i>
FNCLCC*	<i>Federation Nationale des Centres de Lutte Contre le Cancer</i>
GIST	<i>Gastrintestinal stromal tumors</i>
GO	<i>Gene Onthology</i>
HE	<i>Hematoxiniln-Eosin</i>
HHV-8	<i>Human Herpesvirus 8</i>
HIV	<i>Human Immunodeficiency Virus</i>
IPA	<i>Ingenuity Pathway Analysis</i>
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
mRNA	<i>Messenger RNA</i>
NCCN	<i>National Comprehensive Cancer Network</i>
RT-qPCR	<i>Quantitative Reverse Transcriptase-Polymerase Chain Reaction</i>
RNA	<i>Ribonucleic acid</i>
SAM	<i>Significance Analysis for Microarray</i>
ShRNA	<i>Small hairpin RNA</i>
SPSS	<i>Statistics Packet for Social Sciences</i>
TMA	<i>Tissue Microarrays</i>
WHO	<i>World Health Organization</i>

* Em francês

LISTA DE SIGLAS EM PORTUGUÊS

AML	Actina de músculo liso
DEPC	Dietil-pirocarbonato
FHM	Fibrohistiocitoma maligno
FN	Fator de normalização
IHQ	Imunoistoquímica
LM	Leiomiossarcoma
LP	Lipossarcoma
PVC	Policloreto de vinilo
QR	Quantificação relative
SOE	Sem outra especificação
SP	Sarcoma pleomórfico
SPM	Sarcoma de partes moles

ÍNDICE

1	INTRODUÇÃO	1
1.1	Sarcomas de partes moles (SPM)	1
1.1.1	Características clínicas e patológicas	1
1.1.2	Etiologia	2
1.1.3	Diagnóstico	3
1.1.4	Tratamento	4
1.1.5	Classificação	5
1.1.6	Expressão gênica em larga escala	5
1.1.7	Perspectivas futuras	8
1.2	Sarcomas pleomórficos e leiomiossarcomas	8
1.2.1	Sarcomas pleomórficos	8
1.2.2	Leiomiossarcomas	10
1.2.3	Alterações genéticas em sarcomas pleomórficos e leiomiossarcomas	12
1.2.4	Estudos de expressão gênica em larga escala	14
2	JUSTIFICATIVA	26
3	OBJETIVOS	27
4	MATERIAL E MÉTODOS	28
4.1	Caracterização da amostra	28
4.2	Casuística	33
4.3	Extração de RNA	35
4.4	Expressão gênica global (<i>oligoarrays</i>)	36
4.5	RT-PCR quantitativa em tempo real (RT-qPCR)	38
4.5.1	Obtenção de cDNA	38
4.5.2	RT-qPCR em tempo real	39
4.6	Análises estatísticas	42
4.6.1	Dados de expressão gênica global	42

4.6.2	Dados de RT-qPCR em tempo real.....	44
5	RESULTADOS	45
6	DISCUSSÃO	75
7	CONCLUSÃO	99
8	REFERÊNCIAS BIBLIOGRÁFICAS	100

ANEXOS

- Anexo 1** Lista de genes diferencialmente expressos obtidos por análise supervisionada de todos os 44 casos de sarcomas.
- Anexo 2** Lista de genes diferencialmente expressos obtidos por análise supervisionada dos casos (42) após exclusão dos casos com a síndrome de Li-Fraumeni.
- Anexo 3** Lista de genes diferencialmente expressos obtidos por análise supervisionada dos casos (40) após exclusão dos casos com a metástase ao diagnóstico.
- Anexo 4** Lista de genes diferencialmente expressos obtidos por análise supervisionada dos casos (38) após exclusão dos casos com a síndrome de Li-Fraumeni e metástase ao diagnóstico.
- Anexo 5** Lista de genes diferencialmente expressos obtidos por análise supervisionada dos casos (24) de SP e LM mais característicos.

1 INTRODUÇÃO

1.1 SARCOMAS DE PARTES MOLES (SPM)

1.1.1 Características clínicas e patológicas

Os sarcomas de partes moles (SPM) compreendem um grupo heterogêneo e agressivo de neoplasias malignas derivadas de tecidos mesenquimais. Estes tumores mostram grande variação na morfologia, em aspectos genéticos e comportamento clínico (BECK et al. 2010b; GUILLOU e AURIAS 2010). São compostos por mais de 50 tipos histológicos e representam aproximadamente 1% de todas as neoplasias malignas em adultos e 10-15% em crianças e adolescentes (BLAY e CESNE 2009; THWAY 2009). Todos os grupos etários são afetados, com ligeira predominância do sexo masculino em uma proporção de 1,2:1,0 para o sexo feminino, no entanto, mais de 40% dos SPM se manifestam em pacientes acima de 55 anos (SCHMITT e KASPER 2009). TORO et al. (2006) relataram uma incidência aproximada de 4,5-6,5 casos/100 000 habitantes nos EUA numa análise com mais de 25 000 casos de SPM. Os SPM podem surgir em qualquer parte do corpo humano, principalmente nas extremidades (60%), trato gastro-intestinal (GIST, 25%), retroperitônio (15-20%) e região da cabeça e pescoço (National Comprehensive Cancer Network-NCCN 2009).

Pacientes com SPM apresentam uma sobrevida estimada de 5 anos em aproximadamente 50% dos casos (CLARK et al. 2005). Este prognóstico se deve principalmente ao risco de desenvolvimento de metástases a distância ou à presença de metástases ocultas ao diagnóstico (LEHNHARDT et al. 2009). Acredita-se que as

metástases em sarcomas ocorrem em sua maioria por via hematogênica, sendo raras as metástases linfonodais (menos de 5% dos casos), a exceção de alguns subtipos, como o sarcoma sinovial e angiossarcoma (SCHMITT e KASPER 2009; THWAY 2009). Aproximadamente 40-50% dos pacientes com sarcomas de alto grau desenvolvem metástases, principalmente para os pulmões e ossos, enquanto que os pacientes com sarcomas viscerais e retroperitoneais tendem a ter metástases no fígado e peritônio (BRENNAN et al. 2008; COLLINI et al. 2009; LEHNHARDT et al. 2009). Até o presente, o grau histológico é o fator prognóstico mais importante, tanto para prever metástases como para a sobrevida global (PISTERS et al. 1996; CHIBON et al. 2010). A sobrevida média dos pacientes com diagnóstico de doença metastática é de 8-12 meses (WEITZ et al. 2003).

1.1.2 Etiologia

A etiologia da maioria dos sarcomas é desconhecida e grande parte dos estudos etiológicos foi baseado em séries pequenas de pacientes ou até mesmo em relatos de casos. Os modelos atuais sugerem que os SPM não se originam a partir da desdiferenciação de tumores benignos pré-existent, mas surgem *de novo* (SCHMITT e KASPER 2009). Embora os SPM sejam raros e apareçam espontaneamente, alguns fatores de risco foram associados ao desenvolvimento destas neoplasias. Pacientes com doenças hereditárias (neurofibromatose tipo 1, Síndrome de Li-Fraumeni e retinoblastoma hereditário) apresentam maior predisposição a desenvolver certos SPM. Alguns sarcomas estão associados com infecções virais, como o sarcoma de Kaposi (HIV e HHV-8) e um subgrupo de leiomiossarcomas (EBV). A exposição à radiação iônica e a alguns fatores químicos,

como o PVC (policloreto de vinilo) e herbicidas, foram associados com o desenvolvimento de alguns SPM, como sarcoma sinovial, fibrossarcoma e sarcoma pleomórfico. Linfedemas crônicos foram associados ao aparecimento de angiossarcomas (FROEHNER e WIRTH 2001; DEYRUP et al. 2006; SKUBITZ e D'ADAMO 2007; SCOGGINS e PISTERS 2008; THWAY 2009).

1.1.3 Diagnóstico

A raridade, a natureza pleomórfica e o grande número de entidades morfológicas que se sobrepõem, transformam o diagnóstico dos SPM em um desafio para os patologistas, dificultando sua classificação (BECK et al. 2010b). Atualmente, os sarcomas são diagnosticados baseados na histopatologia, complementada pelo uso de técnicas moleculares e de imunoistoquímica em casos selecionados (GRAADT et al. 1999; FLETCHER et al. 2002; WEST 2010). A imunoistoquímica é utilizada para confirmar a linhagem de diferenciação e técnicas moleculares podem identificar translocações cromossômicas específicas ou mutações (NIELSEN e WEST 2010). Enquanto algumas neoplasias, como o GIST, apresentam na maioria das vezes características diagnósticas bem definidas (localização anatômica, morfologia, fenótipo imunoistoquímico e mutações genéticas chave), outros tumores, como os leiomiossarcomas, são definidos por características morfológicas e alguns poucos biomarcadores imunoistoquímicos inespecíficos (WEST 2010). A classificação correta dos SPM é essencial, pois entidades diagnósticas distintas podem requerer estratégias diferentes para um tratamento apropriado (NIELSEN e WEST 2010; WARDELMANN 2010).

1.1.4 Tratamento

O tratamento principal para pacientes com SPM é a ressecção cirúrgica completa do tumor primário, com ou sem radioterapia, que resulta em uma taxa de cura de 50% (GANJOO e JACOBS 2010). As chances de sobrevida são maiores na ausência de comprometimento neoplásico das margens cirúrgicas. A radioterapia é utilizada como forma de reduzir a incidência de recorrência local. A quimioterapia pode ser utilizada com intenção paliativa em tumores avançados (metastáticos) ou para reduzir o tamanho da neoplasia, possibilitando a cirurgia radical da lesão (ERIKSSON 2010). A doxorrubicina e a ifosfamida são os quimioterápicos que apresentam as melhores taxas de resposta em SPM (20-30%), contudo, o tratamento quimioterápico traz benefícios limitados e geralmente não melhora a sobrevida dos pacientes, a exceção de alguns subtipos que acometem crianças e adultos jovens, como o osteossarcoma e sarcoma de Ewing (BECK et al. 2010b; GANJOO e JACOBS 2010; NIELSEN e WEST 2010). Devido a sua raridade, os SPM, com raras exceções, foram agrupados como uma única entidade para fins de estadiamento, prognóstico e investigação clínica, com o objetivo de alcançar significância estatística (WUNDER et al. 2007; VERWEIJ e BAKER 2010). No entanto, análises moleculares revelam claramente diferentes perfis genéticos entre os SPM. Um exemplo modelo é a utilização do imatinib para o tratamento de tumores gastro-intestinais (GIST) (NIELSEN e WEST 2010; VERWEIJ e BAKER 2010).

1.1.5 Classificação

A aplicação de técnicas em citogenética clássica e molecular permitiu classificar os SPM em duas categorias principais: (1) tumores que apresentam alterações genéticas específicas e cariótipos simples, incluindo a fusão de genes devido a translocações recíprocas; representam 15-20% dos tumores mesenquimais (Ex., *EWSR1-FLI1* em sarcoma de Ewing, *SYT-SSX* em sarcoma sinovial, *TLS-CHOP* em lipossarcoma mixóide e *PAX3-FKHR* em rabdomiossarcoma alveolar), e mutações somáticas específicas (*KIT* e *PDGFRA* em tumores estromais gastrointestinais) (MITELMAN et al. 2007; NIELSEN e WEST 2010); (2) tumores com alterações genéticas inespecíficas e cariótipos complexos, que apresentam características histológicas e imunoistoquímicas semelhantes. Este grupo representa 50% dos SPM e é composto principalmente por tumores pleomórficos de células alongadas, incluindo leiomiossarcomas, sarcomas pleomórficos e fibrossarcomas. Adicionalmente, alguns lipossarcomas desdiferenciados apresentam tanto anormalidades cromossômicas específicas, como cariótipos complexos e aleatórios (IWASAKI et al. 2009; BOVÉE e HOGENDOORN 2010; GUILLOU e AURIAS 2010).

1.1.6 Expressão gênica em larga escala

O câncer é uma doença relacionada com alterações nas funções de genes associados a proliferação celular, entre outras funções. Independentemente da origem destas alterações (mutações em ponto, deleções, translocações, amplificações gênicas ou alterações no padrão de metilação), o fenótipo maligno é mediado pelo padrão característico de expressão gênica (ANTONESCU 2008). Características

morfológicas são geralmente bons indicadores do comportamento biológico, contudo, a propensão a metástases, crescimento tumoral e resposta à quimioterapia podem ser determinados pelos padrões de expressão gênica que não são bem correlacionados com a morfologia (SKUBITZ et al. 2008).

A tecnologia de *arrays* de expressão gênica foi desenvolvida em meados da década de 1990 e permite avaliar os níveis de expressão de milhares de genes simultaneamente em amostras tumorais e é uma das formas mais eficientes para detectar novos marcadores tumorais específicos (SCHENA et al. 1995; LOCKHART e WINZELER 2000; VAN DE RIJN e FLETCHER 2006; ANTONESCU 2008; CONLEY et al. 2008). As principais plataformas de *arrays* utilizadas nos estudos de expressão gênica envolvem a deposição robótica de fragmentos de DNA em lâminas de vidro ou a síntese *in situ* de oligonucleotídeos em matrizes de alta densidade utilizando fotolitografia (SCHENA et al. 1995). Embora todas as células tenham os mesmos genes, apenas 5% deles são expressos em uma célula em particular em um determinado momento, desta forma, os estudos de expressão gênica proporcionam informações privilegiadas do estado e da função das células tumorais (OSUNA e ÁLAVA 2009). Tumores com histologia e morfologia parecidas nem sempre apresentam comportamento biológico semelhante, com discrepâncias na taxa de crescimento, resposta ao tratamento e potencial metastático. Em nível molecular, tumores aparentemente idênticos podem apresentar padrões de expressão gênica diferentes (THWAY 2009).

Em sarcomas, os objetivos principais dos estudos de perfis de expressão gênica são: (1) entender melhor as vias biológicas anômalas que contribuem com a sarcomatogênese, (2) aperfeiçoar o sistema de classificação dos sarcomas para

refletir melhor a biologia destes tumores e (3) identificar biomarcadores para ampliar e melhorar o diagnóstico, prognóstico e resposta terapêutica (TSCHOEP et al. 2007; BECK et al. 2010b). Estudos de expressão gênica permitiram a separação de tumores em grupos com distintos padrões de expressão, incluindo rabdomiossarcomas e sarcomas de Ewing (DAVICIONI et al. 2006; HANCOCK e LESSNICK 2008; DAVICIONI et al. 2009). A combinação desta tecnologia com a investigação da expressão de proteínas em microarranjos de tecidos (do inglês, *tissue microarray* - TMA) permitiu a identificação da apolipoproteína D como marcador diagnóstico em dermatofibrossarcoma protuberans e da proteína TLE1 em sarcomas sinoviais (WEST et al. 2004b; TERRY et al. 2007). Analisando o padrão de expressão gênica, WEST et al. (2004a) relataram que o gene *DOG1* (*TMEM16A*) era expresso especificamente em GIST e validaram os resultados com ensaios de TMA e FISH. Estudos posteriores confirmaram que os anticorpos anti-DOG1 são mais sensíveis e específicos para diagnosticar GIST do que os anticorpos anti-KIT (ESPINOSA et al. 2008; MIETTINEN et al. 2009).

Os estudos de expressão gênica em SPM têm sido limitados pela escassez de material congelado, pelo grande número de tipos histológicos reconhecidos e pelos custos experimentais. Como resultado, a maioria destas pesquisas inclui um número relativamente pequeno de amostras por tipo histológico (3 a 10 casos). O número limitado de espécimens e a utilização de amostras de várias localizações anatômicas, sem relação biológica, afetam dramaticamente a habilidade das pesquisas em SPM de identificar subgrupos e alterações recorrentes em nível molecular. Além disso, a grande maioria destes estudos não informa se as amostras são obtidas de pacientes pré ou pós-tratamento quimioterápico e ou radioterápico. Isto contrasta com os

resultados obtidos em outras classes tumorais, como os carcinomas, em que há uma maior incidência da doença e é possível restringir a pesquisa a um único sítio anatômico (WEST 2010; WEIGELT et al. 2010).

1.1.7 Perspectivas futuras

Embora a histologia, combinada com a imunoistoquímica, continue sendo o “padrão ouro” no diagnóstico dos SPM, técnicas de biologia molecular que detectam alterações genéticas específicas são ferramentas auxiliares importantes na rotina diagnóstica. A maioria dos sarcomas de alto grau (que incluem mixofibrossarcomas, leiomiossarcomas e sarcomas pleomórficos) carecem de alterações genéticas específicas, dificultando seu diagnóstico e a identificação de potenciais alvos terapêuticos (NIELSEN e WEST 2010). Pesquisas em nível molecular, incluindo estudos de perfil de expressão gênica, permitirão um enorme progresso na compreensão da biologia dos sarcomas, e prometem mudanças significativas na forma como os SPM são diagnosticados e classificados. O aumento do conhecimento sobre as vias envolvidas na sarcomatogênese e na progressão destas neoplasias pode levar ao desenvolvimento de novos indicadores prognósticos, biomarcadores e alvos terapêuticos moleculares (THWAY et al. 2009).

1.2 SARCOMAS PLEOMÓRFICOS E LEIOMIOSSARCOMAS

1.2.1 Sarcomas pleomórficos

Os sarcomas pleomórficos (SP) são tumores compostos por fibroblastos (células alongadas), histiócitos (células poliédricas), células indiferenciadas, mitoses

anormais e necrose, arrançados em um padrão de crescimento estoriforme (ROSENBERG 2003; SCAPOLAN et al. 2008; THWAY 2009). Estas neoplasias ocorrem geralmente em adultos de meia-idade, acomete as extremidades e apresentam comportamento altamente agressivo (WEISS e ENZINGER 1978; THWAY 2009). O termo *sarcoma pleomórfico* foi proposto pela Organização Mundial da Saúde (WHO) em 2002 para substituir a denominação antiga de “Fibrohistiocitoma Maligno” (FHM) (FLETCHER 2006; DEI TOS 2006).

O termo *fibrohistiocitoma maligno* foi introduzido em meados dos anos 60 para designar um grupo de neoplasias que apresentavam histiócitos (macrófagos) que adquiriam características estruturais e funcionais de fibroblastos, sendo classificados em cinco subtipos histológicos (OZZELLO et al. 1963; O'BRIEN e STOUT 1964; MATUSHANSKY et al. 2009). O surgimento destes tumores a partir de histiócitos, fibroblastos ou células mesenquimais indiferenciadas ainda é desconhecido, porém, estudos mais recentes sugerem que os SP derivam de células da medula óssea, provavelmente de células tronco mesenquimais (MATUSHANSKY et al. 2007; LI Q et al. 2010).

A existência do SP como entidade clínico-patológica bem definida se tornou rapidamente popular após o trabalho de WEISS e ENZINGER (1978). Em meados dos anos 1980 o SP (FHM, como era designado) representava o diagnóstico de sarcoma mais comum em adultos (ENZINGER 1986; FLETCHER 1992). Contudo, uma série de estudos começaram a questionar a existência do SP como entidade distinta (SNOVER et al. 1982; BROOKS 1986; DEHNER 1988), cujo ápice foi o trabalho publicado por FLETCHER (1992), que analisou uma série de 159 “FHM” e demonstrou que 63% dos casos poderiam ser reclassificados (com o apoio da

imunoistoquímica e microscopia eletrônica) em tipos específicos de sarcomas pleomórficos, e 12,6% como tumores não mesenquimais (DEI TOS 2006). Subsequentemente, a incidência dos SP como entidade diagnóstica diminuiu drasticamente (FLETCHER et al. 2002).

Atualmente, os SP são vistos como um grupo heterogêneo de tumores pobremente diferenciados sem uma linhagem específica, que inclui variantes de leiomiossarcomas, lipossarcomas e mais raramente rabdomiossarcomas (THWAY 2009). O diagnóstico é realizado por exclusão (DEI TOS 2006; FLETCHER 2006). Até uma década atrás, o SP do subtipo estoriforme-pleomórfico representava 40% dos SPM em adultos, contudo, após a alteração conceitual, este subtipo (atual SP com alto grau de indiferenciação), representa apenas 5-7% das neoplasias em tecidos moles (DEI TOS 2006. FLETCHER 2006; GUILLOU e AURIAS 2010).

1.2.2 Leiomiossarcomas

Os leiomiossarcomas (LM) são neoplasias malignas heterogêneas com diferenciação de músculo liso, compostas por células alongadas com abundante citoplasma eosinófilo e núcleo alongado (YANG et al. 2009). Os LM pobremente diferenciados apresentam grande atividade mitótica, necrose e pleomorfismo nuclear (REN et al. 2003). A grande maioria destes sarcomas expressa a proteína SMA (actina de músculo liso), desmina e caldesmon em ensaios de imunoistoquímica, contudo, nenhum destes marcadores é absolutamente específico para a diferenciação em músculo liso e diagnóstico de LM, podendo ser expressos por outros tumores de partes moles (YANG et al. 2010; WEST 2010). Em casos pobremente diferenciados,

estes marcadores são expressos em menor intensidade ou se encontram ausentes (LARRAMENDY et al. 2008).

Os LM acometem principalmente adultos de meia idade (50-60 anos), adultos jovens e até mesmo crianças, sendo mais comuns em mulheres, mesmo em localizações não-uterinas (LARRAMENDY et al. 2006). Estas neoplasias podem se desenvolver em qualquer parte do corpo, assim como em órgãos viscerais (próstata, bexiga e útero) (GUILLOU e AURIAS 2010). Excluindo-se os LM uterinos, os LM de partes moles são raros (aproximadamente 5-10% dos SPM) e constituem uma percentagem significativa dos sarcomas de retroperitônio e pelvis (JEMAL et al. 2009; THWAY 2009). Em extremidades, os LM são menos comuns e respondem por 10-15% de todos os casos de sarcomas neste sítio anatômico (YANG et al. 2009).

Os LM que se desenvolvem na derme são menos agressivos e raramente metastatizam, já os LM profundos, principalmente de alto grau histológico, estão associados a uma maior taxa de mortalidade, devido ao aumento do risco de recorrência local e metástases, principalmente para o fígado e pulmão (WANG et al. 2001; FARSHID et al. 2002; YANG et al. 2009; GUILLOU e AURIAS 2010). Até o presente, o grau histológico, o tamanho do tumor e a localização anatômica são os únicos fatores patológicos que são independentemente preditivos da evolução dos pacientes (FARSHID et al. 2002; SVARVAR et al. 2007). O tratamento padrão para os LM é a ressecção da lesão, já o tratamento por quimioterapia citotóxica convencional se mostra ineficiente, com resposta em 50% dos pacientes e sobrevida a longo prazo menor que 10% (TRENT et al. 2007).

Após a última reclassificação da WHO (2002), que considerou os “FHM” meramente como sinônimo de sarcomas pleomórficos indiferenciados (SP), os LM

passaram a representar a segunda neoplasia de partes moles mais comum em adultos após os lipossarcomas (DEI TOS 2006; SKUBITZ e D'ADAMO, 2007).

1.2.3 Alterações Genéticas em Sarcomas Pleomórficos e Leiomiossarcomas

Dados citogenéticos em SP e LM mostraram cariótipos com múltiplas anormalidades numéricas e estruturais (cromossomos em anel, cromossomos dicêntricos, *double minutes* e regiões homogeneamente coradas). Ganhos, perdas e amplificações cromossômicas são comuns a estas neoplasias. Apesar dos rearranjos cariotípicos complexos, que podem envolver 30-35% do genoma, nenhuma alteração específica foi identificada para ser utilizada no diagnóstico destes tumores (MERTENS et al. 1998; LARRAMENDY et al. 2006; YANG et al. 2009; MITELMAN et al. 2010).

Estudos de CGH (do inglês, *comparative genomic hybridization*) cromossômico em SP revelaram ganhos em 1p31-p32, 1q21-q22, 5p15, 7q21-q31, 12q12-q15, 17q23, 19p e 20p, e regiões comuns de perdas cromossômicas em 1q32, 2q32-q35, 9p21-pter, 10q e principalmente 13q21-q22 (FORUS et al. 1995; LARRAMENDY et al. 1997; MAIRAL et al. 1999; PARENTE et al. 1999; SIMONS et al. 2000; WENG et al. 2003; LARRAMENDY et al. 2008). Ganhos em 1p31 foram associados com a diminuição da sobrevida global e ganhos em 7q32 com o aumento do risco de metástases e recorrência local (LARRAMENDY et al. 1997). Baseados em estudos anteriores, MAIRAL et al. (1999) sugeriram o gene *RB1* (13q14) como candidato a supressor tumoral nestas neoplasias. Por outro lado, SIMONS et al. (2000) propuseram a perda do gene *CDKN2A* (9p21) como ponto crítico no desenvolvimento dos SP. WENG et al. (2003) relacionaram o ganho em

17q com o aumento da sobrevida livre de doença e com risco diminuído de metástases nestes tumores.

Em LM, estudos de CGH revelaram perdas nas regiões 2p, 2q, 10q, 11q e 13q, assim como ganhos nas regiões 1p, 5q, 8q e 17p (PACKENHAM et al. 1997; EL-RIFAI et al. 1998; LEVY et al. 2000; OTAÑO-JOOS et al. 2000; WANG et al. 2001; DERRÉ et al. 2001). Perdas nas regiões 2p, 10q e 12p, assim como ganhos em 1q, 5p e 17p foram associados com o desenvolvimento mais agressivo dos LM (HU et al. 2001; HU et al. 2005). Por outro lado, perdas na região 13q14-q21 e ganhos em 5p14-pter foram relacionadas à sobrevida diminuída (WANG R et al. 2003).

As similaridades genéticas e a tendência de sobreposição dos fenótipos imunoistoquímicos sugerem que os SP e LM compartilham vias oncogênicas comuns, correspondendo a estágios distintos de diferenciação da mesma entidade tumoral e que a maioria dos SP seriam LM pobremente diferenciados (DERRÉ et al. 2001; CHIBON et al. 2003; LARRAMENDY et al. 2008; CARNEIRO et al. 2009; PÉROT et al. 2009; PÉROT et al. 2010). Algumas alterações genéticas são mais frequentemente observadas em ambos os grupos, como a deleção do gene *RB1* no cromossomo 13 e a deleção da região 10q, enquanto a amplificação de 17p é mais específica em LM (CHIBON et al. 2000; OTAÑO-JOOS et al. 2000; HU et al. 2005).

DERRÉ et al. (2001) avaliaram 27 LM utilizando CGH e compararam as alterações genômicas mais frequentes com os resultados obtidos em estudo prévio de SP. Os autores demonstraram que ambos os tipos de neoplasias apresentavam alterações cromossômicas similares. Da mesma forma, LARRAMENDY et al. (2008) analisaram 102 SP e 82 LM pela técnica de CGH e encontraram um perfil

semelhante de alterações genéticas entre os dois grupos tumorais. A análise de clusterização não foi capaz de diferenciar estes dois grupos de tumores.

Estudos pelas técnicas de CGH, imunoistoquímica e FISH (do inglês, *fluorescence in situ hybridization*) também revelaram perfis genômicos semelhantes entre SP e lipossarcomas (LP), principalmente a amplificação da região 12q13-q15 (característica de LP), reclassificando muitos casos de SP em LP (CHIBON et al. 2002; COINDRE et al. 2003; COINDRE et al. 2004). Análises de agrupamentos não separam os SP dos LP e sugerem que a maioria dos SP, localizados no abdômen e retroperitônio, corresponderiam a LP indiferenciados (CHIBON et al. 2003; IBDAIH et al. 2005).

1.2.4 Estudos de Expressão Gênica em Larga Escala

A Quadro 1 resume os principais estudos que utilizaram análises em larga escala de expressão de transcritos (cDNA-array ou *oligoarray*) em SP e LM. Recentemente, o grupo francês publicou três artigos em sarcomas, utilizando a mesma plataforma de expressão gênica, *Human Genome U133 Plus 2.0 array* (Affymetrix). CHIBON et al. (2010) analisaram os perfis genômicos e de expressão gênica em 183 sarcomas de genótipo complexo (como em LM e SP), onde ocorrem múltiplas alterações genômicas, e identificaram uma assinatura de expressão gênica, composta por 67 genes (CINSARC), que é capaz de prever o aparecimento de metástases nestes sarcomas. Os autores validaram os resultados em um grupo independente de 127 sarcomas de genótipo complexo. PÉROT et al. (2010) utilizaram as técnicas de aCGH, sequenciamento e expressão gênica para avaliar o estado do *TP53*, *p14*, *p15* e *p16* em uma série de 34 LM e 109 SP. Os autores relataram que deleções e mutações eram mais frequentes em LM; 50% destes

tumores apresentavam inativação bialélica do *TP53*, enquanto a maioria dos SP retinha um alelo selvagem (87,2%). Adicionalmente, em tumores sem mutação no *TP53* foram detectadas deleção de ou co-inibição de *p15* e *p16* com o *p14*. GIBAULT et al. (2011) utilizaram uma análise integrada de aCGH e dados de expressão gênica para estudar 160 sarcomas com genótipo complexo. Análise não supervisionada dos dados de expressão gênica permitiu identificar cinco subgrupos tumorais, um deles composto por LM bem diferenciados (grupo A) e os outros quatro formados por SP, LM pobremente diferenciados e outros tumores pleomórficos (B-D). A deleção do gene *PTEN* foi um evento recorrente nos grupos A, C e D. Já o gene *DKK1* apresentou expressão aumentada nos grupos D e E, e redução significativa nos grupos A, B e C.

Dois estudos, SHMULEVICH et al. (2002) e FRANCIS et al. (2005), avaliaram se a heterogeneidade intratumoral interferia no perfil de expressão gênica em LM e SP. Análises de clusterização hierárquica, utilizando o coeficiente de correlação de Pearson e a distância Euclidiana, demonstraram variações intratumorais mínimas se comparadas com as variações entre tecidos. As amostras provenientes de um mesmo paciente, LM ou SP, estavam presentes em um mesmo grupo. Os autores concluíram que análises confiáveis de expressão gênica podem ser realizadas em amostras tumorais únicas, porém, a heterogeneidade intratumoral pode interferir em resultados obtidos a partir de séries pequenas, limitando a validade dos achados. Em séries amplas de tumores, os efeitos relacionados à heterogeneidade tumoral e variabilidade experimental são minimizados, aumentando a probabilidade da obtenção de padrões de expressão reprodutíveis.

MATUSHANSKY et al. (2007) compararam o padrão de expressão gênica de células tronco mesenquimais com perfis genéticos de sarcomas de outro estudo e mostraram que a assinatura genética das células tronco estava associada significativa e exclusivamente com os SP. DAIGELER et al. (2009) analisaram 10 amostras de SP pela metodologia de *oligoarray* e correlacionaram os resultados obtidos com dados clinico-patológicos. Os resultados não mostraram nenhum tipo de correlação entre a clusterização dos tumores e qualquer achado clínico ou histopatológico. No entanto, foi observada uma variedade de genes com expressão aumentada (*SLC39A14*, *SLC2A3*, *SH3*, *ENO2* e *MMP-16*) e diminuída (*CD36*, *LPL*, *GATA3*), que foram associados com pior prognóstico, desenvolvimento de metástases e recorrência da doença.

Quadro 1 - Principais estudos de expressão gênica em larga escala que incluíram sarcomas pleomórficos (SP) e leiomiossarcomas (LM).

Entidade Tumoral	Metodologia de Expressão Gênica/aCGH	Plataforma - Número de sequências plotadas	Validação dos resultados	Referência
3LM	cDNA-array	2K	_____	Shmulevich et al. (2002)
14SP , 13SS, 10 LM , 3 LPP, 3LPD e 4SW	cDNA-array	23K	RT-PCR semi-quantitativa e RT-qPCR em tempo real	Nagayama et al. (2002)
41 SPM, incluindo 8 SS monofásico, 1 LPD, 1 LPM, 2 LPP, 11 LM , 8 SP , 2 SW e 8 GIST	cDNA-array	22K e 42K	_____	Nielsen et al. (2002)
9 SP , 9 LM e 9 SS	cDNA-array	5K	_____	Lee et al. (2003)
11 LM (5 do retroperitônio, 5 de tecidos moles de extremidades e 1 de pescoço).	Oligoarray	Affymetrix-22K	RT-PCR	Ren et al. (2003)
11 SP , 8 FBS, 6 LM , 4 LP de células redondas, 3 LPP, 5 LPD, 4 SCC, 5 SS e 5 GIST	Oligoarray	Affymetrix-12K	In silico (HOOCV)	Segal et al. (2003)
4 LMU, 4 LM , 19 LOU, 46 MN e 250 outras amostras oriundas de 18 tipos de tecidos diferentes.	Oligoarray	Affymetrix-60K	_____	Skubitz e Skubitz (2003)
37 LM (incluindo tumores primários, recorrência local de tumores primários e sarcomas metastáticos).	cDNA-array	5K	In silico (LOOCV)	Lee et al. (2004)
4 SS, 11LP, 8 SW, 9 LM , 12 FB, 3 GIST, 2 DFBP, 1 SE, 11 MEN, 3 MES, 29 SP	Oligoarray	Affymetrix-40K	_____	Skubitz e Skubitz (2004)
16 classes de SPM, incluindo 19 SEW, 7 FBS, 5 GIST, 17 LM , 33LP e 38 SP	cDNA-array	Agilent-12K	IHQ	Baird et al. (2005)
9 SP , 6 LM , 4 SS, 7 FBM, 2 GIST, 4 LPM, 3 LPD, 3 LPP e 14 Tecidos Normais	Oligoarray	Affymetrix-22K	Cultura de células, IHQ e IFC	Detwiller et al. (2005)
21 SP e 17 LM	cDNA-array	27K	_____	Francis et al. (2005)

96 SPM, incluindo 4 SW, 3 LPBD, 3 SP, 9 LM , 11 OS e 10 SS monofásico	Oligoarray	Affymetrix-22K	In silico (LOOCV)	Henderson et al. (2005)
177 SPM de 13 clases, incluindo 40 LM, 64 SP , 32 SS, 8 TMBNP, 3 GIST, 4 LPM e 1 SE	cDNA-array	27K	In silico (LOOCV)	Francis et al. (2007)
105 SPM, incluindo 16 SS, 19 LPM, 3 LPBD, 15 LPD, 15 FBSM, 6 LM, 4 FBS e 21 SP	Oligoarray	Affymetrix-22K	RT-qPCR em tempo real	Nakayama et al. (2007)
37 GIST e 31 LM	Oligoarray	Agilent-44K	RT-qPCR em tempo real	Price et al. (2007)
8 GIST, 8 LM , 7 SS, 8 FB, 5 TFS, 5 DFBP e 10 TTCG	Oligoarray	HEEBO-44K	IHQ	Lee et al. (2008)
9 LM, 16 SP , 12 LP, 4 SS e 12FA	Oligoarray	Affymetrix-40K	_____	Skubitz et al. (2008)
10 SP	Oligoarray	Affymetrix-22K	_____	Daigeler et al. (2009)
18 LM e 31 SP	cDNA-array/BAC-array	27K/42K	Dados de expressão gênica de 7 LP	Carneiro et al. (2009)
8 GIST, 16 LM ; 7 SS, 8 FB, 5 TFS, 5 DFBP e 10 TTCG	Oligoarray	HEEBO-44K	FISH e IHQ	Espinosa et al. (2009)
16 LM e 16 SP	Oligoarray/BAC-array	Affymetrix-47K	RT-qPCR em tempo real, FISH, Western blot, siRNA e shRNA	Pérot et al. (2009)
51 LM	Oligoarray/Oligoarray	Agilent-44K	IHQ	Beck et al. (2010a)
71 SP, 52 LM , 44 LPD e 16 outros SPM	Oligoarray/BAC-array	Affymetrix-47K/4K	Subgrupo de 127 SPM	Chibon et al. (2010)
31 LM e 109 SP	Oligoarray/BAC-array	Affymetrix-47K/4K	RT-qPCR em tempo real, FISH, Western blot e IHQ	Pérot et al. (2010)

32 GIST e 25 LM	Oligoarray/Oligoarray	Agilent-60K/Agilent-44K	_____	Yang et al. (2010)
42 GIST e 30 LM	Oligoarray/Oligoarray	Agilent-44K	_____	Ylipaa et al. (2010)
51 LM, 70 SP , 22 MF, 9 LPP e 8 RML	Oligoarray/BAC-array	Affymetrix-47K	_____	Gibault et al. (2011)

DFBP: Dermatofibrossarcoma protuberans; FA: fibromatose agressiva; FB: Fibromatose; FBS: fibrossarcoma; FBSM: fibrossarcoma mixóide; GIST: tumores gastrointestinais; HOOCV: *hold-one-out cross validation*; IFC: Imunofluorescência; IHQ: imunoistoquímica; LOU: leiomioma uterino; LM: leiomiossarcoma; LMU: leiomiossarcoma uterino; LOOCV: *leave-one out cross-validation*; LP: lipossarcoma; LPBD: lipossarcoma bem diferenciado; LPD: lipossarcoma desdiferenciado; LPM: lipossarcoma mixóide; LPP: lipossarcoma pleomórfico; MEN: Meningioma; MES: Mesotelioma; MF: Mixofibrossarcoma; MN: miométrio normal; OS: Osteossarcoma; RML: rabdomiossarcoma pleomórfico; SCC: sarcoma de células claras; SE: Sarcoma epitelióide; SEW: Sarcoma de Ewing; shRNA: *small airping RNA*; siRNA: *small interfering RNA*; SPM: Sarcomas de partes moles; SP: sarcoma pleomórfico; SS: sarcoma sinovial; SW: schwanoma; TFS: Tumor fibroso solitário; TMBNP: Tumores malignos de bainha de nervo periférico; TTCCG: Tumores tenosinoviais de células gigantes.

Em negrito: SP e LM utilizados nos estudos de expressão gênica.

Em LM, REN et al. (2003) avaliaram a expressão gênica de 35 amostras de tecidos de origem mesenquimal, incluindo 11 casos de LMs de diferentes graus histológicos e sítios anatômicos, e identificaram 92 genes que distinguiam LM de baixo grau dos LM de alto grau e LM metastáticos. SKUBITZ e SKUBITZ (2003) utilizaram a plataforma da Affymetrix de 60K e identificaram um grupo de genes diferencialmente expressos em LM, leiomiomas e endométrio normal. Seis genes (*CDKN2A*, *DIAPH3*, *DCX*, *CAPN6*, *IL17B* e *PLP1*) apresentaram expressão aumentada em LM se comparados com o endométrio normal e 18 outros tecidos. Mais recentemente, BECK et al. (2010a) combinaram os resultados dos perfis de expressão gênica, aCGH, e imunoistoquímica para identificar e caracterizar subtipos moleculares de LM de partes moles em 51 casos. A análise de expressão gênica utilizando agrupamento não supervisionado mostrou três grupos de LM. O subgrupo mais evidente foi caracterizado por genes envolvidos principalmente com a função muscular, além de genes codificadores de fosfoproteínas e proteínas quinases (incluindo *CALD1*, *ACTG2*, *CFL2* e *MYLK*). Análise de aCGH mostrou um aumento significativo na variação do número de cópias de DNA neste subgrupo, principalmente perdas em 16q24 e 1p36. Adicionalmente, ensaios de imunoistoquímica identificaram cinco marcadores que, por sua vez, estavam associados com uma melhora da sobrevida dos pacientes.

LEE et al. (2004) identificaram 355 genes diferencialmente expressos entre 20 tumores primários e 7 sarcomas metastáticos, cujo perfil de expressão indicava o desenvolvimento de metástases em LM primários. A maioria dos genes associado ao perfil metastático codificava proteínas envolvidas no ciclo celular (*CDC27* e *CDK2AP1*), transdução de sinal (*IFNAR2* e *PAK2*) e apoptose (*BCL2A1*). Genes

associados com macrófagos (*CD163* e *CD68*), com a citocina CSF1 (induz a proliferação e diferenciação de macrófagos e monócitos) e com as proteínas de resposta a este fator (*CD163*, *FCGR3a* e *CTSL-1*) foram correlacionados a um pior prognóstico em LM ginecológicos e não ginecológicos. (LEE et al. 2008; ESPINOSA et al. 2009). NIELSEN et al. (2002) classificaram os LM em dois grupos. O primeiro subgrupo foi composto de seis amostras com expressão diferenciada em 24 genes, relacionados principalmente com a estrutura muscular (*ACTG2*, *MYH11*, *MYLK* e *CNN*), enquanto o segundo subgrupo (cinco casos) foi agrupado com amostras de SP e LP e apresentava genes envolvidos principalmente na síntese de citocinas (*MIG*, *SCYA5*), metabolismo (*LOXL1*, *MME*) e com características histiocíticas (*CD11B*, *CD68*).

O estudo de NAGAYAMA et al. (2002) comparou os perfis de expressão gênica de 13 sarcomas sinoviais e 34 sarcomas de células alongadas. A análise de agrupamento não-supervisionado dos genes diferencialmente expressos não foi capaz de separar os SP, LM e LP em classes distintas. Por outro lado, LEE et al. (2003) identificaram dois grupos diferentes de LM. O primeiro subgrupo foi agrupado com os SP, enquanto o segundo apresentou padrões de expressão distintos. Também identificaram um novo subgrupo de SP, caracterizado por 21 genes diferencialmente expressos, relacionados principalmente com metabolismo (*CTPS*, *SYK* e *ACAT2*) e resposta ao estresse oxidativo (*HSPE1* e *HIF1A*). SEGAL et al. (2003) analisaram 51 SPM utilizando uma plataforma de *oligoarray* (Affymetrix) e mostraram uma grande heterogeneidade entre os tumores pleomórficos, ao mesmo tempo que conseguiram identificar um grupo de SP como uma entidade distinta. Em outro estudo, após análise de agrupamento hierárquico dos padrões de expressão gênica, algumas

amostras de SP foram agrupadas com LP e schwannomas, enquanto classes distintas foram observadas contendo apenas amostras de SP. Os autores concluíram que os perfis de expressão gênica podem ser úteis na classificação deste conjunto raro de tumores (SKUBITZ e SKUBITZ 2004).

BAIRD et al. (2005) encontraram 2766 sondas diferencialmente expressas entre 16 classes de SPM analisados sendo identificados subgrupos de LM, LP e SP. HENDERSON et al. (2005) avaliaram o padrão de expressão gênica em 96 neoplasias mesenquimais, representativas de 19 tipos tumorais e mostraram a grande variabilidade presente em osteossarcomas, SP e LM. Não encontraram qualquer marcador específico que diferenciasse os LM dos SP. NAKAYAMA et al. (2007) utilizaram uma plataforma de *oligoarray* contendo 22283 sequências para avaliar o perfil de expressão gênica em 105 amostras de 10 tipos de SPM. Os resultados obtidos confirmaram a heterogeneidade do perfil de expressão gênica dos SP e concluíram que a maioria destas neoplasias podem ser reclassificados em subtipos pleomórficos de outros sarcomas. FRANCIS et al. (2007) avaliaram o perfil de expressão gênica de 177 SPM pela técnica de cDNA array e encontraram que os sarcomas de natureza pleomórfica, principalmente SP e LM, eram caracterizados por perfis heterogêneos, embora sugerissem que os genes relacionados com a hipóxia e envolvidos no ciclo celular, proliferação, homeostase e outras funções, parecem ter um papel importante nos subtipos pleomórficos. DETWILLER et al. (2005) demonstraram que o padrão de expressão gênica de 107 genes, relacionados com hipóxia, permitia distinguir SPM de tecidos normais. Neste estudo os SP e LM ficaram agrupados dentro de um único cluster.

Análise de clusterização realizadas por SKUBITZ et al. (2008) em SPM, utilizando três conjuntos de genes (*OVCA*, *RCC* e *AF*), sugeriu a existência de subgrupos de SP com significado biológico. Adicionalmente, alguns genes que são alvos terapêuticos foram encontrados diferentemente expressos nos SPM. CARNEIRO et al. (2009) utilizaram plataformas de BAC arrays e cDNA-array para identificar assinaturas moleculares específicas e marcadores prognósticos em 31 SP e 18 LM. Os perfis de expressão gênica, assim como a variação no número de cópias de DNA se revelaram muito similares nos dois tipos tumorais. A análise de expressão gênica mostrou nove genes diferencialmente expressos, entre eles o *TPM2*, porém, eles não permitiram distinções entre os subtipos tumorais. Adicionalmente, perdas na região 4q31 (*SMAD1*) e 18q22, além da presença de necrose tumoral, foram identificados como preditores independentes de metástases.

PÉROT et al. (2009) utilizaram a análise integrada de dados genômicos (aCGH e FISH) e transcriptômicos para examinar 19 SP e 17 LM bem diferenciados. As análises moleculares mostraram que o gene da miocardina (*MYOCD*) se encontra altamente amplificado e expresso em LM retroperitoniais. A proteína MYOCD é um cofator transcricional de SRF (fator de resposta sérica) que regula a diferenciação de músculo liso e pode estar envolvido nas altas taxas de metástase em LM, uma vez que está fortemente associada com a habilidade de migração das células (WANG Z et al. 2003; PETIT et al. 2008; PÉROT et al. 2009). Em estudo mais recente, KONSTANTINOPOULOS et al. (2010) analisaram cinco conjuntos de dados independentes de *arrays* de expressão (325 tumores) e por meio de ferramentas de bioinformática desenvolveram um preditor (170 genes), que foi utilizado para reclassificar SP e sarcomas sem outra especificidade. A maioria destas neoplasias

foram reclassificadas como LP (46%), fibrossarcomas (29%) e LM (14%). Esta classificação revelou linhas de diferenciação tecidual não detectadas anteriormente (adipócitos, fibroblastos, músculo liso) e foi validada em espécimens de parafina. Os SP reclassificados compartilhavam padrões similares de ativação de vias oncogênicas com seus subtipos correspondentes.

Estudos de expressão gênica também podem ser ferramentas importantes no diagnóstico diferencial de GIST e LM. Previamente, os GIST foram agrupados com SPM de genótipo complexo, principalmente LM (PRICE et al. 2007). Contudo, em anos recentes, o GIST emergiu como uma classe distinta de tumor de partes moles, associado com mutações no *KIT* e *PDGFRA* (HEINRICH et al. 2003; TRENT et al. 2007). PRICE et al. (2007) analisaram o perfil de expressão gênica de 68 amostras tumorais (37 GIST e 31 LM) e identificaram dois genes, *OBSCN* e *C9orf65*, capazes de diferenciar corretamente os GIST e LM, inclusive melhor que o diagnóstico baseado em *KIT*. Estudos mais recentes integraram dados de aCGH e expressão gênica e demonstraram diferenças substanciais entre estes tipos tumorais. Em quatro regiões cromossômicas (1p, 14q, 15q e 22q), a expressão gênica e as alterações genômicas foram significativamente diferentes entre os GIST e LM. Recentemente, análises de escala multidimensional mostraram que o perfil de expressão gênica destas quatro regiões distinguia os GIST dos LM com muita precisão (YANG et al. 2010; YLIPAA et al. 2010).

Novas informações moleculares podem levar à reclassificação de muitos tipos tumorais. Neoplasias previamente classificadas como SP mostraram ter padrões de expressão gênica similares a outros tumores, principalmente leiomiossarcomas e lipossarcomas (COINDRE et al. 2004; CARNEIRO et al. 2009; GIBAULT et al.

2011). Análises de expressão gênica em larga escala são cruciais para determinar estes padrões moleculares. Estudos em amostras coletadas antes de qualquer tipo de tratamento (radioterápico ou quimioterápico) fornecerão informações que permitam correlacionar resultados com padrões moleculares (KENNEY et al. 2009; WEST 2010). A análise genética provou ser uma ferramenta valiosa no diagnóstico e na gestão clínica dos SPM. No futuro, as decisões clínicas serão cada vez mais baseadas em uma combinação de critérios histológicos e na identificação molecular de anormalidades genéticas (JAIN et al. 2010). Na última década, muitos tipos de sarcomas foram estudados por técnicas baseadas em *microarrays*. Contudo, estes resultados ainda estão sendo “traduzidos” para permitir um melhor entendimento da sarcomatogênese e assim facilitar a identificação de marcadores prognósticos e diagnósticos, além de esquemas terapêuticos mais efetivos (ANTONESCU 2008; BECK et al. 2010b).

2 JUSTIFICATIVA

Os sarcomas pleomórficos são neoplasias agressivas que apresentam um comportamento clínico variável. Atualmente são caracterizados por um padrão morfológico comum a outros tipos de sarcomas, sendo o diagnóstico diferencial realizado por exclusão. A heterogeneidade destes tumores dificulta seu diagnóstico. Além disso, os dados citogenéticos e moleculares têm se mostrado inconsistentes, limitando sua utilização na identificação de marcadores genéticos. O diagnóstico preciso dos SP tem valor prognóstico e terapêutico, pois os tumores divergem tanto no seu comportamento clínico como na resposta a terapia. No presente estudo foi avaliado o perfil de expressão gênica em um grupo de sarcomas (sarcomas pleomórficos e leiomiossarcomas) que apresenta características histológicas semelhantes, inclusive por marcadores imunoistoquímicos, pela metodologia de *oligorrays*, com o intuito de caracterizar de forma mais acurada estes tumores, tanto do ponto de vista molecular quanto biológico. Estas abordagens poderão contribuir para um diagnóstico mais efetivo e, conseqüentemente, para o delineamento de tratamentos mais específicos para cada paciente, aumentando suas chances de sobrevida.

3 OBJETIVOS

- Comparar o padrão de expressão gênica entre sarcomas pleomórficos e leiomiossarcomas para avaliar se há diferenças entre eles;
- Utilizar os dados de expressão para investigar a associação entre os transcritos alterados em redes e vias genéticas e moleculares;
- Selecionar e investigar o padrão de expressão de transcritos específicos para confirmar os achados da análise de expressão gênica global;
- Comparar e correlacionar os resultados da pesquisa com os dados clínico-patológicos.

4 MATERIAL E MÉTODOS

4.1 CARACTERIZAÇÃO DA AMOSTRA

A casuística inicialmente coletada foi de 111 casos, porém, após revisão histológica e dos prontuários verificou-se que 67 amostras eram provenientes de pacientes que receberam tratamento prévio ou em que não foi possível a obtenção de RNA de qualidade ou quantidade suficiente para o estudo. A casuística final foi composta de 44 amostras de tecido congelado, pertencentes a 43 pacientes, e composta de 22 sarcomas pleomórficos e 22 leiomiossarcomas. Quatro amostras de tecidos adjacentes normais também foram incluídas. Estes tecidos foram obtidos durante o procedimento cirúrgico de três casos de SP (SP2T, SP3T e SP6T) e um de LM (LM4T) , os quais fazem parte deste estudo. Todos os espécimens utilizados neste estudo fazem parte do Banco de Tumores do Hospital AC Camargo (39), SP e do Hospital de Câncer de Barretos (5), SP e foram coletadas entre 2000 e 2010. Este estudo foi aprovado pelo Comitê de Ética e Pesquisa de ambos os Hospitais (CEP HCACC 1240/09). As amostras tumorais foram provenientes de biópsias (6 casos), ressecção do tumor primário (26 casos) e cirurgia de recidiva (12 casos), cuja lesão primária foi operada em outro serviço. O Quadro 2 sumariza as informações clínico-patológicas de cada paciente. Estes pacientes não receberam tratamento quimioterápico ou radioterápico prévios à obtenção da amostra. Dois pacientes foram caracterizados como portadores da Síndrome de Li-Fraumeni (LM2T e LM22T). A idade média dos pacientes foi de 59,1 anos (variando de 4 a 90 anos); sendo 25 do

sexo masculino e 19 feminino. A maioria dos tumores localiza-se no retroperitônio (19 casos) e nas extremidades inferiores (16 casos). Em relação ao grau histológico, 40 casos foram classificados como sendo de alto grau (G2 e G3) e quatro casos de LM, foram classificados como de baixo grau (G1). Setenta por cento dos tumores foram profundos e maiores que 5 cm (T2b). O tempo médio de seguimento clínico foi de 29,5 meses (variando de 1 a 180 meses) (Quadro 2).

Quadro 2 - Dados clínicos e patológicos dos pacientes com sarcomas pleomórficos (SP) e leiomiossarcomas (LM).

Caso	Idade	Sexo	Origem do material coletado	Topologia	TNM	Grau Histológico	Segundo Tu Primário	Recidiva e Metástase	I. Perineural	I. Angiolinfática	Follow Up (meses)
SP1T ^c	51	F	Recidiva (4 meses após o Tu primário)	Extremidade Inferior	T1aN0M0	III	—	—	—	—	VSD (105)
SP2T ^e	52	F	Recidiva (10 meses após o Tu primário)	Retroperitônio	T2bN0M0	III	—	Local	—	—	OPD (24)
SP3T	49	M	Recidiva (2 meses após o Tu primário)	Retroperitônio	T2bN0M0	III	—	Local, Linfonodo	Sim	—	OPD (20)
SP4T	90	M	Tumor Primário	Extremidade Inferior	T2aN0M0	III	—	Linfonodo	—	—	OPD (10)
SP5T	50	M	Recidiva (1 mês após o Tu primário)	Retroperitônio	T2bN0M0	III	—	Local	—	—	OPD (7)
SP6T ^d	56	M	Biópsia	Extremidade Inferior	T2bN0M0	III	—	—	—	—	VSD (42)
SP7T ^e	58	F	Recidiva (6 meses após o Tu primário)	Cabeça e Pescoço	T2aN0M0	III	—	—	Sim	—	PS (29)
SP8T	63	M	Tumor primário	Cabeça e Pescoço	T1aN0M0	III	—	—	—	—	OPD (15)
SP9T	72	M	Biópsia	Extremidade Inferior	T2bN0M0	III	—	—	—	—	VSD (12)
SP11T ^d	70	F	Tumor Primário	Retroperitônio	T2bN0M0	III	—	—	—	—	OPD (2)
SP12T ^d	57	M	Tumor Primário	Retroperitônio	T2bN0M0	III	—	Região paravertebral, Pulmão?	—	—	OPD (5)
SP13T ^c	63	M	Recidiva (16 meses	Extremidade	T2bN0M0	III	—	Local,	—	—	OPD (18)

SP14T	32	F	após o Tu primário) Tumor Primário	Inferior Tronco	T1aN0M0	III	_____	Pulmão Pulmão?	_____	_____	VSD (30)
SP15T	80	M	Recidiva (9 meses após o Tu primário)	Retroperitônio	T2bN0M0	III	Ca próstata, Adenoma adrenal	_____	_____	_____	PS (1)
SP16T ^c	60	M	Biópsia	Tronco	T2aN0M0	III	_____	_____	_____	_____	VSD (1)
SP17T	56	M	Biópsia	Extremidade Inferior	T2bN0M0	III	_____	_____	_____	_____	VCD (4)
SP18T^e	63	M	Tumor Primário	Extremidade Inferior	T2bN0M1	III	SP de couro cabeludo?	Pulmão, Adrenal	Sim	_____	OPD (15)
SP19T ^e	77	M	Tumor Primário	Extremidade Superior	T2bN0M0	III	_____	_____	_____	_____	VSD (32)
SP-01	41	M	Tumor Primário	Retroperitônio	T2bN0M0	III	_____	_____	_____	_____	VSD (6)
SP-03	82	F	Tumor Primário	Extremidade Inferior	T2bN0M0	III	_____	_____	_____	_____	OPD (3)
SP-04	78	M	Tumor Primário	Extremidade Superior	T2bN0M0	III	Adenoma tubular de rim	Local, Pulmão	_____	_____	VCD (11)
SP-05	60	M	Recidiva (38 meses após o Tu primário)	Extremidade Inferior	T2aN0M0	III	_____	Pulmão, Ossos	_____	_____	VCD (5)
LM2T ^b	50	F	Recidiva (144 meses após o Tu primário)	Retroperitônio	T1bN0M0	III	CEC Couro Cabeludo, Sarcoma de hipocôndrio, Lipoma de dorso	_____	_____	_____	VSD (180)
LM3T	60	F	Recidiva (22 meses após o Tu primário)	Extremidade Superior	T1aN0M0	II	_____	_____	_____	_____	VSD (28)
LM4T	89	F	Recidiva (6 meses após o Tu primário)	Retroperitônio	T2bN0M0	II	_____	Pulmão	_____	_____	OPD (12)

LM5T ^a	77	F	Tumor Primário	Tronco	T2bN0M0	III	—	—	—	—	OPD (51)
LM6T ^f	54	F	Tumor Primário	Retroperitônio	T2bN0M1	III	—	Fígado	—	—	OPD (14)
LM7T ^f	37	M	Biópsia	Extremidade Inferior	T1aN0M0	III	—	—	—	—	VSD (27)
LM8T	61	F	Tumor Primário	Retroperitônio	T2bN0M0	III	—	Fígado?	—	—	OPD (20)
LM9T	62	M	Tumor Primário	Extremidade Inferior	T2bN0M0	II	Adenocarcinoma de Próstata	—	—	—	VSD (99)
LM10T	65	M	Tumor Primário	Extremidade Inferior	T1aN0M0	III	—	—	—	—	PS (74)
LM11T	55	F	Tumor Primário	Retroperitônio	T2bN0M0	I	—	Ossos, Pulmão, Fígado, Coxa	—	—	OPD (43)
LM13T	45	F	Tumor Primário	Retroperitônio	T2bN0M0	I	Miomas uterinos	—	—	—	VSD (121)
LM14T	55	M	Tumor Primário	Extremidade Inferior	T1bN0M0	II	—	—	—	—	PS (1)
LM15T	49	M	Tumor Primário	Extremidade Inferior	T2bN0M0	I	—	Pulmão?	—	—	PS (22)
LM16T ^f	50	M	Recidiva (85 meses após o Tu primário)	Retroperitônio	T2bN0M0	III	—	—	—	—	VSD (93)
LM17T	48	M	Tumor Primário	Retroperitônio	T2bN0M0	III	—	—	—	—	PS (4)
LM18T	61	M	Tumor Primário	Retroperitônio	T2bN0M0	III	—	Nódulos peritoneais	—	—	VSD (16)
LM19T	4	M	Tumor Primário	Extremidade Inferior	T2bN0M0	III	—	—	—	—	VSD (7)

LM20T ^f	81	F	Biópsia	Extremidade inferior	T2bN0M1	III	—	Pulmão, Cavidade abdominal	—	—	OPD (10)
LM21T	52	F	Tumor Primário	Retroperitônio	T2bN0M0	I	Tu benigno de útero	—	—	—	VSD (9)
LM22T ^b	74	F	Tumor Primário	Retroperitônio	T2bN0M0	II	Ca Útero	—	—	—	PS (2)
LM23T	45	F	Tumor Primário	Retroperitônio	T2bN0M0	III	Leiomiomas uterinos	Ossos?	—	—	VCD (3)
LM-02	58	F	Tumor Primário	Extremidade Superior	T2aN0M0	III	Carcinoma ductal de mama	—	—	—	VCD (48)

Tecidos Adjacentes Normais

Caso	Origem do material coletado	Topologia	Avaliação histopatológica
SP2N	Tecido adjacente normal do caso SP2T	Retroperitônio	Tecido fibroso misturado com gordura
SP3N	Tecido adjacente normal do caso SP3T	Retroperitônio	Tecido muscular esquelético não neoplásico
SP6N	Tecido adjacente normal do caso SP6T	Extremidade Inferior	Tecido mesenquimal composto principalmente de gordura
LM4N	Tecido adjacente normal do caso LM4T	Tronco	Tecido muscular esquelético misturado com tecido fibroso

Legendas – VCD: Vivo com Doença, VSD: Vivo sem Doença, OPD: Óbito pela Doença, PS: Perda de Seguimento.

Em negrito: Tumores pertencentes ao mesmo paciente. Há dúvidas se o caso SP8T é um segundo tumor primário ou é uma metástase do caso SP18T.

^a Paciente com história de LM de reto, operado em outro serviço. Há dúvidas se o caso LM5T é um tumor primário de retroperitônio ou é uma recidiva de LM de reto.

^b Pacientes com a Síndrome de Li-Fraumeni.

^c Casos de SP com áreas mixóides focais.

^d Caso de SP com morfologia que sugere lipossarcoma.

^e Casos de SP com diagnóstico mais complicado.

^f Casos de LM com diagnóstico mais complicado.

4.2 CASUÍSTICA

Foram selecionados para a análise de expressão gênica global 22 leiomiossarcomas e 22 sacomas pleomórficos. Todos os casos foram reavaliados pela patologista Dra. Isabela Werneck da Cunha Werneck, Departamento de Patologia, Hospital AC Camargo, SP. O critério diagnóstico foi baseado nas recomendações da Organização Mundial da Saúde e levou em consideração a morfologia e os resultados da análise da expressão de proteínas por imunoistoquímica (FLETCHER et al. 2002). O grau histológico foi definido segundo as recomendações da *Federation Nationale des Centres de Lutte Contre le Cancer* (FNCLCC), que utiliza três variáveis: índice mitótico, necrose tumoral e diferenciação celular (GUILLOU et al. 1997). Exemplos representativos dos aspectos morfológicos de alguns casos que compõem este estudo estão apresentados na Figura 1.

Todos os casos de sarcomas de alto grau histológico com morfologia fusocelular e pleomórfica, negativos para os imunomarcadores musculares (como AML, HHF35 e desmina), vasculares (CD34, CD31), epiteliais (EMA e citoceratinas) e de diferenciação neural (proteína S100) foram classificados como Sarcomas Pleomórficos (SP2T, SP4T, SP9T, SP15T, SP-01, SP-03, SP-04 e SP-05). Os sarcomas com morfologia totalmente pleomórfica e expressão apenas focal de marcadores musculares também foram considerados SP (SP7T, SP11T, SP14T, SP17T, SP18T, SP19T). Nos casos SP3T, SP5T e SP8T não foram realizados exames de imunoistoquímica, porém, a morfologia completamente pleomórfica permitiu sua classificação como SP.

Os casos de leiomiossarcomas apresentaram morfologia variando da clássica forma de feixes entrecortados até uma morfologia mais pleomórfica, mas mantendo expressão pelo menos moderada de marcadores musculares (AML, HHF35 ou desmina). Casos com morfologia pleomórfica, mas com áreas de feixes associados a expressão focal de marcadores musculares (AML, HHF35 ou desmina) foram considerados como LM (LM6T, LM7T, LM16T e LM20T). Alguns sarcomas pleomórficos (três casos), apresentaram áreas mixóides focais, podendo eventualmente corresponder a mixofibrossarcomas de alto grau histológico (SP1T, SP13T e SP16T).

Dois casos de SP, localizados no retroperitônio, podem também corresponder a lipossarcomas desdiferenciados (SP11T e SP12T) e um caso, localizado na extremidade inferior, pode ser classificado como um lipossarcoma mixóide (SP6T), entretanto, como não foram observados componentes de lipossarcomas bem diferenciados justapostos nestes casos, após exaustiva análise histopatológica, eles foram classificados como SP.

Para a análise de transcritos específicos, foram selecionadas 30 amostras dependendo da disponibilidade da amostra e resultados da análise de expressão gênica global, incluindo 14 SP (SP1T, SP2T, SP3T, SP4T, SP5T, SP6T, SP8T, SP12T, SP14T, SP15T, SP19T, SP-01, SP-03 e SP-04), 12 LM (LM2T, LM4T, LM5T, LM6T, LM8T, LM11T, LM13T, LM15T, LM17T, LM22T, LM23T e LM-02) e 4 tecidos adjacentes normais (SP2N, SP3N, SP6N e LM4N).

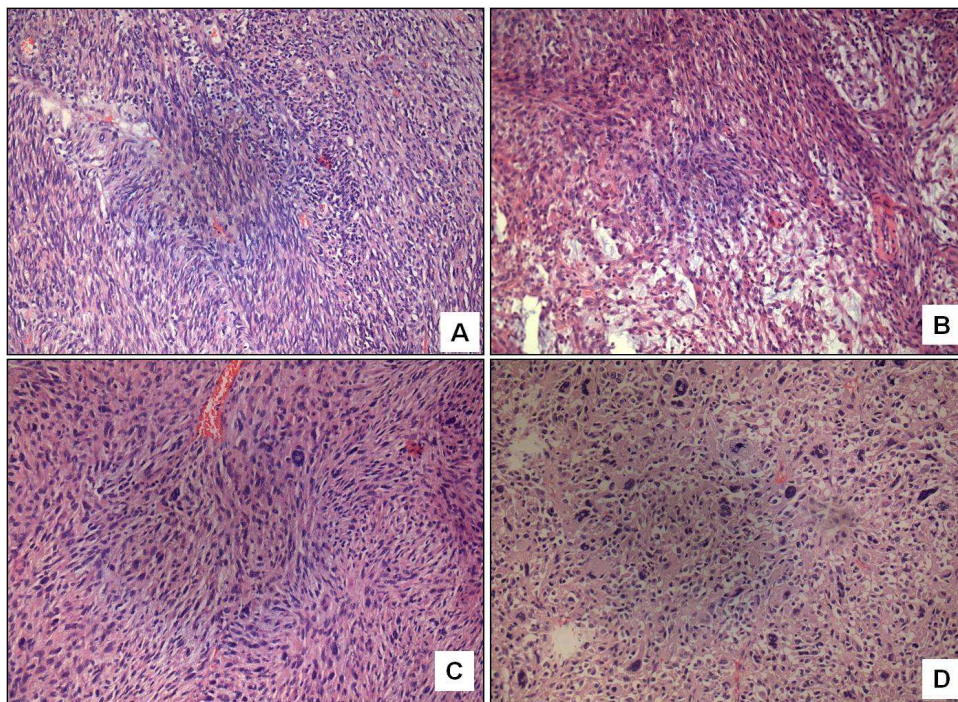


Figura 1 - Aspecto morfológico de quatro casos de sarcomas pleomórficos e leiomiossarcomas. Coloração HE, objetiva de 40x. **(A)** Leiomiossarcoma. Observa-se grande número de feixes de fibras musculares lisas, em lençóis, caracterizando um leiomiossarcoma bem diferenciado. **(B)** Sarcoma pleomórfico com áreas mixóides, com células de aspecto gorduroso. **(C)** Leiomiossarcoma com áreas pleomórficas. Presença de células musculares lisas, de moderado grau de diferenciação, em arranjo de ninho. Há anisocariose, megacariose com hiper cromasia. **(D)** Sarcoma pleomórfico, composto por células com anisocitose, anisocariose, megacariose, algumas com núcleo de cromatina condensada. O citoplasma é amplo e as vezes vacuolado.

4.3 EXTRAÇÃO DE RNA

A extração do RNA foi realizada no Laboratório de Biomoléculas, Hospital AC Camargo, SP sob a responsabilidade da Dra Dirce Carraro. Foram utilizados os protocolos estabelecidos pelo *Trizol®reagent* (Invitrogen) e *Rneasy Mini Kit* (Qiagen). A qualidade e a pureza do RNA foram avaliadas utilizando-se o espectrofotômetro NanoDrop (ND-1000 Spectrophotometer v.3.0.1, Labtrade) e o Bioanalyzer (RNA 6000 NanoLabChip kit 2100 Agilent Technologies),

respectivamente. A concentração ideal obtida para a realização dos experimentos é de 0,8 µg/uL.

4.4 EXPRESSÃO GÊNICA GLOBAL (OLIGOARRAYS)

O perfil de expressão gênica foi avaliado utilizando-se a plataforma *Whole Human Genome 4x44K* (Agilent). Os procedimentos de marcação das amostras, hibridação e detecção seguiram o protocolo *Two-color microarray-based gene expression analysis* (Agilent Technologies) que utiliza o *Quick Amp Labeling Kit* (Agilent Technologies) recomendado pelo fornecedor. Para a pré-síntese de cDNA foram obtidos 0,8 µg de RNA total das amostras teste (sarcomas e tecidos adjacentes normais) e RNA de Referência Universal, que foram marcadas com cianina 3 (Cy-3) e cianina 5 (Cy-5), respectivamente. O RNA de referência universal utilizado contém uma mistura de quantidades iguais de RNA total obtido de 15 linhagens celulares diferentes (Quadro 3), já utilizado em outros estudos (GOMES et al., 2005; STOLF et al., 2006; BOCCARDO et al., 2010). A quantidade relativa de cada mRNA na amostra foi calculada como a razão entre a intensidade na amostra sobre a intensidade na referência; e as amostras foram comparadas entre si pela razão das razões (POLLACK, 2002).

Quadro 3 - Linhagens celulares utilizadas na composição do RNA de referência universal.

Linhagens celulares	Origem
Daudi	Linfoma de Burkitt
DLD-1	Adenocarcinoma de Cólon
DU 145	Carcinoma de Próstata
FaDu	Carcinoma Epidermóide de Faringe
GM 637	Fibroblasto
H 146	Carcinoma de Pulmão
H 1080	Fibrossarcoma
HB4α	Célula Luminal de Mama
HEK 293	Rim embrionário Humano
Jurkat	Leucemia Aguda de Células T
Saos-2	Osteossarcoma
SK-BR-3	Adenocarcinoma de Mama
SK-MEL-28	Melanoma
T24	Carcinoma de Bexiga
T98G	Glioblastoma

Foram adicionados 2 μ L de RNA *spike A* às amostras teste e 2 μ L de RNA *spike B* às amostras referência, além de 1,2 μ L de T7 *Promoter Primer*, o qual se liga à cauda poli-A do mRNA. Essa reação foi incubada no termociclador PTC-100 (Peltier - Effect Cycling - MJ Research, Inc., USA) por 10 min a 65 °C e colocada em gelo por 5 min. A síntese de cDNA consistiu na adição de 4 μ L de 5X *First Strand Buffer* (250 mM Tris-HCl, pH 8,3, 375 mM KCl, 15 mM MgCl₂), 2 μ L de DTT (0.1M), 1 μ L de dNTP mix (10mM), 0,5 μ L da enzima *RNaseOut*, que promove a inativação das RNases e 1 μ L da enzima MMLV-RT, responsável pelo início da transcrição reversa. As reações foram incubadas em um termociclador por 2 horas a 40°C, para realizar a transcrição reversa, e a 65°C por 15 min, para inativar a ação da enzima, e foram transferidas para o gelo. As amostras de cDNA foram transcritas para cRNA em um volume final de 80 μ L por reação contendo 15,3 μ L de

água estéril, 20 μ L de 4X *Transcription Buffer*, 6,4 μ L de 50% PEG, 6 μ L de DDT (0.1M), 8 μ L de NTP mix, 0,5 μ L de *RNaseOUT*, 0,6 μ L de *Inorganic Pyrophosphatase*, 2,4 μ L de Cy-3 ou Cy-5 e 0,8 μ L da enzima T7 RNA Polimerase, que incorpora as cianinas à fita de cRNA. As reações permaneceram em termociclador a 40°C por 2 horas. A purificação das sondas foi realizada utilizando o *RNeasy Mini Kit* (Quiagen). Após as etapas de purificação na coluna, o cRNA purificado foi coletado em 30 μ L de água estéril. Os procedimentos de hibridação e lavagem seguiram as recomendações do fabricante. A captura das imagens foi obtida após o uso do *DNA microarray Scanner with SureScan High Resolution Technology* (Agilent Technologies) e o *Scan Control Software* 8.1. A extração dos dados foi possível usando o *Software Feature Extraction v10.1.1.1* (Agilent Technologies). Este programa além de extrair os valores de intensidade a partir das imagens geradas, também possibilita a aplicação de testes de qualidade das hibridações utilizando os RNAs *spike* adicionados durante a marcação. Este procedimento também faz as correções necessárias dos valores de intensidade provenientes de cada corante (usando o algoritmo Lowess), para eliminar as diferenças de eficiência da emissão dos corantes fluorescentes Cy3 e Cy5; assim como a subtração do *background* local (valor de intensidade ao redor de cada *spot*).

4.5 RT-PCR QUANTITATIVA EM TEMPO REAL (RT-qPCR)

4.5.1 Obtenção de cDNA

Previamente à reação de transcrição reversa para obtenção do DNA complementar (cDNA), 2 μ g de cada amostra foi diluída com H₂O DEPC a uma

concentração final de 0,125 ug/ μ L e a seguir tratadas com *DNAseI Amplification Grade* (Invitrogen - 1 U/ μ L). A Síntese de cDNA foi realizada em um volume final de 32 μ L contendo 5x *First Strand Buffer* (250 mM Tris-HCl, pH 8,3; 375 mM KCl; 15 mM MgCl₂), 10 mM de cada dNTP, 0,5 ug/ μ L de Oligo(dT)₁₈, 0,1M ditioneitol (DTT), 100 ng/ μ L de *random primer* e 200U da transcriptase reversa *SuperScriptTM III* (Invitrogen Life Technologies, Inc., Carlsbad, CA). A transcrição reversa foi realizada a 42°C por 60 min e a reação foi subsequentemente inativada por 15 min a 70°C. O cDNA foi estocado a -20°C.

4.5.2 RT-qPCR em tempo real

O desenho dos iniciadores para os genes alvo e endógenos foi realizado utilizando o *Software Primer Express 2.0* (*Applied Biosystems*). Primeiramente foi realizada uma análise *in silico* baseada em pesquisa de bancos de dados públicos, como o *National Center for Biotechnology Information* (<http://ncbi.nlm.nih.gov>) e o *Ensembl*, para verificar se estes genes se expressavam em sarcomas e para avaliar a presença de possíveis isoformas. As sequências *forward* e *reverse* dos iniciadores foram avaliadas pela ferramenta BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>), garantindo a especificidade dos iniciadores para cada gene a ser analisado pela RT-qPCR. Baseados em estudos prévios do grupo de sarcomas do Hospital AC Camargo, três genes endógenos foram testados para a normalização da expressão dos genes alvo, os quais foram selecionados segundo os resultados da análise de expressão gênica global (*cluster* supervisionado e análise de redes gênicas). Para este fim se utilizou o *software* geNorm (<http://medgen.ugent.be/~jvdesomp/genorm/>),

desenvolvido por VANDESOMPELE et al. (2002). As sequências dos iniciadores dos genes alvo e de referência estão apresentadas no Quadro 4.

Quadro 4 - Sequências de bases dos iniciadores utilizados nos experimentos de expressão dos transcritos alvo *SRF*, *SRC*, *BAD* e *TOP2A* por RT-qPCR.

Gene	Iniciador <i>Forward</i>	Iniciador <i>Reverse</i>
<i>GAPDH</i> ^E	5'-CGTCTTCACCACCATGGAGA-3'	5'-CGGCCATCACGCCACAGTTT-3'
<i>HMBS</i> ^E	5'-GGCAATGCGGCTGCAA-3'	5'-GGGTACCCACGCGAATCAC-3'
<i>HPRT</i> ^E	5'-TCATTATGCTGAGGATTTGGAAAG-3'	5'-GGCCTCCCATCTCCTTCATC-3'
<i>SRF</i> ^A	5'-CGCTACACGACCTTCAGCAAG-3'	5'-CATAGGCCTTCTTGATGCC-3'
<i>SRC</i> ^A	5'-CAAGGTGCCAAATTCAAAAT-3'	5'-GTCAGCAGGATCCCGAAGG-3'
<i>BAD</i> ^A	5'-CTTTAAGAAGGGACTTCCTCGCC-3'	5'-AAGACTCGCGTCCAGCTGG-3'
<i>TOP2A</i> ^A	5'-AGAAGACAGCAGCAAAAAGTCAGTC-3'	5'-ACCAGAATTCAAAGCTGGATCC-3'

^A Genes alvo

^E Genes endógenos

As reações de amplificação foram realizadas no Termociclador 7500 *Real Time PCR System* (*Applied Biosystems*). Foram utilizadas placas de 96 poços, contendo duplicatas de cada amostra e controle negativo, num volume final de 12,5 µL para cada reação sendo 5 µL do reagente SYBR Green Master Mix (*Applied Biosystems*); 0,2 µL de cada iniciador 10 µM (iniciadores *forward* e *reverse*) e 4 µL de cDNA diluído (concentração de 0,005 µg/µL). As condições de reação de amplificação para todos os iniciadores foram: 95°C por 10 min, seguido de 40 ciclos a 95°C por 15 seg e 60°C por 1 min. Ao final de cada reação, a curva de dissociação para cada gene testado foi realizada de modo a observar a especificidade da reação por meio da identificação de um único pico referente a um único produto amplificado. A quantificação relativa (QR) da expressão gênica é usada para descrever mudanças na expressão de um gene de interesse em um grupo de amostras

em relação a uma amostra referência. Foi utilizado um *pool* de 10 amostras de sarcomas como referência. Para a quantificação relativa dos transcritos de interesse foi utilizado o modelo proposto por PFAFFL (2001), conforme fórmula abaixo descrita.

$$QR = \frac{(E+1)alvo^{\Delta Ct\ alvo}}{(E+1)normalizador^{\Delta Ct\ normalizador}}$$

Onde:

E = Eficiência da reação de amplificação

Alvo = transcrito alvo avaliado

Normalizador = transcrito utilizado como normalizador

ΔCt = Diferença entre o Ct (Ciclo onde a fluorescência ultrapassou o limiar estabelecido, do inglês *Cycle threshold*) obtido para a amostra referência e o CT obtido para a amostra em análise

A eficiência de amplificação para os iniciadores utilizados foi igual a 100% (exatamente uma duplicação ao final de cada ciclo, $E=1$). Ao utilizar mais de um gene endógeno na normalização das reações, utilizou-se a média geométrica dos valores de expressão relativa (FN: fator de normalização gerado pelo programa geNorm) dos genes endógenos mais apropriados.

$$QR = \frac{2^{\Delta Ct\ alvo}}{FN}$$

4.6 ANÁLISES ESTATÍSTICAS

4.6.1 Dados de expressão gênica global

A análise da qualidade das hibridações foi realizada utilizando os relatórios *QC Reports* gerados pelo *Software Feature Extraction v10.1.1.1 (Agilent Technologies)*. Foi analisado um conjunto de 43.376 sondas de cDNA presentes nas matrizes 44K. Destas, foram utilizadas apenas 28.226 sondas que tiveram medidas válidas (ou seja, foram detectados acima do fundo de matriz média, conforme determinado por um conjunto de controles negativos representados nas matrizes Agilent).

O *software R 2.9.2 (R Development Core Team 2009)* foi utilizado para cálculo da média entre as sondas que cobrem um mesmo gene e remoção dos controles e genes que apresentaram um número de *Nas (not applied, na ausência de sinal)* igual a 0 (zero) para o grupo normal e menor ou igual a 10 para o grupo tumoral permitindo assim a formação de uma única matriz de dados em log na base 2. Após a remoção de *outliers* pouco reprodutíveis entre as réplicas, utilizou-se a ferramenta *Significance Analysis for Microarray (SAM)*. Esta análise permite a identificação de genes diferencialmente expressos com significância estatística entre conjuntos de experimentos de *microarrays*, por permutações repetidas dos casos de expressão de cada transcrito, entre as amostras de dois ou mais grupos que estão sendo comparados. Para cada gene, o valor real da correlação medida entre os grupos (valor observado) é comparada a um valor de correlação obtido pela permutação dos dados (valor esperado ao acaso) para checar o valor estatístico associado ao conjunto

de transcritos encontrados como diferencialmente expresso entre duas condições analisadas.

Se o valor de correlação observado excedia o valor esperado, a diferença de expressão era considerada significativa. Para isso foi utilizado um P-valor de 0,05. O programa também permite o cálculo da taxa de falsa descoberta (*False Discovery Rate* - FDR), que expressa a porcentagem de falsos-positivos dentro do grupo de transcritos considerado diferencialmente expresso. Para a seleção de conjunto de transcritos foi considerada uma taxa de falso positivo próximo a zero entre os dados presentes nos grupos tumorais.

Após a identificação de um conjunto de transcritos diferencialmente expressos, foram feitas análises de agrupamentos hierárquicos (*Hierarchical clustering Analysis*) utilizando o programa TMEV e como medida de similaridade foi utilizada a distância Euclidiana e 1000 permutações. Foram construídos mapas de expressão (*Heat Maps*) dos transcritos expressos diferencialmente utilizando-se o mesmo programa. Outro método utilizado para a obtenção de significância estatística das diferenças de expressão (valores de P) foi calculado por *bootstrap resampling*, ou seja, análise dos valores dos genes significantes expressos após 1000 permutações aleatórias das amostras (REIS et al. 2004).

Todos os genes com alteração de expressão gênica tiveram o *fold change* calculado e foram classificados em grupos utilizando as anotações das vias do *Kyoto Encyclopedia of Genes and Genomes*-KEGG e a base de dados *Gene Ontology* (GO). Genes com *fold change* > 0,7 foram submetidos à análise de redes, vias gênicas, doenças relacionadas e função celular ou molecular com a utilização do

software Ingenuity Pathway Analysis (IPA) (Ingenuity Systems). Os *scores* para cada item foram calculados pelo teste exato de *Fisher* e são iguais a $-\log(\text{valor de } P)$.

4.6.2 Dados de RT-qPCR em tempo real

As análises estatísticas foram realizadas utilizando o programa SPSS (*Statistics Packet for Social Sciences*), versão 17.0 (SPSS, Chicago, IL, USA). Na análise do nível dos transcritos, os valores de expressão relativa foram convertidos em logaritmo natural ($\ln$), no intuito de diminuir o impacto de valores extremos. Pelo fato dos grupos não apresentarem distribuição normal, foi utilizado o teste não-paramétrico de Mann-Whitney para determinar a significância estatística da diferença do nível de expressão entre as amostras de SP e LM. O teste não-paramétrico de Wilcoxon foi utilizado na comparação entre os quatro sarcomas e seus respectivos tecidos normais adjacentes, devido ao pequeno número de amostras ($n=8$).

5 RESULTADOS

No presente estudo foram analisados 44 sarcomas, provenientes de 43 pacientes incluindo 22 sarcomas pleomórficos (duas provenientes do mesmo paciente, SP8T e SP18T) e 22 leiomiossarcomas. O Quadro 5 apresenta os resultados da análise de expressão protéica por imunoistoquímica para os marcadores relacionados com a função muscular nos casos do presente estudo. Estão incluídos também quatro amostras de tecidos normais adjacentes (SPN ou LMN) provenientes de pacientes desta casuística (3 SP e 1 LM). A análise de expressão de transcritos foi realizada com sucesso em todos os casos utilizando a plataforma de oligoarrays 4x44K (Agilent). O teste estatístico *SAM* de uma e duas classes, com valor de $FDR=0$, foi utilizado para analisar e identificar os genes diferencialmente expressos nos dois grupos de tumores.

O primeiro *cluster* foi gerado a partir de teste não supervisionado entre os 44 casos (Figura 2). Por esta análise foram observados quatro agrupamentos e um único caso que se agrupou separadamente. Todos os agrupamentos apresentaram tanto amostras de SP como de LM, não sendo possível a diferenciação por entidade diagnóstica. O primeiro grupo foi composto por nove amostras: seis sarcomas pleomórficos e três leiomiossarcomas. Entre os três leiomiossarcomas, dois (LM13T e LM15T) são de baixo grau histológico de malignidade (GI). Entre as amostras de SP, apenas dois casos foram considerados SP típicos quanto à morfologia e padrão de expressão de proteínas por imunoistoquímica (SP-01 e SP-02) (Quadro 3). Dois casos classificados como SP (SP11T e SP12T) apresentaram morfologia sugestiva de

Quadro 5 - Resultados da análise da expressão proteica por imunohistoquímica para os marcadores relacionados com a função muscular nas amostras do presente estudo.

Casos	Imunohistoquímica
SP1T	Não testado
SP2T	AML - / Desmina - / HHF-35 -
SP3T	Não testado
SP4T	Desmina - / HHF-35 -
SP5T	Não testado
SP6T	AML -
SP7T	AML + focal / Desmina - / HHF-35 + focal
SP8T	Não testado
SP9T	AML -
SP11T	Desmina + focal
SP12T	Não testado
SP13T	AML - / Desmina - / HHF-35 - / Caldesmon -
SP14T	AML + focal
SP15T	AML - / Desmina -
SP16T	AML + focal / Desmina -
SP17T	AML + focal / Desmina + focal
SP18T	AML + focal / HHF-35 + focal
SP19T	AML + focal / Desmina + focal / HHF-35 -
SP-01	Desmina - / HHF-35 -
SP-03	Desmina - / HHF-35 - / Calponina + focal
SP-04	Desmina -
SP-05	Desmina - / HHF-35 -
LM2T	AML + / Desmina +
LM3T	Não testado
LM4T	AML + / HHF-35 +
LM5T	AML + / Desmina - / HHF-35 -
LM6T	AML + / Desmina -
LM7T	AML + / Desmina +
LM8T	Desmina - / HHF-35 +
LM9T	AML + focal / Desmina + / HHF-35 + focal
LM10T	AML + / HHF-35 -
LM11T	Não testado
LM13T	AML + / Desmina + / HHF-35 +
LM14T	AML + / Desmina + / HHF-35 +
LM15T	AML + / Desmina + / HHF-35 + / Calponina +
LM16T	AML + difuso/ Desmina + difuso
LM17T	AML + / Desmina + / HHF-35 +
LM18T	AML + / Desmina +
LM19T	AML + / Desmina - / HHF-35 +
LM20T	AML + / Desmina + focal
LM21T	AML + / Desmina +
LM22T	AML + / Desmina +
LM23T	AML + / Desmina + / Caldesmon +
LM-02	Desmina + / HHF-35 + focal

+: imunomarcagem positiva, -: imunomarcagem negativa. Análise realizada sob a responsabilidade da Dra Isabela Werneck da Cunha, Departamento de Patologia, Hospital AC Camargo, SP.

lipossarcoma desdiferenciado, enquanto o caso SP13T não pode ser excluído quanto ao diagnóstico de mixofibrossarcoma. Por outro lado, o caso SP7T foi considerado como SP devido a morfologia pleomórfica, apesar de apresentar padrão de expressão focal das proteínas AML e HHF-35.

O segundo grupo foi composto por 10 casos de LM e três de SP. Entre os casos de LM, dois (LM6T e LM16T) apresentaram morfologia e padrão de expressão protéica por imunoistoquímica não característicos. O caso LM6T apresentou positividade para AML e negatividade para desmina, enquanto o caso LM16T apresentou positividade apenas difusa tanto para AML como para desmina. Entre as três amostras de SP, o caso SP19T apresentou morfologia pleomórfica, porém, com expressão protéica focal para AML e desmina. Em relação ao caso SP8T, não está claro se este representa um segundo tumor primário ou se é uma metástase do caso SP18T.

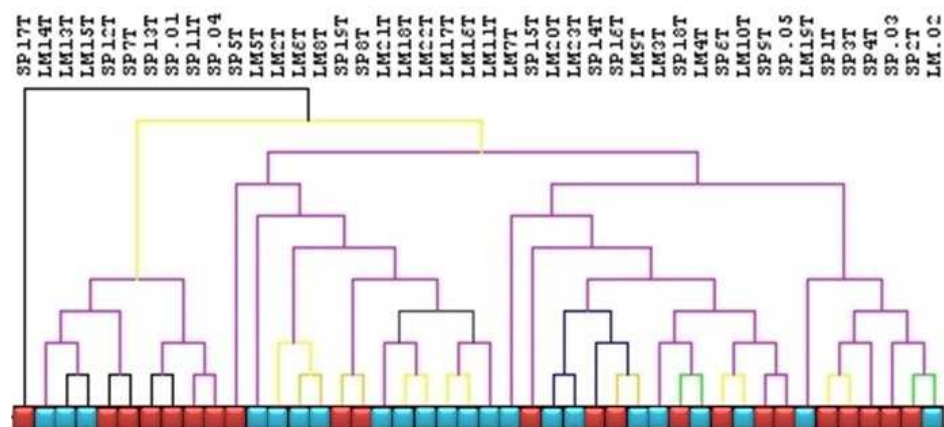


Figura 2 - Agrupamento entre os 44 casos de sarcomas por análise não supervisionada. Os SP estão representados em vermelho e os LM em azul.

O terceiro grupo apresentou número de casos de LM (LM3T, LM4T, LM7T, LM9T, LM10T, LM20T e LM23T) igual ao de SP (SP6T, SP9T, SP14T, SP15T, SP16T, SP18T e SP-05). Dois casos (LM7T e LM20T), foram considerados como

LM quanto a sua morfologia e padrão de expressão protéica. Por outro lado, um caso de SP pode representar um mixofibrossarcoma (SP16T) e outro um lipossarcoma mixóide (SP6T). O caso SP18T foi considerado como SP por apresentar um padrão morfológico pleomórfico, apesar da positividade, em ensaios de imunoistoquímica, para AML e HHF35 em raras células neoplásicas. Não foi descartada a possibilidade deste caso (SP18T) ser um tumor primário da amostra SP8T (que também poderia ser uma metástase) a qual se encontra no segundo agrupamento.

O quarto grupo foi composto por cinco SP (SP1T, SP2T, SP3T, SP4T e SP-03) e dois LM (LM-02 e LM19T). O caso SP1T pode corresponder a um mixofibrossarcoma com células gigantes. O caso SP2T foi o sarcoma mais difícil diagnosticar e foi considerado como SP pela negatividade de expressão das proteínas desmina, AML e HHF-35, apesar do padrão morfológico não ser característico quanto ao pleomorfismo. O caso SP17T foi o único que se apresentou isolado no *cluster*. Esta amostra foi caracterizada como SP devido à morfologia pleomórfica, apesar da positividade focal para as proteínas desmina e AML.

A análise supervisionada dos resultados da expressão dos transcritos dos 44 casos também não foi capaz de separar os dois tipos tumorais. Foram observados 390 genes diferencialmente expressos, sendo 16 com expressão diminuída e 374 com expressão aumentada. Esta análise mostrou dois grupos, um formado por 16 casos de LM e outro composto por 22 casos de SP e os restantes seis casos de LM (Figuras 3A e 3B). Interessante notar que o grupo composto apenas de LM englobou todos os LM retroperitoniais da casuística (10). Os seis LM que se agruparam juntamente com os SP eram todos de extremidades.

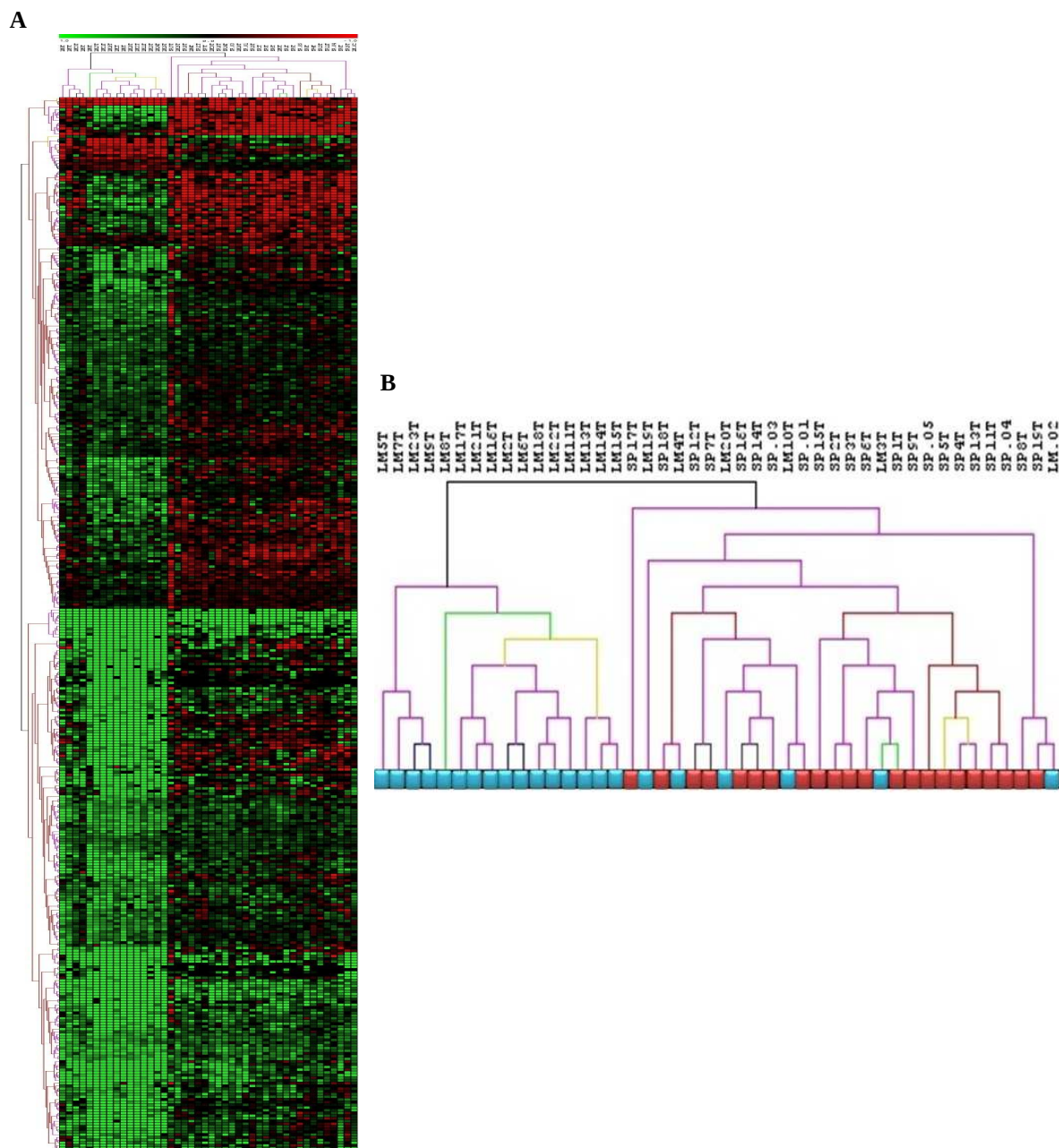


Figura 3 - A) Análise supervisionada baseada no tipo tumoral resultou em 390 genes diferencialmente expressos. Em verde os genes com expressão aumentada e em vermelho os genes com expressão diminuída. **B)** Esta análise resultou em dois grupos. Os SP estão representados em vermelho e LM em azul.

A análise dos dados usando as anotações das vias do KEGG (www.funnet.info) demonstrou que os genes com expressão diminuída fazem parte de vias metabólicas da arginina, prolina e porfirina, enquanto os genes com

expressão aumentada se relacionam com vias de sinalização (ErbB, cálcio e MAPK), adesão e contração muscular (Figura 4).

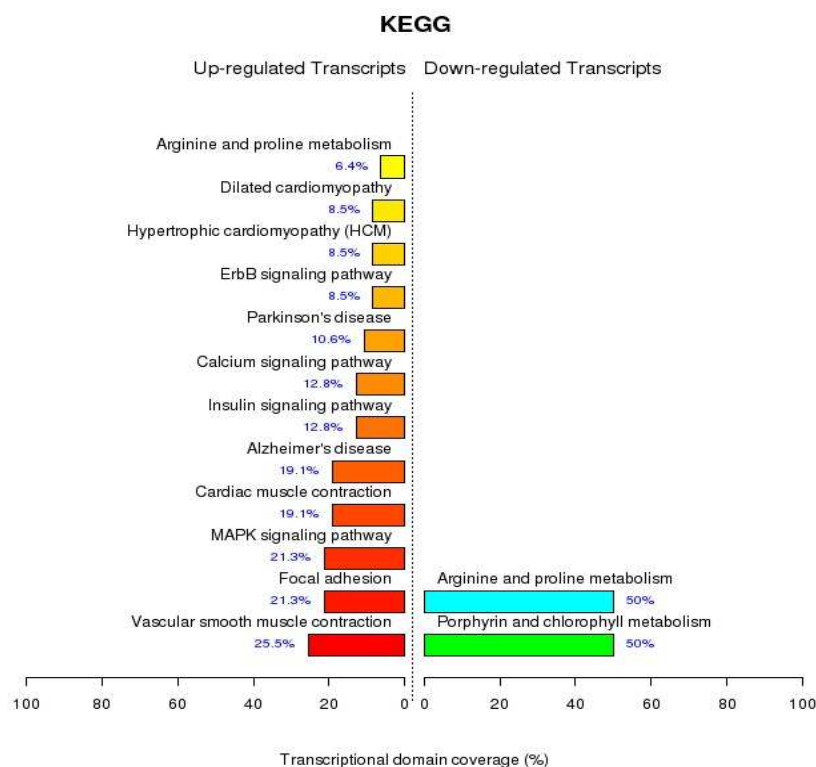


Figura 4 - Anotações das vias do KEGG (<http://www.funnet.info>) para os genes diferencialmente expressos entre o *cluster* de LM e o *cluster* de SP e LM.

A pesquisa usando a base de dados *Gene Ontology* (GO) revelou que a maioria dos genes com expressão aumentada participava de processos biológicos relacionados com a diferenciação celular, contração muscular e desenvolvimento do organismo, enquanto os genes com expressão diminuída estavam associados com respostas ao stress oxidativo e hipóxia (Figura 5A). Em relação à função molecular, metade dos genes com expressão diminuída estavam associados com funções de atividade de oxiredutase, enquanto 20% dos genes com expressão aumentada tiveram como função a ligação a actina, seguido pela atividade de oxiredutase e constituinte estrutural do músculo (Figura 5B).

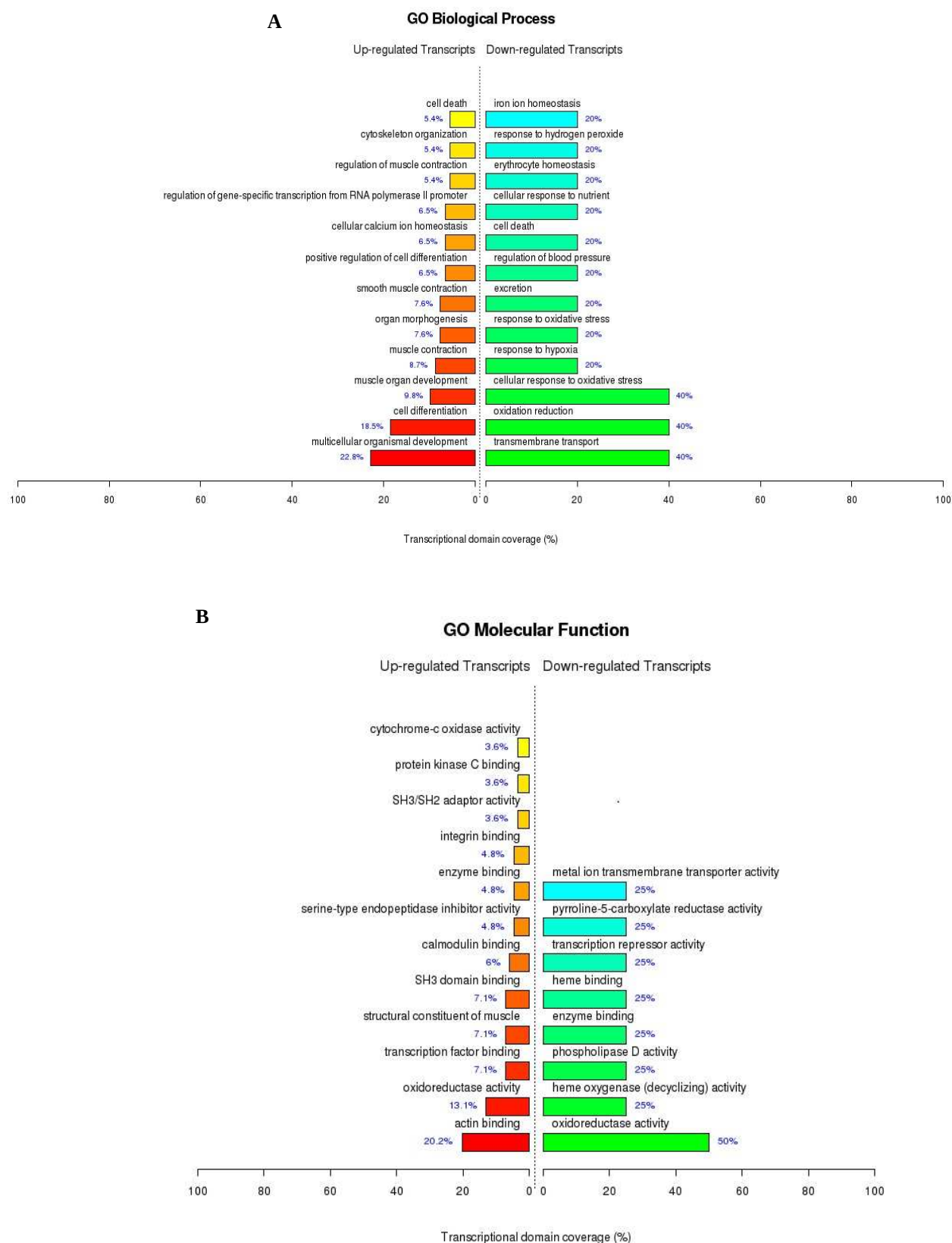


Figura 5 - Resultado da análise dos dados usando a base de dados *Gene Ontology* (<http://www.funnet.info>) resultando nos processos biológicos (A) e na função molecular (B) a que os genes diferencialmente expressos estão envolvidos.

Foi detectado que o *cluster* composto pelos 16 LM resultante da análise supervisionada continha genes com expressão aumentada associados a estrutura e função muscular, alguns específicos da musculatura lisa, incluindo actina, miosina e calponina. A Quadro 6 mostra esses genes e seu respectivo *fold change*. A maioria dos LM que se agrupou conjuntamente com os SP apresentava expressão diminuída desses marcadores, apesar de serem positivos para desmina ou SMA após análise imunoistoquímica.

Quadro 6 - Genes com expressão aumentada relacionados com a estrutura e função muscular observado no grupo dos 16 LM.

Gene	Nome	Fold change
<i>MYLK</i>	<i>myosin light chain kinase</i>	0,5115786
<i>CNN1</i>	<i>calponin 1, basic, smooth muscle</i>	0,5202636
<i>TPM2</i>	<i>tropomyosin 2 (beta)</i>	0,5782486
<i>CA3</i>	<i>carbonic anhydrase III, muscle specific</i>	0,58289635
<i>LMOD1</i>	<i>leiomodulin 1 (smooth muscle)</i>	0,5985782
<i>ACTG2</i>	<i>actin, gamma 2, smooth muscle, enteric</i>	0,59997165
<i>MYL9</i>	<i>myosin, light chain 9, regulatory</i>	0,61344564
<i>TAGLN</i>	<i>Transgelin</i>	0,6339436
<i>ACTBL2</i>	<i>actin, beta-like 2</i>	0,6363338
<i>TMOD1</i>	<i>tropomodulin 1</i>	0,6426155
<i>ACTC1</i>	<i>actin, alpha, cardiac muscle 1</i>	0,64602077
<i>ACTA1</i>	<i>actin, alpha 1, skeletal muscle</i>	0,65006465
<i>ACTA2</i>	<i>actin, alpha 2, smooth muscle, aorta</i>	0,67180085
<i>SVIL</i>	<i>Supervillin</i>	0,681278
<i>DSTN</i>	<i>destrin (actin depolymerizing factor)</i>	0,68825006
<i>SGCA</i>	<i>sarcoglycan, alpha</i>	0,68979406
<i>SYNPO2</i>	<i>synaptopodin 2</i>	0,69446224
<i>MYOM2</i>	<i>myomesin (M-protein) 2</i>	0,73326117
<i>MYOCD</i>	<i>Myocardin</i>	0,73374933
<i>ACTN1</i>	<i>actinin, alpha 1</i>	0,739447
<i>CALD1</i>	<i>caldesmon 1</i>	0,7411432
<i>SMTN</i>	<i>Smoothelin</i>	0,7662147
<i>MYL6</i>	<i>myosin, light chain 6, alkali, smooth muscle and non-muscle</i>	0,8094318

Quando se considera a localização anatômica dos sarcomas deste estudo, observa-se que na análise não supervisionada não houve uma separação clara entre os grupos, porém, entre os quatro grupos observados, o segundo apresentou a maior quantidade de sarcomas retroperitoneais (10 de 13 amostras), enquanto o terceiro e

quarto grupos apresentaram mais casos de sarcomas de extremidades (nove de 14 e cinco de sete casos, respectivamente). Os dois casos de sarcomas na região da cabeça e pescoço se localizaram no primeiro e segundo grupo, enquanto os três casos de sarcomas da região do tronco se associaram ao segundo e terceiro grupo.

Os casos que desenvolveram metástases ou recidivas durante o seguimento clínico se distribuíram de forma homogênea entre os quatro agrupamentos.

A análise supervisionada considerando a topologia dos casos revelou a presença de dois grandes grupos contendo 961 genes diferencialmente expressos: um composto por todos os sarcomas retroperitoniais, os três sarcomas de tronco e dois sarcomas de extremidades, enquanto o segundo grupo continha os 16 casos de sarcomas de extremidades, os dois casos de sarcomas da região da cabeça e pescoço e apenas dois casos de sarcomas do retroperitônio (Figura 6).

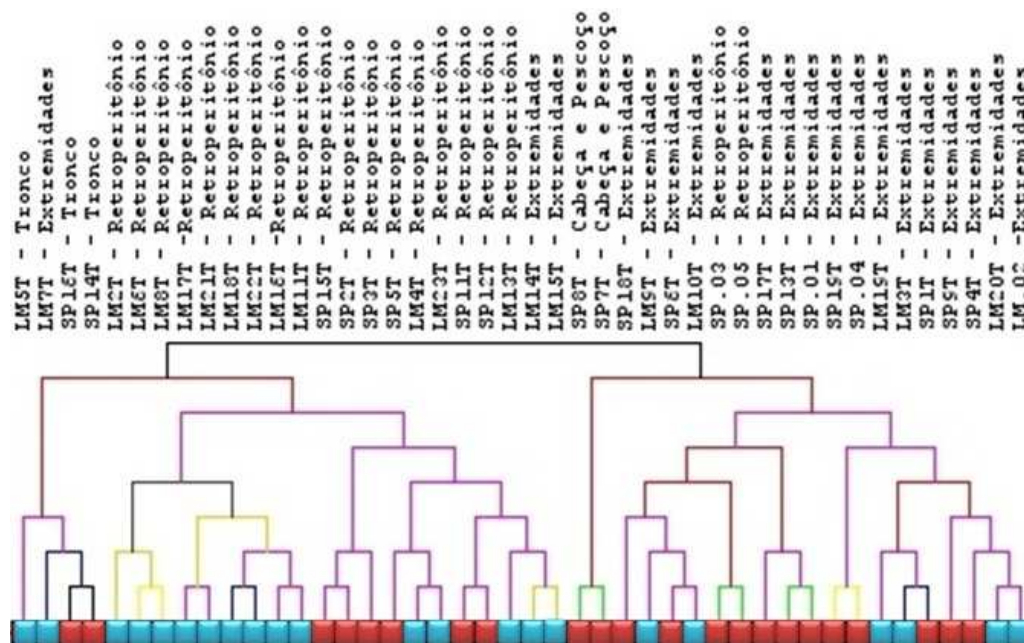


Figura 6 - Análise supervisionada dos 44 casos de sarcomas, considerando a topologia das amostras. Os sarcomas pleomórficos (SP) estão representados em vermelho e os leiomiossarcomas (LM) em azul.

Em uma nova tentativa de diferenciar molecularmente os SP dos LM, outros testes estatísticos de uma e duas classes, com valor de FDR=0, foram realizados, mas em nenhum deles foi possível separar os dois grupos tumorais. Inicialmente foram excluídos os dois casos com a síndrome de Li-Fraumeni (LM2T e LM22T) e foram realizadas análises supervisionadas e não supervisionadas para os 42 casos restantes. Na análise não supervisionada foi possível distinguir quatro grupos que não diferiram substancialmente dos anteriormente obtidos considerando todos os casos (Figura 7A). Por outro lado, a análise supervisionada revelou 129 genes diferencialmente expressos (expressão aumentada) que separavam as amostras em dois grandes grupos, semelhante ao detectado na análise de todas as amostras (Figura 7B).

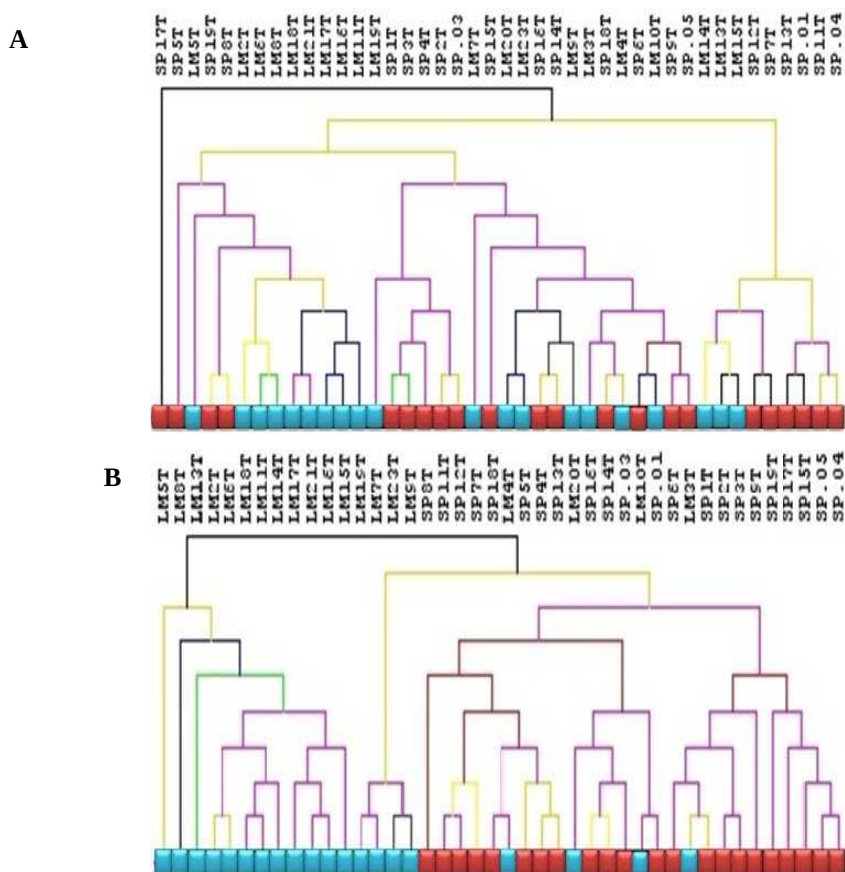


Figura 7 - Agrupamento entre os 42 casos de sarcomas por análise não supervisionada (A) e supervisionada (B), após exclusão dos casos com a síndrome de Li-Fraumeni. Os sarcomas pleomórficos (SP) estão representados em vermelho e os leiomiossarcomas (LM) em azul.

Da mesma forma, testes não supervisionados e supervisionados, excluindo os casos com metástase ao diagnóstico (LM6T, LM20T, SP18T e SP8T) e posteriormente com a exclusão dos casos com a síndrome de Li-Fraumeni e com metástase ao diagnóstico, apresentaram resultados semelhantes aos obtidos nas análises anteriores (Figuras 8A, 8B, 9A e 9B). Na análise supervisionada após a exclusão dos casos com metástase ao diagnóstico, o *cluster* que nas outras análises agrupava apenas LM, desta vez agrupou também um caso de SP (SP-01). O teste supervisionado revelou 83 genes diferencialmente expressos (expressão aumentada) no caso da análise excluindo os casos com metástase ao diagnóstico, enquanto o mesmo teste revelou 165 genes com expressão aumentada quando se excluía os casos com a síndrome de Li-Fraumeni e metástase ao diagnóstico.

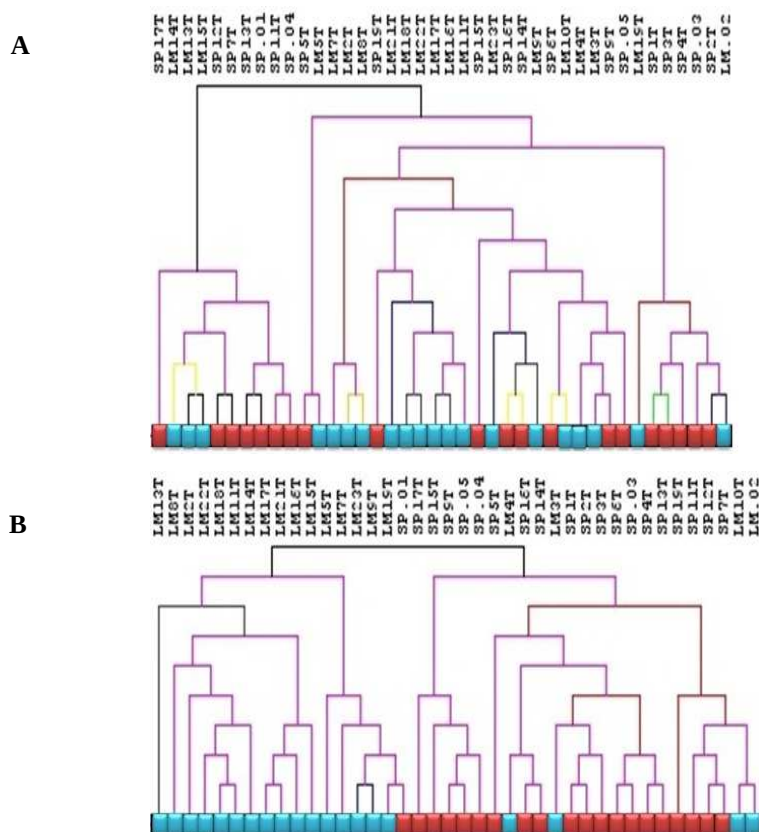


Figura 8. Agrupamento entre os 40 casos de sarcomas por análise não supervisionada (A) e supervisionada (B), após exclusão dos casos com metástase ao diagnóstico.

morfologia típica e positividade pelo menos para um marcador imunoistoquímico de músculo liso (AML, HHF-35 ou desmina) (14 LM). Nestas novas análises não foi possível separar os casos de SP e LM.

A análise não supervisionada gerou um único agrupamento, dividido em vários subgrupos, cujo padrão de expressão gênica não mostrou nenhuma relação com topologia, recidiva ou presença de metástase (Figura 10). Após a análise supervisionada foram detectados 127 genes diferencialmente expressos, nove com expressão diminuída e 118 com expressão aumentada, que separavam um grupo formado por 12 LM de um grupo formado pelos 10 SP e os dois LM restantes (Figura 11). Os dois LM que se agruparam com os SP (LM4T e LM10T) são casos que também se agruparam junto com SP em testes supervisionados anteriores. Entre os genes com expressão aumentada, alguns se associaram com a estrutura muscular, como o *MYLK*, *MYL9*, *ACTC1*, *SYNM* e *TPM2*, enquanto outros genes se relacionaram especificamente com a função do músculo liso, como o *SMTN*, *ACTG2* e *CNN1*.

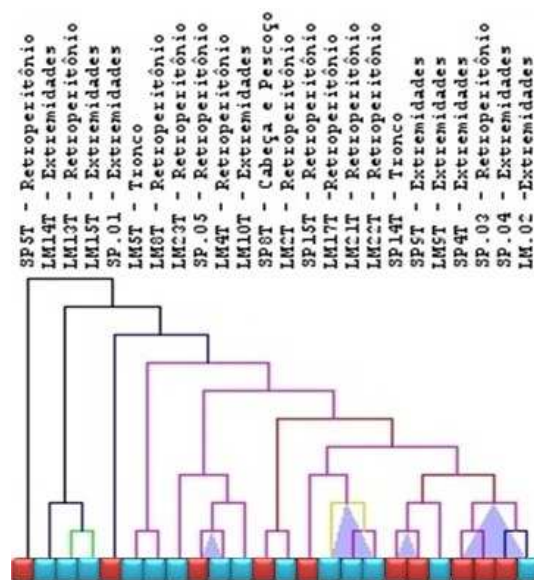


Figura 10 - Agrupamento entre os 24 casos de SP e LM mais característicos por análise não supervisionada. Os sarcomas pleomórficos (SP) estão representados em vermelho e os leiomiossarcomas (LM) em azul.

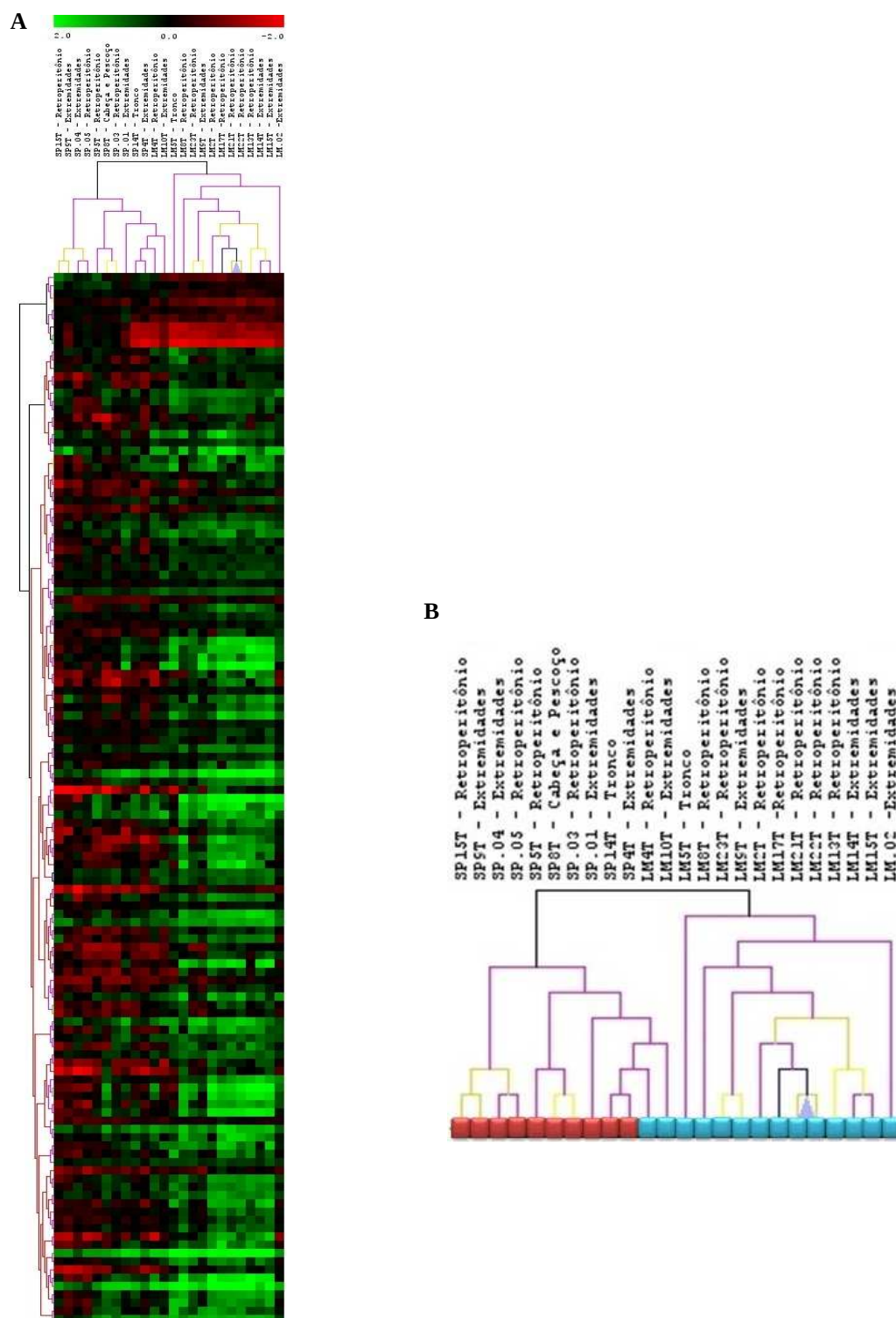
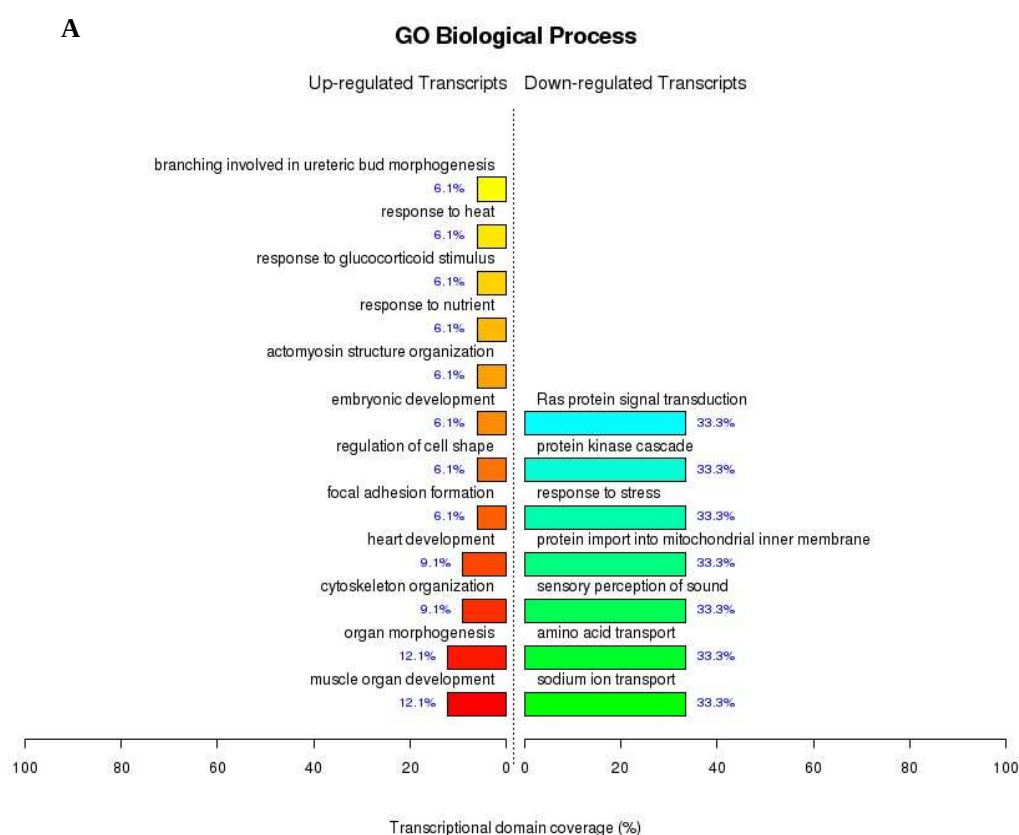


Figura 11 - A) Resultados da análise supervisionada entre os 24 casos de sarcomas demonstrando 127 genes diferencialmente expressos. Em verde os genes com expressão aumentada e em vermelho os genes com expressão diminuída. **B)** Agrupamento entre os 24 casos de sarcomas por análise supervisionada. Os sarcomas pleomórficos (SP) estão representados em vermelho e os leiomiossarcomas (LM) em azul.

Com o objetivo de avaliar os processos biológicos e funções moleculares relacionados aos genes diferencialmente expressos obtidos após as análises descritas acima, foi utilizado o banco de dados *Gene Ontology* (GO). Foi verificado que um terço dos genes com expressão diminuída se associavam a processos biológicos relacionados com a resposta ao stress e o transporte de íons sódio e aminoácidos, enquanto os genes com expressão aumentada atuavam no desenvolvimento do músculo e organização do citoesqueleto (Figura 12A). A função molecular da maioria dos genes com expressão aumentada estava relacionada com a ligação à actina, constituição estrutural do músculo e atividade motora. Por outro lado, metade dos genes com expressão diminuída estavam associados à ligação a íons sódio e atividade de MAP quinase (Figura 12B).



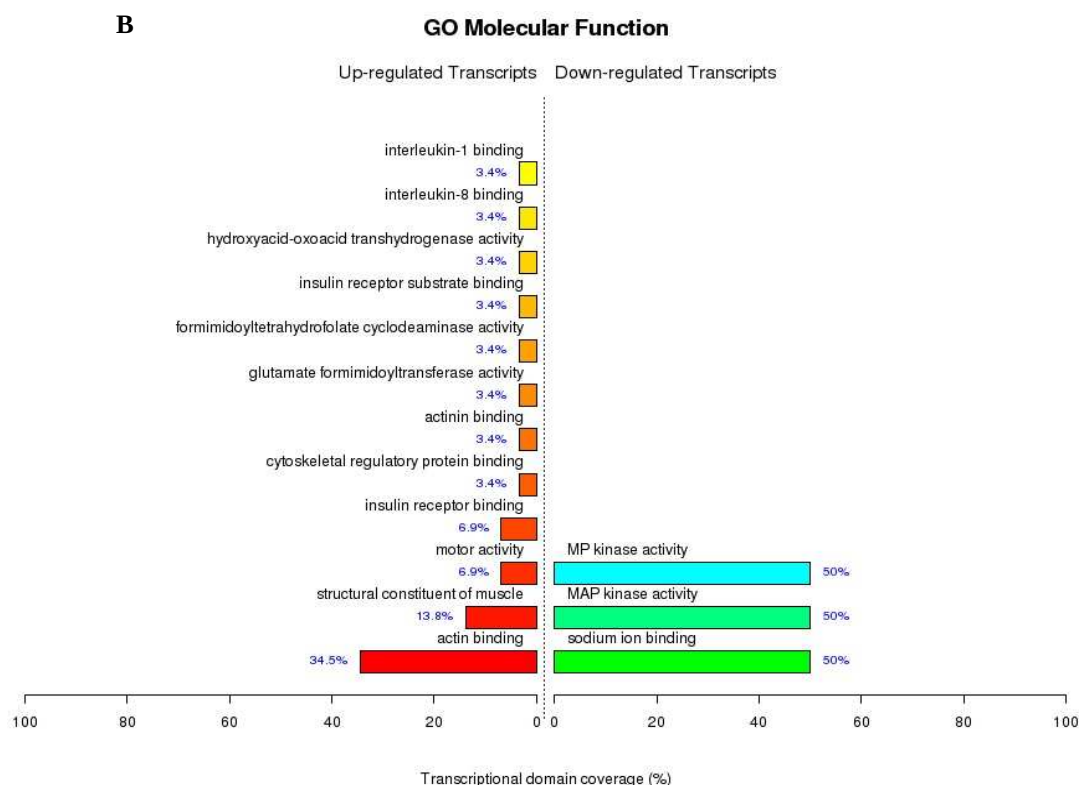


Figura 12 - Resultados da busca por processos biológicos (A) e funções moleculares (B) dos genes diferencialmente expressos usando o banco de dados *Gene Ontology* (ww.funnet.info).

Quatro amostras de tecido normal adjacente ao tumor também foram avaliadas, sendo três provenientes de casos de SP e um de paciente com LM. Inicialmente selecionamos doze amostras de tecidos adjacentes normais, mas só foi possível obter RNA de boa qualidade de quatro casos. O teste *SAM* de duas classes, com $FDR=0$, foi utilizado para fazer uma clusterização hierárquica pareada, ou seja, comparou-se os quatro tecidos adjacentes normais e os sarcomas (três SP e um LM) dos quais foram obtidos esses tecidos (Figura 13A). Neste caso foram detectados 40 genes que diferenciavam ambos os grupos, com *fold change* variando de -74 a 34 (Quadro 7). Utilizou-se as anotações das vias do KEGG para classificar os genes com expressão diferenciada (Figura 13B) e observou-se que os genes com expressão aumentada nos casos tumorais em relação aos normais adjacentes estavam associados às vias de

sinalização do *TP53* e ao ciclo celular, enquanto os genes com expressão diminuída se associavam a vias metabólicas em geral.

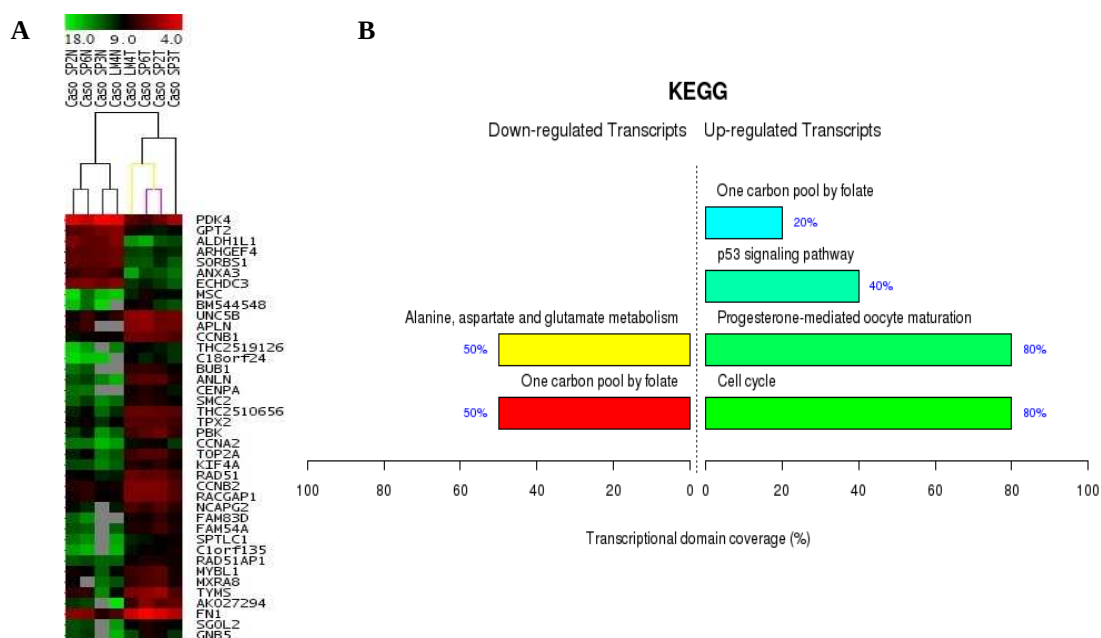


Figura 13 - A) Análise de clusterização hierárquica pareada, utilizando as quatro amostras tumorais e seus respectivos tecidos adjacentes normais. Em vermelho os genes com expressão aumentada e em verde os genes com expressão diminuída. **B)** Anotações das vias do KEGG para os genes diferencialmente expressos nos grupos pareados, quatro tecidos adjacentes normais e quatro amostras tumorais.

Quadro 7 - Genes diferencialmente expressos entre o grupo de quatro amostras tumorais e seus respectivos tecidos adjacentes normais. Os *fold changes* positivos indicam aumento de expressão e os negativos diminuição da expressão dos transcritos.

Gene	Nome	Fold change
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1	- 74,4772733
ECHDC3	Enoyl Coenzyme A hydratase domain containing 3	- 48,9213174
SORBS1	Sorbin and SH3 domain containing 1	- 32,8040662
ARHGEF4	Rho guanine nucleotide exchange factor (GEF) 4	- 26,5985556
ANXA3	Annexin A3	- 22,4217464
PDK4	Pyruvate dehydrogenase kinase, isozyme 4	- 18,2308636
GPT2	Glutamic pyruvate transaminase (alanine aminotransferase) 2	- 16,1014477
GNB5	Guanine nucleotide binding protein (G protein), beta 5	2,30829034
SGOL2	Shugoshin-like 2 (S. pombe)	4,01496831
MYBL1	V-myb myeloblastosis viral oncogene homolog (avian)-like 1	4,96373123
APLN	Apelin, AGTRL1 ligand	6,0378378
RAD51AP1	RAD51 associated protein 1	6,51213316
CCNA2	Cyclin A2	6,65779131
SPTLC1	Serine palmitoyltransferase, long chain base subunit 1	6,69094831
THC2519126	Unknown	6,82993572
KIF4A	Kinesin family member 4A	7,13589061
NCAPG2	Non-SMC condensin II complex, subunit G2	7,14424967
CENPA	Centromere protein A	7,25490715
SMC2	Structural maintenance of chromosomes 2	7,99155721
FAM54A	family with sequence similarity 54, member A	8,25235738
TPX2	TPX2, microtubule-associated, homolog (Xenopus laevis)	8,61069537
FAM83D	Family with sequence similarity 83, member D	8,87947882
CCNB1	Cyclin B1	9,9931811
CCNB2	Cyclin B2	10,3797209
BM544548	Unknown	10,6886367
C18orf24	Spindle and kinetochore associated complex subunit 1	11,1413156
(SKA1)		
BUB1	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)	11,5988692
UNC5B	Unc-5 homolog B (C. elegans)	11,8637196
RACGAP1	Rac GTPase activating protein 1	12,0174904
MSC	Musculin (activated B-cell factor-1)	12,7990932
RAD51	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)	12,9213419
C1orf135	Chromosome 1 open reading frame 135	13,5360162
TOP2A	Topoisomerase (DNA) II alpha 170kDa	14,3535456
FN1	Fibronectin 1	14,5511624
MXRA8	Matrix-remodelling associated 8	14,5777292
TYMS	Thymidylate synthetase	14,7236862
THC2510656	Unknown	20,0346618
PBK	PDZ binding kinase	22,3958544
AK027294	Unknown	26,4073934
ANLN	Anillin, actin binding protein	34,2140361

O teste supervisionado realizado nos 44 casos de sarcomas separou um grupo de 16 LM de um grupo composto pelos 22 SP e os seis LM. Entre os 390 genes resultantes desta análise foram selecionados aqueles com *fold change* maiores que 0,7 (211) e menores que -1 (16) e submetidos a análise de redes, vias gênicas, doenças relacionadas e função celular ou molecular com a utilização do software Ingenuity Pathways Analysis (IPA) (Ingenuity Systems). Os scores para cada item foram calculados pelo teste exato de Fisher e são iguais a $-\log(\text{valor de } P)$. As principais doenças e funções celulares ou moleculares relacionadas aos genes com aumento ou diminuição de expressão estão representadas na Figura 14 (maiores scores). A avaliação dos resultados mostrou grande número de moléculas alteradas presentes em doenças neuronais, no câncer e em distúrbios genéticos, enquanto houve predominância de genes com função no desenvolvimento, crescimento, proliferação, morfologia, organização e morte celular.

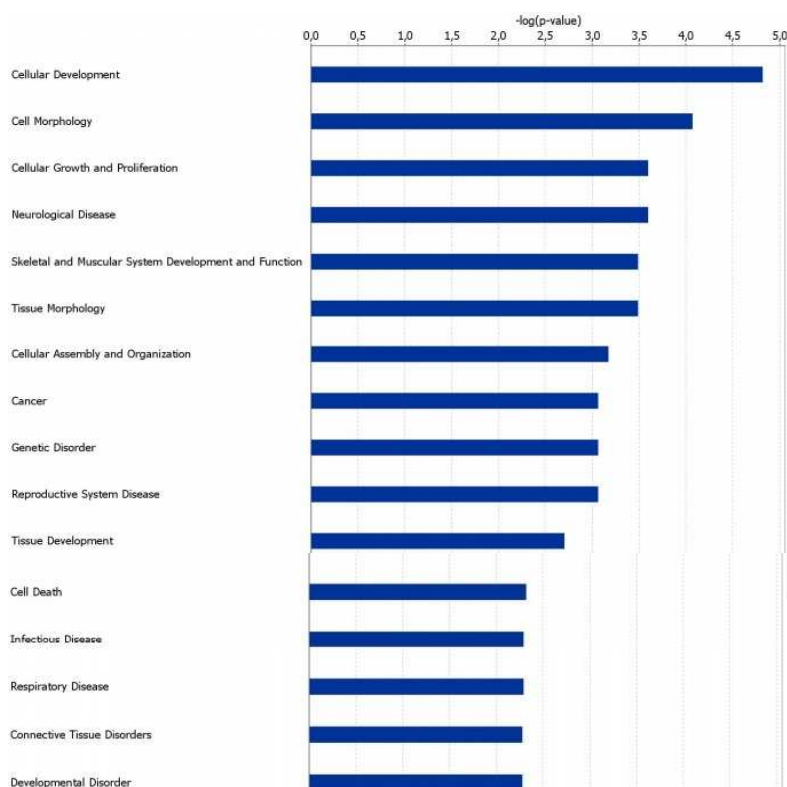


Figura 14 - Genes diferencialmente expressos entre os 44 casos de sarcomas com $-1 > \text{fold change} > 0,7$ associados as doenças e funções celulares e moleculares que apresentaram os maiores scores. Os scores são calculados pelo teste exato de Fisher e são iguais a $-\log(\text{valor de } P)$. Imagem gerada pelo *software* Ingenuity Pathways Analysis (IPA).

Os genes diferencialmente expressos também se correlacionaram com importantes vias canônicas previamente descritas. As vias com maiores *scores* estão representadas na Figura 15 e Quadro 8. O *score* mais alto foi obtido para a via de sinalização do gene *ILK* (8 genes alterados), seguida por genes associados às seguintes vias de sinalização: fosfolipase C (9 genes alterados), hipertrofia cardíaca (8 genes alterados) e proteína quinase A (8 genes alterados).

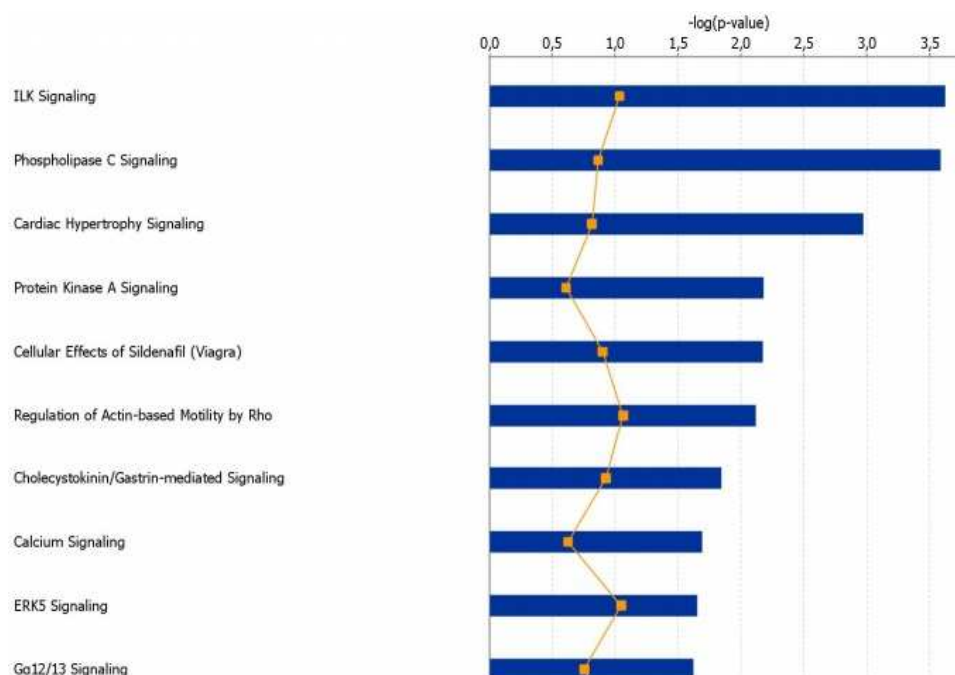


Figura 15 - Vias canônicas correlacionando com os genes diferencialmente expressos com $-1 > \text{fold change} > 0,7$ em sarcomas e que apresentaram os maiores scores. Os scores para cada via são calculados pelo teste exato de Fisher e são iguais a $-\log(\text{valor de } P)$. Imagens geradas pelo software *Ingenuity Pathways Analysis* (IPA).

Quadro 8 - Relação dos genes diferencialmente expressos nas quatro vias canônicas que apresentaram maiores scores.

Via canônica	Genes alterados
Sinalização do gene <i>ILK</i>	↑ <i>MUC1</i> , ↑ <i>MYL9</i> , ↑ <i>TGFB11I</i> , ↑ <i>MYL6</i> , ↑ <i>FLNA</i> , ↑ <i>ILK</i> , ↑ <i>FBNP1</i> , ↑ <i>ACTN1</i>
Sinalização da fosfolipase C	↑ <i>MYL9</i> , ↓ <i>HMOX1</i> , ↑ <i>SRC</i> , ↑ <i>MYL6</i> , ↑ <i>PPP1R12A</i> , ↑ <i>MEF2C</i> , ↑ <i>FBNP1</i> , ↑ <i>HDAC5</i> , ↑ <i>CALM1</i>
Sinalização da hipertrofia cardiac	↑ <i>MYL9</i> , ↑ <i>SRF</i> , ↑ <i>CACNA1C</i> , ↑ <i>MYL6</i> , ↑ <i>MEF2C</i> , ↑ <i>FBNP1</i> , ↑ <i>HSPB1</i> , ↑ <i>CALM1</i>
Sinalização da proteína quinase A	↑ <i>MYL9</i> , ↑ <i>PYGM</i> , ↑ <i>BAD</i> , ↑ <i>MYL6</i> , ↑ <i>FLNA</i> , ↑ <i>PPP1R12A</i> , ↑ <i>PHKG1</i> , ↑ <i>CALM1</i>

↑Aumento de expressão; ↓Diminuição de expressão

Nove redes gênicas foram obtidas com scores significativos na análise entre os genes resultantes da análise de expressão gênica. A Figura 16 ilustra as quatro principais redes envolvendo os genes alterados quanto ao padrão de expressão

(maiores *scores*). Os genes com alteração de expressão gênica em cada uma das vias estão descritos na Quadro 9.

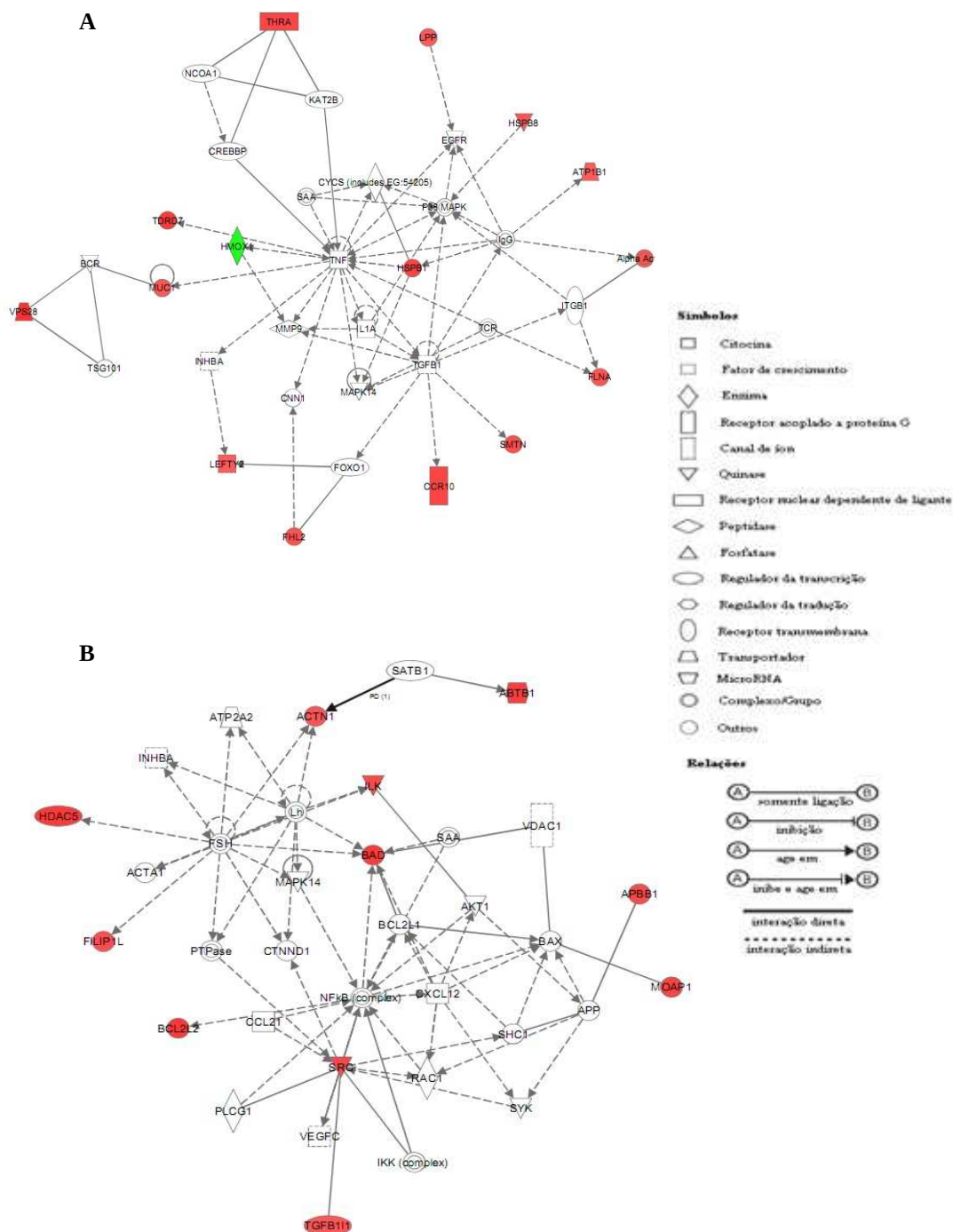
A principal rede alterada quanto a expressão gênica está relacionada a doenças do sistema reprodutivo, desenvolvimento e morte celular (Figura 16A). Essa rede é formada por 15 genes alterados, incluindo aumento de expressão de transcritos em genes da família das *heat shock proteins* (*HSPB1* e *HSPB8*) e em genes associados com a função muscular (*Alpha actinin* e *SMTN*, este específico de músculo liso). O gene *HMOX1* foi o único que apresentou diminuição de expressão e faz parte da família de oxigenases do grupo heme.

A segunda rede contém 35 genes relacionados com o ciclo, morfologia e morte celular (Figura 16B). Genes conhecidos e associados com morte celular (*TGPB111*, *BCL2L2* e *BAD*) apresentam aumento de expressão, assim como o oncogene *SRC* e o gene *ACTN1*, o qual está associado com a função muscular. Outro gene importante com aumento de expressão foi o *ILK*, o qual se relaciona com inúmeros processos celulares.

A terceira rede contém 10 genes alterados, os quais estão envolvidos em processos de replicação, recombinação e reparo a danos no DNA, além do ciclo e morte celular (Figura 16C). Nesta rede foi detectado aumento de expressão de genes supressores tumorais (*RBBP7* e *TP53INP1*), associados a metástase (*MTA3*) e relacionados com a replicação celular (*CHEK1*).

Na quarta rede gênica estão representados 7 genes associados com a movimentação celular, tráfego de células imunes e com o desenvolvimento e funcionamento do sistema hematopoiético (Figura 16D). Nesta rede estão incluídos os genes *MYOCD*, *MEF2C*, *SRF* e *PPP1R12A* (todos com aumento de expressão),

que se relacionam com o processo de desenvolvimento e diferenciação da musculatura lisa. O único gene que apresentou diminuição de expressão foi o *S100A9*, o qual faz parte da família de proteínas S100 e se relaciona com o ciclo, progressão e diferenciação celular.



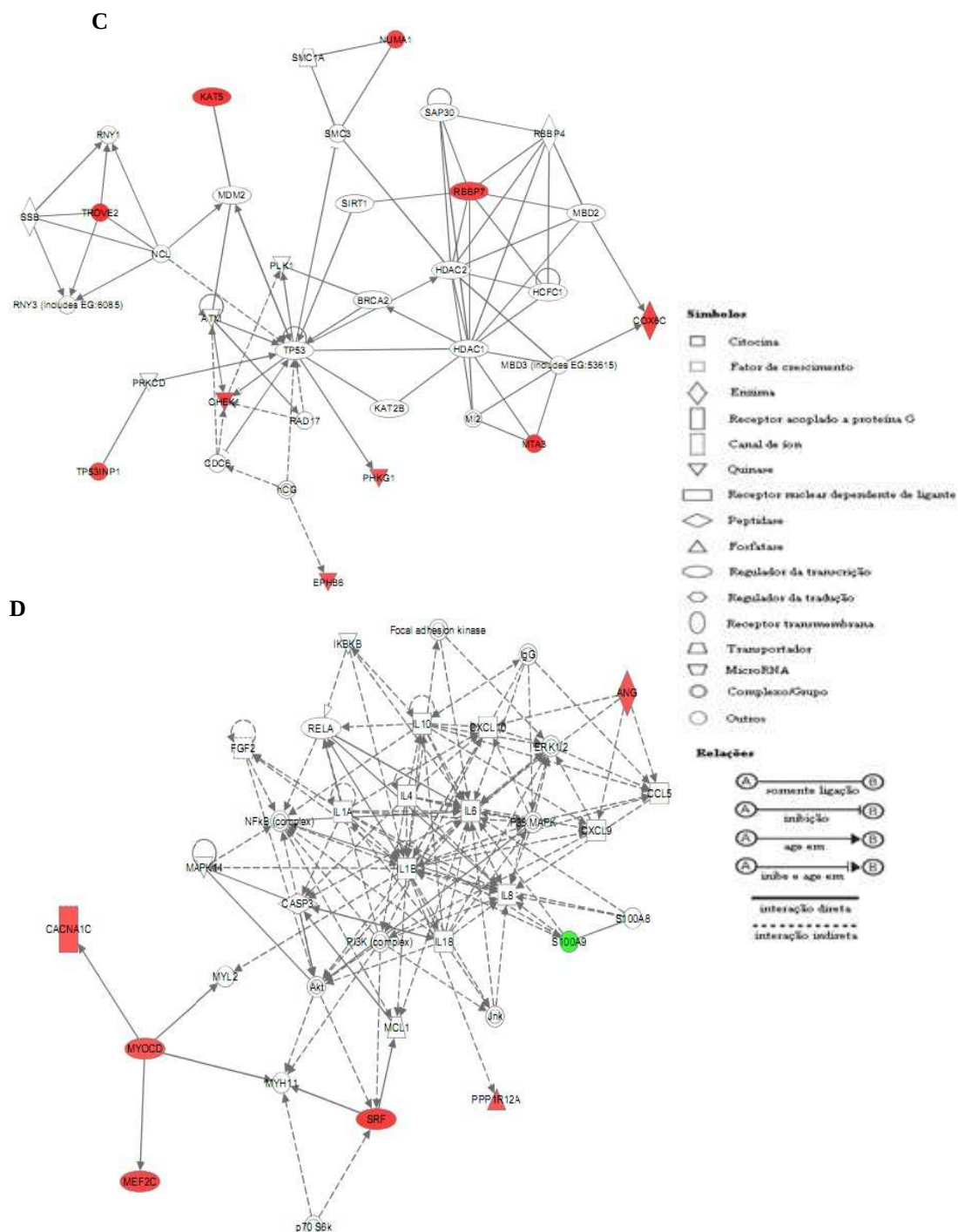


Figura 16 - Representação esquemática de quatro redes baseadas nos genes diferencialmente expressos ($-1 > \text{fold change} > 0,7$). Em **(A)** está representada a rede associada com doenças do sistema reprodutivo/desenvolvimento e morte celular, em **(B)** ciclo/morfologia/morte celular, em **(C)** replicação/recombinação/reparo do DNA/ciclo/morte celular e em **(D)** Movimentação celular/tráfego de células imunes/desenvolvimento e função do sistema hematopoiético. Em vermelho estão representados genes com aumento de expressão, verde para diminuição de expressão, branco para moléculas incorporadas com base no banco de dados do *software* e cinza para genes com alteração de expressão que não atingiram o critério de *fold change*. Imagens geradas pelo *software Ingenuity Pathways Analysis (IPA)*.

Quadro 9 - Descrição dos genes com e sem alteração de expressão gênica presentes em cada uma das quatro redes gênicas com maiores scores.

Rede gênica	Funções	No. Genes alterados	Genes ^a
1	Doenças do sistema reprodutivo/desenvolvimento e morte celular	15	↑ Alpha Actinin , ↑ ATP1B1 , BCR, ↑ CCR10 , CNN1, CREBBP, CYCS, EGFR, ↑ FHL2 , ↑ FLNA , FOXO1, ↓ HMOX1 , ↑ HSPB1 , ↑ HSPB8 , IgG, IL1A, INHBA, ITGB1, KAT2B, ↑ LEFTY2 , ↑ LPP , MAPK14, MMP9, ↑ MUC1 , NCOA1, P38MAPK, SAA, ↑ SMTN , TCR, ↑ TDRD7 , TGFB1, ↑ THRA , TNF, TSG101, ↑ VPS28
2	Ciclo/morfologia/morte celular	11	↑ ABTB1 , ACTA1, ↑ ACTN1 , AKT1, ↑ APBB1 , APP, ATP2A2, ↑ BAD , BAX, BCL2L1, ↑ BCL2L2 , CCL21, CTNND1, CXCL12, ↑ FILIP1L , FSH, ↑ HDAC5 , IKK (complex), ↑ ILK , INHBA, Lh, MAPK14, ↑ MOAP1 , NFkB (complex), PLCG1, PTPase, RAC1, SAA, SATB1, SHC1, ↑ SRC , SYK, ↑ TGFB1I1 , VDAC1, VEGFC
3	Replicação/recombinação/repairo do DNA/ciclo/morte celular	10	ATM, BRCA2, CDC6, ↑ CHEK1 , ↑ COX6C , ↑ EPHB6 , HCFC1, hCG, HDAC1, HDAC2, ↑ KAT5 , KAT2B, MBD2, MBD3, MDM2, Mi2, ↑ MTA3 , NCL, ↑ NUMA1 , ↑ PHKG1 , PLK1, PRKCD, RAD17, RBBP4, ↑ RBBP7 , RNY1, RNY3, SAP30, SIRT1, SMC3, SMC1A, SSB, TP53, ↑ TP53INP1 , ↑ TROVE2
4	Movimentação celular/Tráfego de células imunes/desenvolvimento e função do sistema hematopoiético	7	Akt, ↑ ANG , ↑ CACNA1C , CASP3, CCL5, CXCL9, CXCL10, ERK1/2, FGF2, Focal adhesion kinase, IgG, IKBKB, IL4, IL6, IL8, IL10, IL18, IL1A, IL1B, Jnk, MAPK14, MCL1, ↑ MEF2C , MYH11, MYL2, ↑ MYOCD , NFkB (complex), P38 MAPK, p70S6k, PI3K (complex), ↑ PPP1R12A , RELA, S100A8, ↓ S100A9 , ↑ SRF

^a: os genes em destaque apresentaram alteração de expressão; ↑: aumento de expressão gênica; ↓: diminuição de expressão gênica

A confirmação dos resultados obtidos pela técnica de expressão gênica global foi realizada por experimentos de RT-qPCR. A seleção das 30 amostras (26 sarcomas e 4 tecidos normais adjacentes) foi baseada no teste supervisionado dos 44 sarcomas e principalmente na disponibilidade de RNA. Os genes *BAD*, *SRC* e *SRF* se associaram com duas das quatro principais redes gênicas geradas pelo *software* Ingenuity Pathways Analysis (IPA) e foram selecionados para a confirmação dos resultados obtidos pela comparação de SP e LM. O *TOP2A* foi um dos genes com maior *fold change* na comparação entre os quatro sarcomas e seus respectivos tecidos adjacentes normais sendo selecionado para confirmar estes resultados.

Para avaliar a expressão dos genes alvo, inicialmente investigou-se a estabilidade de expressão de três genes candidatos a normalizadores (*GAPDH*, *HPRT* e *HMBS*) em 26 casos previamente analisados por expressão gênica global. Todos apresentaram estabilidade adequada para a normalização dos genes alvo (coeficiente de estabilidade menor que 1,5), sendo que o gene *HPRT* apresentou expressão mais estável (M=1,144) e o *GAPDH* menos estável (M=1,473) (Quadro 10). A melhor dupla indicada pelo programa foi *HPRT* e *HMBS* (M=0,856), sendo considerado o conjunto mais adequado de genes normalizadores a ser utilizado (Figura 17).

Quadro 10 - Expressão relativa, fator de normalização (FN) e coeficiente de estabilidade de expressão (M).

Amostras	<i>GAPDH</i>	<i>HPRT</i>	<i>HMBS</i>	FN
POOL	1.000	1.000	1.000	1.5098
LM-02	2.315	9.170	3.571	6.3897
LM11T	0.784	0.188	0.063	0.3173
LM13T	0.128	0.032	0.085	0.1062
LM15T	0.645	0.228	0.283	0.5232
LM17T	1.190	2.213	1.585	2.4309
LM22T	2.236	3.260	0.905	2.8316
LM23T	1.004	2.811	2.791	3.0040
LM2T	1.602	1.608	1.096	2.1339
LM5T	0.184	0.213	0.137	0.2644
LM6T	1.083	2.221	1.334	2.2268
LM8T	1.539	8.102	1.314	3.8346
SP-03	0.536	2.842	2.262	2.2804
SP-04	0.821	1.984	1.597	2.0764
SP12T	0.823	0.292	0.407	0.6956
SP14T	0.492	2.432	2.383	2.1409
SP15T	1.300	2.803	4.937	3.9560
SP19T	0.398	0.515	0.381	0.6453
SP1T	0.005	0.029	0.026	0.0235
SP4T	0.061	0.080	0.122	0.1270
SP5T	1.617	6.706	8.496	6.8188
SP8T	0.136	0.529	0.329	0.4335
LM4T	0.507	0.544	0.390	0.7180
SP2T	0.107	0.215	0.126	0.2153
SP3T	0.218	0.219	0.151	0.2917
SP6T	1.153	10.875	12.893	8.2437
M<1.5	1.473	1.144	1.185	

Legenda: O gene *HPRT* apresentou expressão mais estável (verde) individualmente (M=1,144) e o *GAPDH* é o mais variável (vermelho) (M=1,473) entre as 25 amostras de sarcomas testadas. O fator de normalização (FN) representa a média geométrica da expressão relativa dos três genes controles avaliados

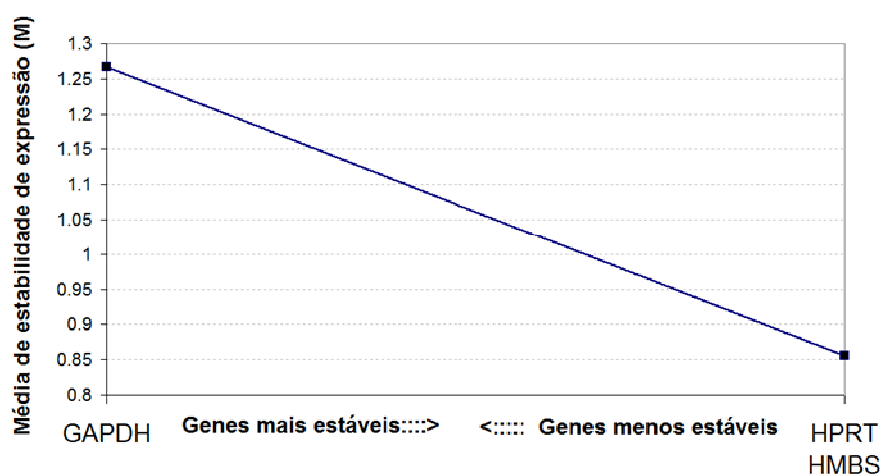


Figura 17 - Média do coeficiente de estabilidade de expressão (M). Cálculo para determinar o melhor par de genes controles, após a exclusão progressiva dos genes de expressão menos estável até permanecer apenas a melhor dupla, indicando que *HPRT* e *HMBS* representam o melhor par de genes ($M=0,856$).

A Figura 18 ilustra a expressão relativa dos genes *SRC*, *SRF* e *BAD* entre as 25 amostras de sarcomas que foram incluídas na análise de RT-qPCR. Pelo teste estatístico não paramétrico de Mann-Whitney a expressão diminuída dos três genes em sarcomas pleomórficos foi altamente significativa, tanto pela técnica de expressão gênica global como por RT-qPCR.

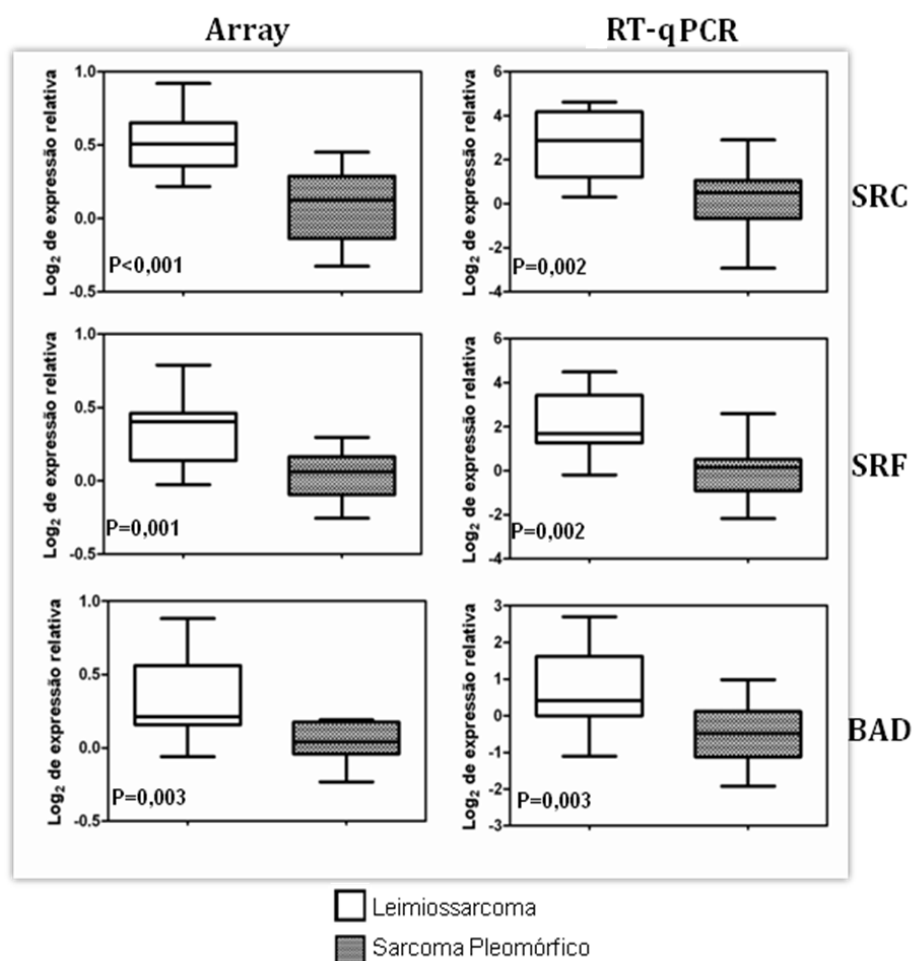


Figura 18 - Expressão relativa dos genes estudados por RT-qPCR e análise de expressão gênica global nas amostras de LM e SP. O gráfico mostra a distribuição de todos os valores entre as barras, com exceção de *outliers* e valores extremos. Foi utilizado o teste de Mann-Whitney para estas comparações.

Para a comparação entre tecido normal adjacente e tumoral, utilizando os 4 pares de amostras analisados por *oligoarray*, repetiu-se o teste pelo geNorm, que mais uma vez considerou o *GAPDH* como menos estável (Quadro 11). Assim, foram utilizados os genes *HPRT* e *HMBS* para a normalização da expressão do *TOP2A*. A expressão do gene *TOP2A* foi aumentada nas amostras de sarcomas em relação aos tecidos normais adjacentes por ambas as técnicas (Figura 19). Entretanto, apesar da evidente expressão aumentada de *TOP2A* nos sarcomas, o teste não-paramétrico de

Wilcoxon não demonstrou significância estatística nos dois casos, com $P=0,068$, provavelmente pelo pequeno número de amostras avaliadas.

Quadro 11 - Coeficiente de estabilidade de expressão (M).

Genes	M	
<i>GAPDH</i>	1.620	-
<i>HPRT</i>	1.465	1,225
<i>HMBS</i>	1.380	1,225

Legenda: O nível de expressão mais estável individualmente foi detectado para *HMBS* ($M=1,380$) e *GAPDH* foi o mais variável ($M=1,620$) nas quatro amostras de sarcomas e seus respectivos tecidos adjacentes normais testados. Após a exclusão do *GAPDH*, o valores de M para o *HMBS* e *HPRT* passam para 1,225 (em verde) representando a melhor dupla de genes.

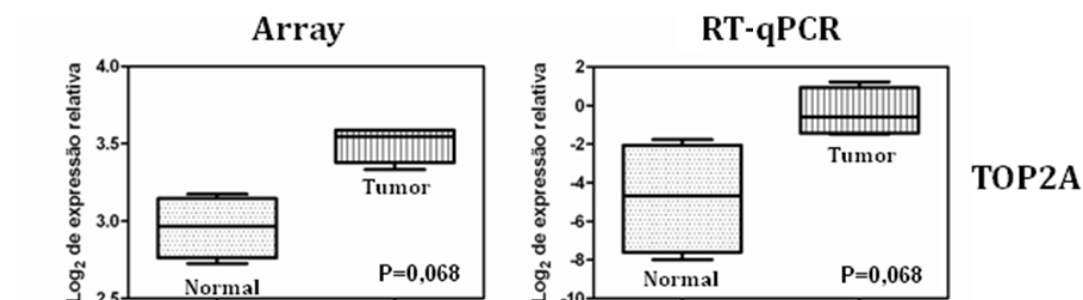


Figura 19 - Expressão relativa do *TOP2A* por RT-qPCR e expressão gênica global nas amostras tumorais e normais. O gráfico mostra a distribuição de todos os valores entre as barras, com exceção de *outliers* e valores extremos. Foi utilizado o teste de Wilcoxon para estas comparações.

6 DISCUSSÃO

A classificação dos sarcomas é baseada primordialmente na análise histológica. Alguns subtipos apresentam morfologia característica e alterações moleculares específicas que facilitam o diagnóstico, como no caso do sarcoma de Ewing (translocação *EWSR1-FLI1*) e GIST (mutação do *KIT*). Por outro lado, a grande maioria dos sarcomas, compostos por células malignas alongadas, não apresenta alterações genéticas específicas, que aliadas à morfologia pouco característica (pleomórfica) representam um grande desafio diagnóstico para os patologistas (NIELSEN e WEST 2010).

A classificação histopatológica tradicional não leva em consideração a complexidade molecular dos sarcomas, deixando uma parcela significativa de neoplasias não classificadas ou agrupadas em entidades controversas, como os SP (GUILLOU e AURIAS 2010; KONSTANTINOPOULOS et al. 2010). Devido à falta de fenótipos que indiquem algum tipo de diferenciação celular, ensaios de imunohistoquímica representam a única forma de obter mais informações para o diagnóstico, que é notoriamente inconsistente ou inconclusiva (TOGUUCHIDA e NAKAYAMA 2009). Neste contexto, estudos de perfis de expressão gênica podem fornecer informações adicionais sobre o comportamento biológico destas neoplasias, identificar biomarcadores diagnósticos e prognósticos e conduzir ao desenvolvimento de novas abordagens terapêuticas (SKUBITZ e D'ADAMO 2007; SKUBITZ et al. 2008; BECK et al. 2010b). Um desafio particular na classificação dos sarcomas é elucidar a natureza dos sarcomas pleomórficos. Atualmente, o termo

SP (antigo FHM) é utilizado para designar SPM que não apresentam uma linha de diferenciação celular específica e onde foram excluídas a diferenciação epitelial, melanocítica e linfóide. Resumidamente, trata-se de um diagnóstico por exclusão (DEI TOS 2006; FLETCHER 2006; THWAY 2009).

Estudos de expressão gênica em larga escala têm sido realizados para caracterizar perfis genéticos. Estes estudos indicam que os SP não representam uma entidade biológica distinta, mas uma coleção heterogênea de sarcomas de alto grau que compartilham padrões de expressão gênica com outros sarcomas de alto grau assim como a presença de cariótipos complexos, principalmente com os LM (NIELSEN et al. 2002; SEGAL et al. 2003; BAIRD et al. 2005; FRANCIS et al. 2007; NAKAYAMA et al. 2007; CARNEIRO et al. 2009; BECK et al. 2010a; KONSTANTINOPOULOS et al. 2010).

Um dos primeiros estudos a avaliar o perfil de expressão gênica pela metodologia de *microarray* foi realizado por SHMULEVICH et al. (2002). Os autores analisaram três amostras de pacientes com LM e concluíram que a heterogeneidade intratumoral não influencia significativamente o padrão de expressão gênica. FRANCIS et al. (2005) avaliaram a heterogeneidade intratumoral e intertumoral de SP e LM. Os autores utilizaram oito amostras de um único caso de SP e 10 amostras de um único caso de LM, comparando os resultados com amostras provenientes de 20 casos de SP e 16 casos de LM. As amostras provenientes de um mesmo caso se agruparam em análises de clusterização hierárquica. Eles concluíram que a heterogeneidade intratumoral parece ter um impacto pequeno na variabilidade dos perfis de expressão gênica, porém, pode ter um impacto maior nos resultados de séries pequenas de tumores e limitar a validade dos achados.

No presente estudo comparou-se o perfil de expressão gênica de 44 amostras tumorais, correspondendo a 22 LM e 22 SP, utilizando a plataforma da Agilent (4 x 44k).

Todos os tumores foram revisados por um mesmo patologista que utilizou análises de imunoistoquímica, na maioria dos casos, para corroborar o diagnóstico histopatológico. Os principais marcadores musculares utilizados foram: AML, HHF-35, desmina e, em raros casos, calponina e caldesmon. Alguns casos poderiam ter sido diagnosticados como lipossarcomas ou mixofibrossarcomas, porém, a falta de evidência histológica e os dados fornecidos pelos imunomarcadores testados levaram à classificação de SP. Entretanto, não é possível descartar completamente as outras possibilidades diagnósticas. Considerando as características morfológicas e imunoistoquímicas, quatro SP (SP2T, SP7T, SP18T e SP19T) e quatro LM (LM6T, LM7T, LM16T e LM20T) foram os mais difíceis de diagnosticar. A morfologia de todos estes casos não foi característica e poderiam ser nomeados apenas como sarcomas de alto grau SOE (*inclassificáveis*). O diagnóstico de SP e LM pouco diferenciados é complexo e muitas vezes a morfologia e os marcadores imunoistoquímicos disponíveis não são suficientes para classificá-los adequadamente. Por isso, outras metodologias, como perfis de expressão gênica em larga escala, podem ser utilizados como ferramentas adicionais para diferenciar os sarcomas pleomórficos.

A análise de expressão gênica utilizando ferramentas de bioinformática para clusterização hierárquica não supervisionada resultou na detecção de quatro grandes agrupamentos não sendo possível separar as duas neoplasias. Todos os grupos apresentaram os dois tipos tumorais, SP e LM. O primeiro e quarto agrupamentos

apresentaram mais amostras de SP do que LM, sete a três e cinco a dois, respectivamente. Por outro lado, o segundo agrupamento apresentou mais amostras de LM (10) do que SP (três). Na tentativa de separar estes subtipos de sarcomas, diversas outros testes não supervisionados foram realizados, seja excluindo os pacientes com a síndrome de Li-Fraumeni, os pacientes com metástase ao diagnóstico ou até mesmo selecionando os casos de SP e LM considerados mais característicos pela análise histopatológica (24). Em nenhuma das análises realizadas foi possível encontrar um grupo formado apenas por um dos subtipos tumorais. Interessante notar que o grupo formado pela análise não supervisionada dos casos mais característicos resultou em um único grande grupo subdividido em vários pequenos subgrupos, lembrando a forma de uma “escada”, sugerindo que todas as amostras estão bastante relacionadas entre si.

Diversos relatos na literatura mostram a grande semelhança entre os SP e outros SPM de células alongadas, principalmente os LM. Desde a década de 1980 começaram a surgir estudos que questionavam, ainda de maneira tênue, o SP como entidade diagnóstica. A partir do começo dos anos 1990 este questionamento aumentou drasticamente devido à publicação de um artigo por FLETCHER et al. (1992), que foram os primeiros a questionar se os SP eram uma entidade diagnóstica distinta. Em uma série de sarcomas pleomórficos, SCHÜRCH et al. (1996) demonstraram que a maioria dos casos que exibe diferenciação miogênica (LM ou rabdomiossarcoma) frequentemente apresenta características histopatológicas semelhantes àsquelas presentes nos SP com morfologia estoriforme-pleomórfica. Resultados obtidos por CGH (hibridação genômica comparativa) cromossômico mostraram que as alterações genômicas eram similares entre SP e LM (FORUS et al.

1995; PARENTE et al. 1999; DERRÉ et al. 2001). Desde o estudo pioneiro de FORUS et al. (1995), que utilizaram pela primeira vez a técnica de CGH para analisar SP e LM, até o recente estudo publicado por LARRAMENDY et al. (2008), todos mostraram grandes semelhanças do padrão de alterações genômicas entre os SP e LM.

LARRAMENDY et al. (2008) avaliaram 102 SP e 82 LM, e verificaram que a distribuição do padrão de alterações genômicas não era capaz de diferenciar os SP dos LM, sugerindo que a maioria dos SP poderia representar o último estado de progressão dos LM. Análises de CGH sugeriram que os SP poderiam ser incluídos em dois subgrupos. Um caracterizado por alterações genômicas complexas que são muito semelhantes às encontradas em LM, ou seja, uma grande proporção dos SP são LM que foram classificados como SP quando indiferenciados. O outro grupo de SP, caracterizado por ganhos recorrentes ou amplificações da região 12q13-q15, e localizados preferencialmente no abdomen e retroperitônio, poderiam corresponder a lipossarcomas (CHIBON et al. 2002; CHIBON et al. 2003). Outros estudos de CGH cromossômico também encontraram grandes semelhanças no padrão de ganhos e perdas genômicas entre lipossarcomas e SP, o que reforça a teoria de que estes dois tipos de sarcomas podem formar uma única entidade diagnóstica (COINDRE et al. 2003; COINDRE et al. 2004; IDBAIH et al. 2005).

O presente estudo apresentou três casos de SP (SP6T, SP11T E SP12T), cujo diagnóstico não descarta a possibilidade de serem LP desdiferenciados ou LP mixóide. As amostras de SP que poderiam corresponder a LP desdiferenciados (SP11T e SP12T) se situaram no retroperitônio e fizeram parte do primeiro *cluster* em análise não supervisionada de todos os casos.

Na literatura, os estudos de perfis de expressão gênica também não foram capazes de separar os SP de LP e principalmente de LM. Um dos primeiros estudos para avaliar o padrão de expressão gênica de SP e LM foi realizado por NIELSEN et al. (2002). Os autores investigaram 41 SPM utilizando uma plataforma de cDNA microarray de 5K. A análise de agrupamentos mostrou que os sarcomas sinoviais, GIST e tumores neurais apresentaram padrões de expressão gênica característicos, enquanto os SP se agruparam juntamente com LM e LP, sem formar entidades distintas. De maneira semelhante, NAGAYAMA et al. (2002) estudaram 47 SPM em plataforma de cDNA microarray de 23K e também não conseguiram diferenciar os SP, LM e LP em análise de agrupamento hierárquico. BAIRD et al. (2005) analisaram 181 amostras, representando 16 tipos de sarcomas, utilizando a plataforma de cDNA-array de 12k e verificaram que os SP não formavam *clusters* distintos, mas se agrupavam juntamente com tumores pobremente diferenciados, como LP desdiferenciado e pleomórfico e tumores da bainha do nervo periférico e alguns leiomiossarcomas. Os LM formaram o grupo adjacente, em análise não supervisionada, ao *cluster* formado pelos SP e demais tumores pobremente diferenciados.

DETWILLER et al. (2005) utilizaram plataforma da Affymetrix de 22K para estudar o padrão de expressão gênica de 38 sarcomas humanos e 14 tecidos normais. A análise de agrupamento hierárquico resultou em três grupos: um formado por 13 tecidos normais e um lipossarcoma, um formado por amostras de GIST, fibrossarcoma, sarcoma sinovial e lipossarcoma e um terceiro grupo composto de amostras de SP, LM, lipossarcomas, fibrossarcoma e uma amostra de pele. HENDERSON et al. (2005) analisaram o padrão de expressão gênica de 96 amostras

de sarcomas e mostraram que os LM e SP apresentaram a maior diversidade genética que se sobrepõe a outros tipos tumorais e nos quais não foi possível encontrar marcadores distintos. NAKAYAMA et al. (2007) utilizaram a tecnologia de *oligorrays* para analisar a expressão gênica de 105 amostras provenientes de 10 tipos de tumores de partes moles. Os SP apresentaram similaridades no padrão de expressão gênica com outros tumores, como LM, LP, fibrossarcomas, mixofibrossarcomas e tumores malignos da bainha do nervo periférico. Amostras destes cinco tipos de sarcomas puderam ser classificados em *clusters* únicos baseados nos perfis de expressão gênica após a exclusão dos SP. Três dos 21 SP apresentaram grande similaridade com um dos cinco tipos de sarcomas tanto do ponto de vista molecular como histológico. Por outro lado, os 18 casos restantes de SP apresentaram pouca ou nenhuma semelhança histológica com os outros tipos tumorais, porém, 12 amostras tiveram similaridade moderada em termos de expressão gênica.

Estudos mais recentes também não foram capazes de diferenciar os SP dos LM. CARNEIRO et al. (2009), utilizaram uma plataforma contendo 27649 sequências de cDNAs para comparar especificamente o padrão de expressão gênica entre 18 amostras de LM e 31 amostras de SP. Os autores também utilizaram uma plataforma de BAC arrays de 32K para avaliar a variação do número de cópias de DNA. A análise de clusterização hierárquica não supervisionada, tanto para o padrão de expressão gênica como para o perfil de variação do número de cópias de DNA, não diferenciou estes dois tipos tumorais. Utilizando o software SAM, foram identificados nove genes (*TAGL3*, *D4S234E*, *KIAA1729*, *PDLIM5*, *TEAD3*, *TPM2*, *ALDH1B1*, *TRDMT1* e *DHODH*) capazes de diferenciar os LM de SP, porém, testes

estatísticos posteriores mostraram que estes não conseguiam classificar corretamente estas neoplasias.

GIBAULT et al. (2011) analisaram 160 amostras de SPM, com genótipo complexo, por análise integrada de CGH-array e perfil de expressão gênica em larga escala. O teste hierárquico não supervisionado dividiu as amostras em cinco grandes grupos. O *cluster* A foi composto quase que exclusivamente por amostras de LM bem diferenciados. O *cluster* B foi formado por poucas amostras de SP e LM, que apresentaram perfis de alterações genômicas (CGH-array) relativamente simples, e foram caracterizadas pela diminuição de expressão de genes relacionados com a reação inflamatória e o controle da mitose. O *cluster* C foi composto por SP e LM pouco diferenciados e de alto grau de malignidade. Estes tumores apresentaram expressão aumentada em genes envolvidos com a adesão e proliferação celular e com a matriz extracelular. Entre 40 e 50% dos casos apresentaram perda nos cromossomos 1q, 2q e 13. O grupo D foi composto por sarcomas de extremidades, envolvendo SP, mixofibrossarcomas e LP pleomórficos e se associaram com a expressão aumentada de genes que codificam os constituintes estruturais da matriz extracelular. Mais de 55% dos casos exibiram perdas nas regiões 13q e 16q. Por último, o grupo E correspondia a sarcomas de extremidades de alto grau histológico, a maioria indiferenciado, caracterizado pela expressão aumentada de genes relacionados com a reação inflamatória, assim como do gene *DKK1* em mais de 70% dos casos.

Baseados nos estudos descritos na literatura e com o objetivo de identificar se há subgrupos de tumores caracterizados por alterações genéticas particulares foram realizadas análises supervisionadas. Inicialmente as análises englobaram todas as

amostras e depois foram excluídos os casos com a Síndrome de Li-Fraumeni e os casos com metástase ao diagnóstico. Finalmente realizou-se uma análise considerando os 24 casos de SP e LM mais característicos quanto à morfologia e imunoistoquímica. Novamente, nenhuma destas análises foi capaz de separar completamente estas duas neoplasias. Este tipo de análise foi capaz de separar um grupo formado apenas de LM, porém, no grupo de SP foram observados seis LM (LM3T, LM4T, LM10T, LM19T, LM20T e LM-02), se consideradas todas as amostras. Estes tumores foram classificados como LM baseados nas características morfológicas e nos resultados da análise imunoistoquímica. A análise para marcadores imunoistoquímicos não foi realizada apenas no caso LM3T, porém, este caso apresentou uma morfologia característica. Por outro lado, o caso LM20T apresentou um desafio maior devido à morfologia e ao padrão de imunomarcadores testados que foram pouco característicos. Estas mesmas amostras de LM também se agruparam com os SP em análises supervisionadas posteriores. Mesmo se considerada apenas a análise supervisionada dos casos típicos de SP e LM, observou-se que as amostras LM4T e LM10T se agruparam com os SP.

Na análise supervisionada para todos os casos, as amostras de LM que se agruparam com os SP foram todas de extremidades, enquanto as 10 amostras de LM de retroperitônio fizeram parte do *cluster* único de LM. Este teste gerou 390 genes diferencialmente expressos apenas em LM quando comparado ao grupo formado por SP e LM, sendo 16 com expressão diminuída e 374 com expressão aumentada. As análises do KEGG e do GO (*Gene Ontology*) revelaram que a maioria dos genes com expressão aumentada estava associada com a função muscular, enquanto os genes com expressão diminuída se associavam com genes de resposta ao stress e hipóxia.

Muitos genes com função muscular apresentaram um ligeiro aumento de expressão no grupo formado apenas de LM (*fold change* < 1). Entre estes genes estão incluídos *ACTA1*, *ACTA2*, *ACTG2*, *CALD1*, *CNN1*, *LMOD1*, *MYL6*, *MYLK*, *MYOCD*, *SMTN*, *TAGLN* e *TPM2*.

Na literatura há relatos descrevendo subgrupos de LM ou SP que se agrupam em *clusters* únicos. Em LM, o padrão de expressão gênica diferenciado diz respeito ao aumento de expressão de genes relacionados com função muscular. NIELSEN et al. (2002) identificaram um grupo de LM composto de seis amostras com expressão diferenciada em 24 genes, relacionados principalmente com a estrutura e função muscular, incluindo *ACTG2*, *MYH11*, *MYPT2*, *MYLK*, *LMOD1*, *ACTA2*, *MYRL2*, *SGCA* e *SLAP*. Por outro lado, LEE et al. (2003) identificaram um subgrupo de três amostras de SP, caracterizado por 21 genes diferencialmente expressos, relacionados principalmente com metabolismo (*CTPS*, *SYK* e *ACAT2*) e resposta ao estresse oxidativo (*HSPE1* e *HIF1A*). Outro subgrupo de nove SP foi identificado por SEGAL et al. (2003), caracterizado pela expressão diferencial de *GFPT2*, *EMP3*, *ANXA1* e *FTSJ1*. Os mesmo autores encontraram um subgrupo de cinco LM caracterizados pela expressão aumentada dos genes *TPM2*, *CNN1*, *LOMD1*, *MYLK*, *LAMA3*, *MYH11* e *TPM1*. BAIRD et al. (2005) identificaram dois subgrupos de SP; um associado com os genes de perfil muscular (*miosina X*, *sarcoglicano B* e *tenascina C*) e o outro relacionado com genes de resposta imune (*HEM1*, *MX1*, *DAP10*, *PLCG2* e *FOLR3*). Em SPM, DETWILLER et al. (2005) examinaram o padrão de expressão gênica de 38 sarcomas e 14 tecidos normais e demonstraram que as amostras tumorais apresentavam um padrão distinto de expressão gênica, composto por 107 genes relacionados com hipóxia, que permitia distinguir estas

neoplasias dos tecidos normais. Entre os vários genes com expressão aumentada pode-se citar o *HIF-1a*. As análises supervisionadas da presente casuística mostraram que o grupo formado apenas de LM apresentou expressão aumentada de vários genes relacionados com a diferenciação de função muscular e que foram descritos em outros estudos, como *ACTG2*, *MYLK*, *LMOD1*, *ACTA2*, *TPM2* e *CNN1*. Por outro lado, o grupo formado por todos os SP e os restantes LM apresentou expressão diminuída dos genes de diferenciação muscular e foi melhor caracterizado por um conjunto de 16 genes com expressão aumentada e que se relacionaram com o metabolismo e transporte transmembrana (*TIMM13* e *SLC39A7*), resposta ao stress oxidativo (*PYCR1*), regulação da transcrição (*HMX1*) e atividade de oxiredutase (*HMOX1*).

Um estudo recente também identificou grupos diferenciados de LM. PÉROT et al. (2009) conseguiram separar, por análise não supervisionada, um conjunto de 19 LM e 17 SP. A única exceção foi um caso de LM que se agrupou com os SP. Vale ressaltar que o diagnóstico destes casos seguiu critérios rígidos e utilizou amostras características para os dois tipos tumorais, tanto do ponto de vista morfológico como imunoistoquímico. Um caso foi considerado LM bem diferenciado quando apresentou morfologia característica e marcação positiva para quatro marcadores de músculo liso em ensaios de imunoistoquímica (SMA, calponina, transgelina e caldesmon). Por outro lado, um caso foi considerado SP por apresentar morfologia característica e negatividade para um ou mais marcadores musculares utilizados. Este estudo mostrou, por análises de CGH-array, FISH e perfil de expressão gênica em larga escala, que o gene da miocardina (*MYOCD*) estava altamente amplificado e com expressão aumentada nos LM retroperitoneais humanos. A inativação deste gene

pela técnica de shRNA ocasionou uma redução dramática da diferenciação muscular lisa e reduziu fortemente a migração celular. Por outro lado, na presença de expressão de *MYOCD* (em três linhagens de SP e uma linhagem de LP) ocorria forte diferenciação da musculatura lisa e um aumento da habilidade migratória das células. Estes resultados demonstraram que os LM retroperitoniais dependem da amplificação e aumento de expressão de *MYOCD*. Os dados sugeriram que nos LM bem diferenciados, o fenótipo de músculo liso pode ser resultado da aquisição de uma alteração genômica, portanto, estes tumores não necessariamente derivam de células inicialmente relacionadas com a diferenciação de músculo liso.

Mais recentemente, GIBault et al. (2011) analisaram o transcriptoma e o genoma de uma série de 160 amostras de sarcomas com genótipo complexo. A análise não supervisionada permitiu a detecção de cinco *clusters*. Com exceção de um SP, o primeiro *cluster* foi composto apenas LM bem diferenciados (27 casos), localizados principalmente no retroperitônio (64% dos casos). Estes tumores apresentaram expressão aumentada de genes relacionados com a organização do citoesqueleto e diferenciação do músculo liso. Em particular, estas neoplasias apresentam expressão aumentada de *MYOCD*. Em nosso estudo foi possível verificar a expressão aumentada de *MYOCD* no grupo formado unicamente por 16 LM, sendo a maioria composto por LM de retroperitônio (10). Interessante notar que os demais 6 LM que se agruparam com os SP e não apresentaram a expressão aumentada deste gene eram todos de extremidades. Estes resultados concordam com os dados apresentados pelo grupo francês de sarcoma, que mostrou a importância deste gene para a diferenciação de músculo liso, principalmente nos LM retroperitoniais (PÉROT et al. 2009; GIBault et al. 2011).

BECK et al. (2010a) combinaram os resultados do perfil de expressão gênica, CGH-array e imunoistoquímica para caracterizar os subtipos moleculares de 51 LM de partes moles. Análises de expressão gênica e de clusterização não supervisionada mostraram três grupos de LM. O primeiro grupo foi caracterizado por genes envolvidos com a diferenciação e função muscular, além de genes codificadores de fosfoproteínas e proteínas quinases (incluindo *CALD1*, *SLMAP*, *DMD*, *ACTG2*, *CFL2* e *MYLK*). Por outro lado, os genes expressos no segundo grupo estavam relacionados com o metabolismo de proteínas e regulação do ciclo celular; enquanto no terceiro grupo, os genes estavam associados com proteínas extracelulares e de resposta a lesões, assim como ao desenvolvimento de órgãos. A análise de CGH-array em 20 casos de LM mostrou que os subtipos moleculares, definidos pela análise de expressão gênica, apresentavam alterações genéticas distintas. Os tumores do primeiro grupo mostraram um aumento significativo na variação do número de cópias de DNA, principalmente perdas em 16q24 e 1p36. Ensaio de imunoistoquímica identificaram cinco marcadores no primeiro grupo de LM (“enriquecidos com genes musculares”) que, por sua vez, estavam associados com uma melhora da sobrevida dos pacientes. Como o objetivo principal do nosso estudo foi separar molecularmente os LM dos SP, não realizamos análises supervisionadas ou não supervisionadas específicas apenas para os casos de LM, porém, a pesquisa usando a base de dados *Gene Ontology* (GO), após análise supervisionada de todos os casos, revelou que o grupo formado apenas de LM apresentou genes diferencialmente expressos referentes à diferenciação de função muscular, metabolismo de proteínas e desenvolvimento de órgãos, sugerindo que a presente

casuística de LM pode conter os três subgrupos moleculares de LM descritos por BECK et al. (2010a).

As análises realizadas não permitiram separar os 44 casos de sarcomas quanto a sua localização anatômica. Nem mesmo os três casos de sarcomas localizados no tronco (LM5T, SP14T e SP16T) ou de cabeça e pescoço (SP7T e SP8T) ficaram agrupados em um único *cluster*. Os sarcomas retroperitoneais (19 casos) se distribuíram pelos quatro grupos, principalmente no segundo *cluster*. Já os sarcomas de extremidades (21 casos) estiveram presentes em todos os grupos, com preferência no terceiro (nove casos) e quarto (cinco casos) grupos. A análise supervisionada dos casos referentes aos sítios anatômicos também não foi capaz de formar *clusters* únicos. Nesta análise os sarcomas de tronco se relacionaram com os sarcomas de retroperitônio, enquanto os sarcomas de cabeça e pescoço se relacionaram com os sarcomas de extremidades. Os casos que desenvolveram metástase ou recidivas durante o seguimento clínico não tiveram relação específica com nenhum dos *clusters* obtidos da análise não supervisionada e se distribuíram por todos grupos. Os dados apresentados concordam com os descritos por BAIRD et al. (2005). Os autores utilizaram análises não supervisionadas na tentativa de associar algumas características clínicas com o padrão de expressão gênica em SP e LM, porém, em nenhum dos casos foi possível distinguir o sítio anatômico, grau tumoral e lesões metastáticas.

REN et al. (2003) foram os primeiros a sugerir que os LM poderiam ser separados por diferenças nos padrões de expressão gênica segundo a diferenciação e a agressividade dos tumores. Os autores identificaram 92 genes que distinguiam LM de baixo grau, alto grau e metastáticos. LEE et al. (2004) realizaram uma análise

hierárquica supervisionada e identificaram 355 genes diferencialmente expressos entre 20 tumores primários e 7 sarcomas metastáticos, cujo perfil de expressão indicava o desenvolvimento de metástases em LM primários. A maioria dos genes do perfil metastático codificava proteínas envolvidas no ciclo celular (CDC27 e CDK2AP1), transdução de sinal (IFNAR2 e PAK2) e apoptose (BCL2A1). No presente estudo não foi realizada uma análise considerando os casos metastáticos (14) e não metastáticos (30) pelo pequeno número de casos, porém, a análise não supervisionada considerando todos os casos não revelou nenhum perfil genético associado a metástase ou recidivas.

LEE et al. (2008) avaliaram o perfil de expressão gênica em uma série de LM e identificaram níveis elevados de expressão de genes associados a macrófagos (*CD163* e *CD68*), sendo confirmado por ensaios de imunoistoquímica em uma plataforma de *tissue microarray* (TMA). As análises revelaram que o aumento de macrófagos nos sarcomas está associado a um pior prognóstico em LM não ginecológicos. Adicionalmente, ESPINOSA et al. (2009) mostraram que a expressão do *CSF1*, citocina que induz a proliferação e diferenciação de macrófagos e monócitos, e das proteínas de resposta a este fator (*CD163*, *FCGR3a* e *CTSL-1*) estão correlacionadas a um pior prognóstico em LM ginecológicos e não ginecológicos. Neste estudo foi detectada a expressão aumentada dos genes relacionados a macrófagos *CD9* e *CD671494* somente nos LM. Entre os pacientes com LM, seis morreram como consequência da doença, cinco por metástases e três menos de um ano após a cirurgia de remoção do tumor. Na tentativa de caracterizar melhor e refinar a análise de expressão gênica, selecionou-se os genes com $-1 > fold\ change > 0,7$ entre os 390 genes obtidos por análise supervisionada dos 44 sarcomas e

que separou um grupo composto unicamente por 16 LM de outro compreendido pelos 22 SP e seis LM. Esta abordagem resultou em 211 genes com expressão aumentada e 16 genes com expressão diminuída os quais foram submetidos a análise de redes, vias gênicas, doenças relacionadas e função celular ou molecular com a utilização do *software Ingenuity Pathways Analysis* (IPA) (Ingenuity Systems). Os scores para cada item foram calculados pelo teste exato de Fisher e são iguais a $-\log$ (valor de P).

Entre 16 LM, os genes com expressão diminuída e que apresentaram os maiores *fold change* estão incluídos o *DUXAP10* (um gene da família de homeobox DUXA que regula aspectos essenciais do desenvolvimento celular) e o *HMOX1* (codifica uma enzima essencial no catabolismo do grupo heme). Por outro lado, entre os genes com expressão aumentada e com maiores *fold change* se encontrou o *HDAC5* (codifica histonas importantes na regulação da transcrição e na progressão do ciclo celular) o *MTA3* (associado a metástase) e os genes *BCL2L2* e *BAD*, que codificam proteínas conhecidas que regulam os processos de apoptose e morte celular. Esta análise também resultou na identificação de nove redes gênicas com scores significativos (*software Ingenuity Pathways Analysis, Ingenuity Systems*) sendo quatro principais. Três destas redes gênicas apresentaram genes associados com o desenvolvimento e função muscular, alguns específicos de músculo liso, demonstrando que estes genes além de conferir a diferenciação muscular aos LM (facilitando o diagnóstico), também regulam importantes vias e funções celulares. A principal rede gênica alterada está relacionada a doenças do sistema reprodutivo, desenvolvimento e morte celular. Os genes *ACTN3*, *LPP* e *SMTN* associados a função muscular se relacionaram nesta rede gênica. A *LPP* foi identificada como

uma proteína altamente expressa em tecidos de músculo liso (GORENNE et al. 2003, GORENNE et al. 2006). Por outro lado, a SMTN foi descrita inicialmente por VAN DER LOOP et al. (1997) como sendo uma proteína do citoesqueleto expressa unicamente em células musculares lisas contráteis. Estudos imunoistoquímicos mais recentes sugerem que a proteína *smootelina* é um marcador específico para neoplasias de músculo liso do trato gastro-intestinal e pode ser utilizada no estadiamento do câncer de bexiga (COCO et al. 2009; COUNCIL e HAMEED 2009; BOVIO et al. 2010). Outros genes alterados (expressão aumentada) encontrados nesta rede foram o *THRA* que codifica um receptor hormonal tireoideano e cuja mutação está associada ao câncer de fígado (CHAN e PRIVALSKY 2010); o *MUC1* que codifica uma fosfoproteína glicosilada ligada à membrana cuja expressão aumentada está associada ao câncer de mama e a um pior prognóstico no câncer gástrico (LACUNZA et al. 2010; RETTERSPLITZ et al. 2010); e o gene da família das *heat shock proteins HSPB1*, que está associado com vários tipos de câncer, como câncer de cabeça e pescoço (ZHU et al. 2010). O único gene com expressão diminuída no grupo de LM presente nesta rede foi o *HMOX1*, cujo aumento de expressão foi detectado em uma série de tumores (incluindo adenocarcinomas, hepatomas, glioblastomas, leucemias mielóides e sarcomas de Kaposi). Este gene atua no processo de apoptose, proliferação, invasão e metástase das células tumorais (JOZKOWICZ et al. 2007).

A segunda rede compreende genes relacionados com o ciclo, morfologia e morte celular. Nesta rede se encontram genes reguladores dos processos de apoptose e morte celular, como o *BCL2L2* e *BAD*, os quais estão associados a vários tipos de câncer como carcinoma de fígado e de pulmão (CRAWFORD et al. 2009;

GALMICHE et al. 2010). Outros genes importantes e com aumento de expressão foram o *SRC* e *ILK*, os quais regulam uma grande variedade de funções celulares. Estudo recente relatou o aumento de expressão da proteína SRC e sua relação com a proliferação e progressão de neoplasias malignas da pele (melanoma maligno, carcinoma de células escamosas e carcinoma de células basais) (LEE et al. 2010). A expressão elevada da ILK foi relatada em uma variedade de cânceres e correlacionada com características de pior prognóstico. PONTIER et al. (2010) relataram correlação entre o aumento de expressão do gene *ILK* com a progressão dos tumores de mama HER2 positivos. Outros estudos associaram o aumento de expressão da ILK com a resistência a drogas quimioterápicas em adenocarcinoma ductal de pâncreas (SCHAEFFER et al. 2010) e com uma menor sobrevida em pacientes com astrocitoma (LI J et al. 2010). Nesta rede, o único gene relacionado com função muscular que apresentou diferença de expressão (aumento) foi o *ACTN1*. SLATER et al. (2007) detectaram a perda total da proteína ACTN1 (*alpha-actinin*) no carcinoma de endométrio e endometriose, indicando uma similaridade entre estas duas categorias no que diz respeito à função das células epiteliais e adesão.

A terceira rede apresentou 10 genes alterados envolvidos em processos de replicação, recombinação e reparo a danos no DNA, além do ciclo e morte celular. Interessante notar que nesta rede se encontrou aumento de expressão dos supressores tumorais *RBBP7* e *TP53INP1* no grupo formado apenas de LM. Por outro lado, também se encontrou aumento de expressão no *MTA3*, um gene associado com metástase. Resultados semelhantes foram descritos em coriocarcinoma, em oposição, expressão diminuída foi relatada em adenocarcinomas de endométrio (BRÜNING et al. 2009, BRÜNING et al. 2010). O gene *CHEK1* encontrado nesta rede foi relatado

como correlacionado ao desenvolvimento do câncer de mama e de cólon (JARDIM et al. 2009; SOLYOM et al. 2010; TONIC et al. 2010). Esta foi a única rede que não apresentou genes associados com a função muscular.

Na quarta rede gênica estão representados genes associados com a movimentação celular, tráfego de células imunes e com o desenvolvimento e funcionamento do sistema hematopoiético. Nesta rede foram verificados genes importantes associados a diferenciação e função muscular, como o *MYOCD*, *MEF2C*, *SRF* e *PPP1R12A* (todos com aumento de expressão no grupo formado apenas de LM). PÉROT et al. (2009) e GIBault et al. (2011) mostraram a importância da expressão do gene *MYOCD* na diferenciação muscular dos LM retroperitoniais. A proteína MYOCD é um cofator transcricional da proteína SRF, o qual estava também presente nesta via (WANG Z et al. 2003; WANG et al. 2004). Estudo mais recente sugere a utilização da miocardina na classificação de tumores mesenquimais do útero (LIANG et al. 2010). Estudos recentes também sugerem a relação dos genes *SRF* e *PPP1R12A* com a progressão do carcinoma hepatocelular (BAI et al. 2008; KWON et al. 2010).

Os genes com função muscular também se associaram às principais categorias, obtidas pela análise dos dados utilizando o *software Ingenuity*. Os genes associados a função e desenvolvimento do sistema esquelético e muscular mostraram a relação entre os genes *MYL9*, *MYL6*, *MYOCD*, *SRF*, *CALD1*, *MEF2C*, *SMTN* e *HDAC5* assim como os associados a morfologia dos tecidos, que foi composta pelos genes *MYL9*, *HMOX1*, *MYOCD*, *SRF* e *SMTN*.

Muitos genes também presentes nas quatro principais redes gênicas descritas acima também estavam presentes nas quatro vias canônicas de maiores *scores*. O

genes *MYL6* e *MYL9* fazem parte de todas as quatro vias canônicas principais. Estes genes codificam uma das cadeias leves da miosina e regulam a contração muscular. Outros genes com função muscular, como o *CALM1*, *MEF2C* e *PPP1R12A* também fazem parte de duas ou mais vias canônicas.

Para complementar as análises de expressão gênica entre os LM e SP, foi utilizado o teste *SAM* de duas classes, com $FDR=0$, para fazer uma clusterização hierárquica pareada entre quatro sarcomas (SP2T, SP3T, SP6T e LM4T) e seus respectivos tecidos adjacentes normais. Neste caso se obteve 40 genes diferencialmente expressos que separavam o grupo tumoral do grupo normal. Os genes com expressão aumentada no grupo tumoral tiveram um *fold change* variando de 2 a 34, enquanto os genes com expressão diminuída apresentaram valores de *fold change* elevados variando de 16 a 74. Nenhum destes 40 genes se mostrou alterado na análise supervisionada entre os casos de SP e LM. As anotações das vias do KEGG revelaram que os genes com expressão aumentada nos tumores estavam associados às vias de sinalização do *TP53* e ao ciclo celular, enquanto os genes com expressão diminuída se associaram a vias metabólicas em geral.

Os quatro genes que apresentaram expressão diminuída e os maiores *fold change* foram o *ALDH1L1*, *ECHDC3*, *SORBS1* e *ARHGEF4*. Uma revisão da literatura revela dois relatos principais que relacionam esses genes com algum tipo de câncer. LIN et al. (2001) associaram o gene *SORBS1* com o desenvolvimento de uma linhagem de hepatoma, enquanto KAWASAKI et al. (2003) relacionaram o gene *ARHGEF4* com o potencial migratório de células do tumor de cólon. Por outro lado, os genes com expressão aumentada no grupo tumoral e que apresentaram os maiores *fold change* foram *ANLN*, *PBK*, *TYMS*, *MXRA8*, *FN1* e *TOP2A*. Estudos

recentes mostraram que o gene *ANLN* se encontra altamente expresso em carcinomas de pâncreas e o gene *TYMS* está associado com a eficácia do bevacizumab no tratamento de pacientes com câncer colorretal metastático (OLAKOWSKI et al., 2009; PANDER et al., 2010). WAALKES et al. (2010) correlacionaram a expressão do gene *FN1* com um estágio avançado da doença em pacientes com câncer renal. O gene *PBK* foi sugerido como marcador diagnóstico entre colangiocarcinomas e carcinomas hepatocelulares, uma vez que está frequentemente expresso em colangiocarcinomas e ausente em hepatocarcinomas. Além disso, a expressão diminuída deste gene em colangiocarcinomas foi associado a sobrevida diminuída dos pacientes (HE et al., 2009). Recentemente, LEMY et al. (2011) avaliaram 100 amostras de câncer de mama e verificaram que os pacientes HER2-positivos/receptores hormonais negativos e sem amplificação do *TOP2A* apresentam os piores prognósticos. Outro estudo correlacionou a expressão do gene *TOP2A* com uma melhor sobrevida livre de doença para os pacientes operados de câncer pulmonar de células não-pequenas e que receberam quimioterapia adjuvante (YAN et al., 2010).

Os estudos de *arrays* de expressão em sarcomas apresentam alguns problemas, principalmente relacionados à raridade de amostras e a dificuldade de classificar os mais de 50 tipos histológicos reconhecidos. Por isso, muitos estudos incluem poucas amostras referentes a um subtipo histológico (3 a 10 casos) e não fazem distinção entre localizações anatômicas. Alguns relatos consideraram um grande número de entidades diagnósticas como uma única categoria (WUNDER et al. 2007; VERWEIJ e BAKER 2010; WEST 2010). Estudos de SHMULEVICH et al. (2002) e FRANCIS et al. (2005) demonstraram que a heterogeneidade tumoral dos sarcomas pode ter uma grande influência nos resultados de análises quando

poucas amostras são consideradas, limitando a validade dos achados. No presente trabalho foi utilizado um número elevado de amostras (44) de apenas duas entidades tumorais: 22 SP e 22 LM. Para o nosso conhecimento, há apenas um estudo que utilizou número maior de casos: 31 SP e 18 LM (CARNEIRO et al. 2009). Os autores utilizaram apenas amostras de extremidades, enquanto aqui foram sarcomas de quatro localizações anatômicas (extremidades, retroperitônio, tronco e região da cabeça e pescoço). As análises supervisionadas e não supervisionadas não mostraram correlação entre os perfis de expressão e os sítios anatômicos.

A comparação dos resultados de expressão gênica com dados da literatura também apresenta dificuldade, pois diferentes autores utilizam diferentes plataformas e estratégias de análise assim como diferentes RNAs de referência. Inicialmente se utilizavam plataformas de expressão *customizadas* com livrarias de cDNA, atualmente existem plataformas comerciais formadas por *oligoarrays*, principalmente da Agilent e Affymetrix contendo um grande número de sequências que são avaliadas em um único experimento. Os *softwares* de análises foram sendo aperfeiçoados com o passar dos anos e atualmente os pesquisadores dispõem de uma gama de ferramentas de bioinformática. A utilização de RNAs de referência diferentes também pode levar a resultados completamente distintos. Uma revisão detalhada da literatura revela a utilização de diferentes tipos de RNA de referência em experimentos de expressão gênica de sarcomas ou até mesmo a sua não utilização, como acontece nos experimentos em plataformas da Affymetrix. Por exemplo, NAGAYAMA et al. (2002) utilizaram uma linhagem de células tronco mesenquimais como RNA de referência; NIELSEN et al. (2002) utilizaram um *pool* composto por 11 linhagens celulares e CARNEIRO et al. (2009) fizeram uso do

RNA de referência comercial (Stratagene, La Jolla, CA, USA). Por outro lado, existem vários trabalhos que não informam ou não deixam claro o RNA de referência utilizado para comparação, como acontece nos relatos de PRICE et al. (2007) e YLIPAA et al. (2010). No presente estudo foi utilizado um *pool* de RNA obtido de 15 linhagens celulares como referência. Adicionalmente foi possível comparar o perfil de expressão gênica de quatro sarcomas e seus respectivos tecidos adjacentes normais. Antes de realizar as comparações todos os tumores e tecidos adjacentes normais foram hibridados com o RNA de referência. Na análise comparativa dos quatro tecidos adjacentes normais e seus respectivos sarcomas era esperado a presença de alguns genes de diferenciação muscular que separassem os grupos, o que não aconteceu, provavelmente devido ao pequeno número de amostras analisadas.

Em resumo, neste estudo foi analisado o padrão de expressão gênica de 22 SP e 22 LM, na tentativa de separar molecularmente estas neoplasias. Atualmente os SP não são considerados uma entidade diagnóstica verdadeira, mas uma coletânea de tumores pobremente diferenciados (LM, lipossarcomas, rabdomiossarcomas) nos quais não é possível identificar uma linhagem celular específica (THWAY 2009). O diagnóstico destas neoplasias representa um verdadeiro desafio para os patologistas e muitas vezes apenas a análise da morfologia não é suficiente para identificar estes tumores, sendo necessário a utilização de outras técnicas para confirmar o diagnóstico e excluir outros tipos tumorais (FLETCHER 2006; WEST 2010). As diferentes análises realizadas não foram capazes de diferenciar estas neoplasias e encontrar relação de algum gene ou grupo de genes com características clínico-patológicas. Os SP e os LM apresentaram uma alta similaridade quanto ao padrão de expressão gênica. Análises não supervisionadas, até mesmo excluindo-se os casos controversos,

falharam em agrupar apenas amostras de SP ou de LM. Por outro lado análises supervisionadas foram capazes de separar um grupo de LM do restante de casos de LM e todos os SP. No grupo de LM foi observado o aumento de expressão de vários genes associados com o desenvolvimento e a função muscular, o que está de acordo com a literatura. Análises mais detalhadas pelo *software* Ingenuity revelaram que estes genes também se associavam a importantes redes gênicas, responsáveis por controlar praticamente toda a maquinaria celular. Outros genes associados a câncer também foram encontrados alterados nas principais redes gênicas e vias canônicas. Os resultados aqui apresentados reforçam a teoria de que os SP podem corresponder a LM pouco diferenciados, assim como existem LM que apresentam diferenciação muscular mais evidente que representaria um grupo mais bem caracterizado e diagnosticado, como sugerido por BECK et al. (2010a).

7 CONCLUSÃO

- As análises de expressão gênica global não foram capazes de diferenciar os Sarcomas Pleomórficos dos Leiomiossarcomas;
- Análises supervisionadas identificaram um subgrupo de LM, caracterizados pela expressão diferencial de vários genes associados com o desenvolvimento e função muscular;
- Os genes com função muscular e que caracterizavam o grupo formado apenas de LM, se associavam com redes e vias gênicas importantes que controlam importantes funções celulares, como o reparo a danos do DNA e a morte celular;
- Não foi possível identificar perfis de expressão gênica associados a características clínico-patológicas, como metástase, recidiva ou sítio tumoral.
- Os genes selecionados para análise por RT-qPCR (*BAD*, *SRC*, *SRF* e *TOP2A*) confirmaram os achados dos *arrays* de expressão em LM e SP.

8 REFERÊNCIAS BIBLIOGRÁFICAS

Antonescu CR. Molecular profiling in the diagnosis and treatment of high grade sarcomas. **Ultrastruct Pathol** 2008; 32:37-42.

Bai X, Wu L, Liang T, et al. Overexpression of myocyte enhancer factor 2 and histone hyperacetylation in hepatocellular carcinoma. **J Cancer Res Clin Oncol** 2008; 134:83-91.

Baird K, Davis S, Antonescu CR, et al. Gene expression profiling of human sarcomas: insights into sarcoma biology. **Cancer Res** 2005; 65:9226-35.

Beck AH, Lee CH, Witten DM, et al. Discovery of molecular subtypes in leiomyosarcoma through integrative molecular profiling. **Oncogene** 2010a; 29:845-54.

Beck AH, West RB, van de Rijn M. Gene expression profiling for the investigation of soft tissue sarcoma pathogenesis and the identification of diagnostic, prognostic, and predictive biomarkers. **Virchows Arch** 2010b; 456:141-51.

Blay JY, Le Cesne A. Adjuvant chemotherapy in localized soft tissue sarcomas: still not proven. **Oncologist** 2009; 14:1013-20.

Boccardo E, Manzini Baldi CV, Carvalho AF, et al. Expression of human papillomavirus type 16 E7 oncoprotein alters keratinocytes expression profile in response to tumor necrosis factor-alpha. **Carcinogenesis** 2010; 31:521-31.

Bovée JV, Hogendoorn PC. Molecular pathology of sarcomas: concepts and clinical implications. **Virchows Arch** 2010; 456:193-9.

Bovio IM, Al-Quran SZ, Rosser CJ, Algood CB, Drew PA, Allan RW. Smoothelin immunohistochemistry is a useful adjunct for assessing muscularis propria invasion in bladder carcinoma. **Histopathology** 2010; 56:951-6.

Brennan MF, Samuel S, Robert GM, Brian OS. Sarcomas of the soft tissue and bone. In: De Vita Jr V, Lawrence TS, Rosenberg AS. De Vita, Hellman, Rosenberg's Cancer principles e practice of oncology. Philadelphia: Lippincott Williams and Wilkins; 2008. p.1741-94.

Brooks JJ. The significance of double phenotypic patterns and markers in human sarcomas. A new model of mesenchymal differentiation. **Am J Pathol** 1986; 125:113-23.

Brüning A, Jückstock J, Blankenstein T, Makovitzky J, Kunze S, Mylonas I. The metastasis-associated gene MTA3 is downregulated in advanced endometrioid adenocarcinomas. **Histol Histopathol** 2010; 25:1447-56.

Brüning A, Makovitzky J, Gengelmaier A, Friese K, Mylonas I. The metastasis-associated genes MTA1 and MTA3 are abundantly expressed in human placenta and chorionic carcinoma cells. **Histochem Cell Biol** 2009; 132:33-8.

Carneiro A, Francis P, Bendahl PO, et al. Indistinguishable genomic profiles and shared prognostic markers in undifferentiated pleomorphic sarcoma and leiomyosarcoma: different sides of a single coin? **Lab Invest** 2009; 89:668-75.

Chan IH, Privalsky ML. A conserved lysine in the thyroid hormone receptor-alpha1 DNA-binding domain, mutated in hepatocellular carcinoma, serves as a sensor for transcriptional regulation. **Mol Cancer Res** 2010; 8:15-23.

Chibon F, Lagarde P, Salas S, et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. **Nat Med** 2010; 16:781-7.

Chibon F, Mairal A, Fréneaux P, et al. The RB1 gene is the target of chromosome 13 deletions in malignant fibrous histiocyoma. **Cancer Res** 2000; 60:6339-45.

Chibon F, Mariani O, Derré J, et al. A subgroup of malignant fibrous histiocyomas is associated with genetic changes similar to those of well-differentiated liposarcomas. **Cancer Genet Cytogenet** 2002; 139:24-9.

Chibon F, Mariani O, Mairal A, et al. The use of clustering software for the classification of comparative genomic hybridization data. an analysis of 109 malignant fibrous histiocyomas. **Cancer Genet Cytogenet** 2003; 141:75-8.

Clark MA, Fisher C, Judson I, Thomas JM. Soft-tissue sarcomas in adults. **N Engl J Med** 2005; 353:701-11.

Coco DP, Hirsch MS, Hornick JL. Smoothelin is a specific marker for smooth muscle neoplasms of the gastrointestinal tract. **Am J Surg Pathol** 2009; 33:1795-801.

Coindre JM, Hostein I, Maire G, et al. Inflammatory malignant fibrous histiocyomas and dedifferentiated liposarcomas: histological review, genomic profile, and MDM2 and CDK4 status favour a single entity. **J Pathol** 2004; 203:822-30.

Coindre JM, Mariani O, Chibon F, et al. Most malignant fibrous histiocyomas developed in the retroperitoneum are dedifferentiated liposarcomas: a review of 25 cases initially diagnosed as malignant fibrous histiocyoma. **Mod Pathol** 2003; 16:256-62.

Collini P, Sorensen PH, Patel S, et al. Sarcomas with spindle cell morphology. **Semin Oncol** 2009; 36:324-37.

Conley AP, Trent J, Zhang W. Recent progress in the genomics of soft tissue sarcomas. **Curr Opin Oncol** 2008; 20:395-9.

Council L, Hameed O. Differential expression of immunohistochemical markers in bladder smooth muscle and myofibroblasts, and the potential utility of desmin, smoothelin, and vimentin in staging of bladder carcinoma. **Mod Pathol** 2009; 22:639-50.

Crawford M, Batte K, Yu L, et al. MicroRNA 133B targets pro-survival molecules MCL-1 and BCL2L2 in lung cancer. **Biochem Biophys Res Commun** 2009; 388:483-9.

Daigeler A, Klein-Hitpass L, Stricker I, et al. Malignant fibrous histiocytoma-pleomorphic sarcoma, NOS gene expression, histology, and clinical course. a pilot study. **Langenbecks Arch Surg** 2010; 395:261-75.

Davicioni E, Anderson MJ, Finckenstein FG, et al. Molecular classification of rhabdomyosarcoma-genotypic and phenotypic determinants of diagnosis: a report from the Children's Oncology Group. **Am J Pathol** 2009; 174:550-64.

Davicioni E, Finckenstein FG, Shahbazian V, Buckley JD, Triche TJ, Anderson MJ. Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas. **Cancer Res** 2006; 66:6936-46.

Dehner LP. Malignant fibrous histiocytoma. Nonspecific morphologic pattern, specific pathologic entity, or both? **Arch Pathol Lab Med** 1988; 112:236-7.

Dei Tos AP. Classification of pleomorphic sarcomas: where are we now? **Histopathology** 2006; 48:51-62.

Derré J, Lagacé R, Nicolas A, et al. Leiomyosarcomas and most malignant fibrous histiocytoomas share very similar comparative genomic hybridization imbalances: an analysis of a series of 27 leiomyosarcomas. **Lab Invest** 2001; 81:211-5.

Detwiller KY, Fernando NT, Segal NH, Ryeom SW, D'Amore PA, Yoon SS. Analysis of hypoxia-related gene expression in sarcomas and effect of hypoxia on RNA interference of vascular endothelial cell growth factor A. **Cancer Res** 2005; 65:5881-9.

Deyrup AT, Lee VK, Hill CE et al. Epstein–Barr virus-associated smooth muscle tumors are distinctive mesenchymal tumors reflecting multiple infection events: a clinicopathologic and molecular analysis of 29 tumors from 19 patients. **Am J Surg Pathol** 2006; 30:75-82.

El-Rifai W, Sarlomo-Rikala M, Knuutila S, Miettinen M. DNA copy number changes in development and progression in leiomyosarcomas of soft tissues. **Am J Pathol** 1998; 153:985-90.

Enzinger FM. Malignant fibrous histiocytooma 20 years after Stout. **Am J Surg Pathol** 1986; 10:43-53.

Eriksson M. Histology-driven chemotherapy of soft-tissue sarcoma. **Ann Oncol** 2010; 21:270-6.

Espinosa I, Beck AH, Lee CH, et al. Coordinate expression of colony-stimulating factor-1 and colony-stimulating factor-1-related proteins is associated with poor prognosis in gynecological and nongynecological leiomyosarcoma. **Am J Pathol** 2009; 174:2347-56.

Espinosa I, Lee CH, Kim MK, Rouse BT, et al. A novel monoclonal antibody against DOG1 is a sensitive and specific marker for gastrointestinal stromal tumors. **Am J Surg Pathol** 2008; 32:210-8.

Farshid G, Pradhan M, Goldblum J, Weiss SW. Leiomyosarcoma of somatic soft tissues: a tumor of vascular origin with multivariate analysis of outcome in 42 cases. **Am J Surg Pathol** 2002; 26:14-24.

Fletcher CD. Pleomorphic malignant fibrous histiocytoma: fact or fiction? A critical reappraisal based on 159 tumors diagnosed as pleomorphic sarcoma. **Am J Surg Pathol** 1992; 16:213-28.

Fletcher CD, Unni KK, Mertens F. **Classification of tumours**; pathology and genetics, tumours of soft tissue and bone. WHO Lyon, France: IARC Press; 2002. (World Health Organization classification of tumours; 4).

Fletcher CD. The evolving classification of soft tissue tumours: an update based on the new WHO classification. **Histopathology** 2006; 48:3-12.

Forus A, Weghuis DO, Smeets D, Fodstad O, Myklebost O, van Kessel AG. Comparative genomic hybridization analysis of human sarcomas: I. Occurrence of genomic imbalances and identification of a novel major amplicon at 1q21-q22 in soft tissue sarcomas. **Genes Chromosomes Cancer** 1995; 14:8-14.

Francis P, Fernebro J, Edén P, et al. Intratumor versus intertumor heterogeneity in gene expression profiles of soft-tissue sarcomas. **Genes Chromosomes Cancer** 2005; 43:302-8.

Francis P, Namløs HM, Müller C et al. Diagnostic and prognostic gene expression signatures in 177 soft tissue sarcomas: hypoxia-induced transcription profile signifies metastatic potential. **BMC Genomics** 2007; 14:8:73.

Froehner M, Wirth MP. Etiologic factors in soft tissue sarcomas. **Onkologie** 2001; 24:139-42.

Galmiche A, Ezzoukhry Z, François C, et al. BAD, a proapoptotic member of the BCL2 family, is a potential therapeutic target in hepatocellular carcinoma. **Mol Cancer Res** 2010; 8:1116-25.

Ganjoo K, Jacobs C. Antiangiogenesis agents in the treatment of soft tissue sarcomas. **Cancer** 2010; 116:1177-83.

Gibault L, Pérot G, Chibon F, et al. New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. **J Pathol.** 2011; 223:64-71.

Gomes LI, Esteves GH, Carvalho AF, et al. Expression profile of malignant and nonmalignant lesions of esophagus and stomach: differential activity of functional modules related to inflammation and lipid metabolism. **Cancer Res** 2005; 65:7127-36.

Gorenne I, Jin L, Yoshida T, et al. LPP expression during in vitro smooth muscle differentiation and stent-induced vascular injury. **Circ Res** 2006; 98:378-85.

Gorenne I, Nakamoto RK, Phelps CP, Beckerle MC, Somlyo AV, Somlyo AP. LPP, a LIM protein highly expressed in smooth muscle. **Am J Physiol Cell Physiol** 2003; 285:674-85.

Graadt van Roggen JF, Bovée JV, Morreau J, Hogendoorn PC. Diagnostic and prognostic implications of the unfolding molecular biology of bone and soft tissue tumours. **J Clin Pathol** 1999; 52:481-9.

Guillou L, Aurias A. Soft tissue sarcomas with complex genomic profiles. **Virchows Arch** 2010; 456:201-17.

Guillou L, Coindre JM, Bonichon F, et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. **J Clin Oncol** 1997; 15:350-62.

Hancock JD, Lessnick SL. A transcriptional profiling meta-analysis reveals a core EWS-FLI gene expression signature. **Cell Cycle** 2008; 7:250-6.

He F, Yan Q, Fan L, et al. PBK/TOPK in the differential diagnosis of cholangiocarcinoma from hepatocellular carcinoma and its involvement in prognosis of human cholangiocarcinoma. **Hum Pathol** 2010; 41:415-24.

Heinrich MC, Corless CL, Duensing A, et al. PDGFRA activating mutations in gastrointestinal stromal tumors. **Science** 2003; 299:708-10.

Henderson SR, Guiliano D, Presneau N, et al. A molecular map of mesenchymal tumors. **Genome Biol** 2005; 6:R76.

Hu J, Khanna V, Jones M, Surti U. Genomic alterations in uterine leiomyosarcomas: potential markers for clinical diagnosis and prognosis. **Gen Chrom Cancer** 2001; 31:117-24.

Hu J, Rao UN, Jasani S, Khanna V, Yaw K, Surti U. Loss of DNA copy number of 10q is associated with aggressive behavior of leiomyosarcomas: a comparative genomic hybridization study. **Cancer Genet Cytogenet** 2005; 161:20-7.

Idbaih A, Coindre JM, Derré J, et al. Myxoid malignant fibrous histiocytoma and pleomorphic liposarcoma share very similar genomic imbalances. **Lab Invest** 2005; 85:176-81.

Iwasaki H, Nabeshima K, Nishio J, et al. Pathology of soft-tissue tumors: daily diagnosis, molecular cytogenetics and experimental approach. **Pathol Int** 2009; 59:501-21.

Jain S, Xu R, Prieto VG, Lee P. Molecular classification of soft tissue sarcomas and its clinical applications. **Int J Clin Exp Pathol** 2010; 3:416-28.

Jardim MJ, Wang Q, Furumai R, Wakeman T, Goodman BK, Wang XF. Reduced ATR or Chk1 expression leads to chromosome instability and chemosensitization of mismatch repair-deficient colorectal cancer cells. **Mol Biol Cell** 2009; 20:3801-9.

Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. **CA Cancer J Clin** 2009; 59:225-49.

Jozkowicz A, Was H, Dulak J. Heme oxygenase-1 in tumors: is it a false friend? **Antioxid Redox Signal** 2007; 9:2099-117.

Kawasaki Y, Sato R, Akiyama T. Mutated APC and Asef are involved in the migration of colorectal tumour cells. **Nat Cell Biol** 2003; 5:211-5.

Kenney RJ, Cheney R, Stull MA, Kraybill W. Soft tissue sarcomas: current management and future directions. **Surg Clin North Am** 2009; 89:235-47.

Konstantinopoulos PA, Fountzilas E, Goldsmith JD, et al. Analysis of multiple sarcoma expression datasets: implications for classification, oncogenic pathway activation and chemotherapy resistance. **PLoS One** 2010; 5:9747.

Kwon CY, Kim KR, Choi HN, et al. The role of serum response factor in hepatocellular carcinoma: implications for disease progression. **Int J Oncol** 2010; 37:837-44.

Lacunza E, Baudis M, Colussi AG, Segal-Eiras A, Croce MV, Abba MC. MUC1 oncogene amplification correlates with protein overexpression in invasive breast carcinoma cells. **Cancer Genet Cytogenet** 2010; 201:102-10.

Lamy PJ, Fina F, Bascoul-Mollevi C, et al. Quantification and clinical relevance of gene amplification at chromosome 17q12-q21 in human epidermal growth factor receptor 2-amplified breast cancers. **Breast Cancer Res** 2011; 13:R15.

Larramendy ML, Gentile M, Soloneski S, Knuutila S, Böhling T. Does comparative genomic hybridization reveal distinct differences in DNA copy number sequence patterns between leiomyosarcoma and malignant fibrous histiocytoma? **Cancer Genet Cytogenet** 2008; 187:1-11.

Larramendy ML, Kaur S, Svarvar C, Böhling T, Knuutila S. Gene copy number profiling of soft-tissue leiomyosarcomas by array-comparative genomic hybridization. **Cancer Genet Cytogenet** 2006; 169:94-101.

Larramendy ML, Tarkkanen M, Blomqvist C, et al. Comparative genomic hybridization of malignant fibrous histiocytoma reveals a novel prognostic marker. **Am J Pathol** 1997; 151:1153-61.

Lee CH, Espinosa I, Vrijaldenhoven S, et al. Prognostic significance of macrophage infiltration in leiomyosarcomas. **Clin Cancer Res** 2008; 14:1423-30.

Lee JH, Pyon JK, Kim DW, et al. Elevated c-Src and c-Yes expression in malignant skin cancers. **J Exp Clin Cancer Res** 2010; 29:116.

Lee YF, John M, Edwards S, et al. Molecular classification of synovial sarcomas, leiomyosarcomas and malignant fibrous histiocytomas by gene expression profiling. **Br J Cancer** 2003; 88:510-5.

Lee YF, John M, Falconer A, et al. A gene expression signature associated with metastatic outcome in human leiomyosarcomas. **Cancer Res** 2004; 64:7201-4.

Lehnhardt M, Daigeler A, Homann HH, et al. MFH revisited: outcome after surgical treatment of undifferentiated pleomorphic or not otherwise specified (NOS) sarcomas of the extremities – an analysis of 140 patients. **Langenbecks Arch Surg** 2009; 394:313-20.

Levy B, Mukherjee T, Hirschhorn K. Molecular cytogenetic analysis of uterine leiomyoma and leiomyosarcoma by comparative genomic hybridization. **Cancer Genet Cytogenet** 2000; 121:1-8.

Li J, Zhang H, Wu J, et al. Prognostic significance of integrin-linked kinase1 overexpression in astrocytoma. **Int J Cancer** 2010; 126:1436-44.

Li Q, Hisha H, Takaki T, et al. Transformation potential of bone marrow stromal cells into undifferentiated high-grade pleomorphic sarcoma. **J Cancer Res Clin Oncol** 2010; 136:829-38.

Liang J, Wu L, Xiao H, et al. Use of myocardin in the classification of mesenchymal tumors of the uterus. **Int J Gynecol Pathol** 2010; 29:55-62.

Lin WH, Huang CJ, Liu MW, et al. Cloning, mapping, and characterization of the human sorbin and SH3 domain containing 1 (SORBS1) gene: a protein associated with c-Abl during insulin signaling in the hepatoma cell line Hep3B. **Genomics** 2001; 74:12-20.

Lockhart DJ, Winzeler EA. Genomics, gene expression and DNA arrays. **Nature** 2000; 405:827-36.

Mairal A, Terrier P, Chibon F, Sastre X, Lecesne A, Aurias A. Loss of chromosome 13 is the most frequent genomic imbalance in malignant fibrous histiocytomas. A

comparative genomic hybridization analysis of a series of 30 cases. **Cancer Genet Cytogenet** 1999; 111:134-8.

Matushansky I, Charytonowicz E, Mills J, et al. MFH classification: differentiating undifferentiated pleomorphic sarcoma in the 21st Century. **Expert Rev Anticancer Ther** 2009; 9:1135-44.

Matushansky I, Hernando E, Socci ND, et al. Derivation of sarcomas from mesenchymal stem cells via inactivation of the Wnt pathway. **J Clin Invest** 2007; 117:3248-57.

Mertens F, Fletcher CD, Dal Cin P, et al. Cytogenetic analysis of 46 pleomorphic soft tissue sarcomas and correlation with morphologic and clinical features: a report of the CHAMP Study Group. **Chromos Morphology Genes Chromos Can** 1998; 22:16-25.

Miettinen M, Wang ZF, Lasota J. DOG1 antibody in the differential diagnosis of gastrointestinal stromal tumors: a study of 1840 cases. **Am J Surg Pathol** 2009; 33:1401-8.

Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. **Nat Rev Cancer** 2007; 7:233-45.

Mitelman F, **Mitelman database of chromosome aberrations in cancer**. Available from: <URL:<http://cgap.nci.nih.gov/Chromosomes/Mitelman>> [2010 dez 12]

Nagayama S, Katagiri T, Tsunoda T, et al. Genome-wide analysis of gene expression in synovial sarcomas using a cDNA microarray. **Cancer Res** 2002; 62:5859-66.

Nakayama R, Nemoto T, Takahashi H, et al. Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. **Mod Pathol** 2007; 20:749-59.

[NCCN] National Comprehensive Cancer Network. **Practice guideline in oncology: soft tissue sarcoma**. 2009. Available from: <URL:<http://www.nccn.org>> [2011 jan 12]

Nielsen TO, West RB, Linn SC, et al. Molecular characterisation of soft tissue tumours: a gene expression study. **Lancet** 2002; 359:1301-7.

Nielsen TO, West RB. Translating gene expression into clinical care: sarcomas as a paradigm. **J Clin Oncol** 2010; 28:1796-805.

O'Brien JE, Stout AP. Malignant Fibrous Xanthomas. **Cancer** 1964; 17:1445-55.

Olakowski M, Tyszkiewicz T, Jarzab M, et al. NBL1 and anillin (ANLN) genes over-expression in pancreatic carcinoma. **Folia Histochem Cytobiol** 2009; 47:249-55.

Osuna D, de Alava E. Molecular pathology of sarcomas. **Rev Recent Clin Trials** 2009; 4:12-26.

Otaño-Joos M, Mechttersheimer G, Ohl S, et al. Detection of chromosomal imbalances in leiomyosarcoma by comparative genomic hybridization and interphase cytogenetics. **Cytogenet Cell Genet** 2000; 90:86-92.

Ozzello L, Stout AP, Murray MR. Cultural characteristics of malignant histiocytomas and fibrous xanthomas. **Cancer** 1963; 16:331-44.

Packenham JP, du Manoir S, Schrock E, et al. Analysis of genetic alterations in uterine leiomyomas and leiomyosarcomas by comparative genomic hybridization. **Mol Carcinog** 1997; 19:273-9.

Pander J, Wessels JA, Gelderblom H, van der Straaten T, Punt CJ, Guchelaar HJ. Pharmacogenetic interaction analysis for the efficacy of systemic treatment in metastatic colorectal cancer. **Ann Oncol** 2010. [Epub ahead of print]

Parente F, Grosgeorge J, Coindre JM, Terrier P, Vilain O, Turc-Carel C. Comparative genomic hybridization reveals novel chromosome deletions in 90 primary soft tissue tumors. **Cancer Genet Cytogenet** 1999; 115:89-95.

Pérot G, Chibon F, Montero A, et al. Constant p53 pathway inactivation in a large series of soft tissue sarcomas with complex genetics. **Am J Pathol** 2010; 177:2080-90.

Pérot G, Derré J, Coindre JM, et al. Strong smooth muscle differentiation is dependent on myocardin gene amplification in most human retroperitoneal leiomyosarcomas. **Cancer Res** 2009; 69:2269-78.

Petit MM, Lindskog H, Larsson E, et al. Smooth muscle expression of lipoma preferred partner is mediated by an alternative intronic promoter that is regulated by serum response factor/myocardin. **Circ Res** 2008; 103:61-9.

Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. **Nucleic Acids Res** 2001; 29:45.

Pisters PW, Leung DH, Woodruff J, Shi W, Brennan MF. Analysis of prognostic factors in 1,041 patients with localized soft tissue sarcomas of the extremities. **J Clin Oncol** 1996; 14:1679-89.

Pollack JR. **DNA microarrays: a molecular cloning manual**. New York: Cold Spring Harbor; 2002; RNA common reference sets; p.168-72.

Pontier SM, Huck L, White DE, et al. Integrin-linked kinase has a critical role in ErbB2 mammary tumor progression: implications for human breast cancer. **Oncogene** 2010; 29:3374-85.

Price ND, Trent J, El-Naggar AK, et al. Highly accurate two-gene classifier for differentiating gastrointestinal stromal tumors and leiomyosarcomas. **Proc Natl Acad Sci U S A**. 2007; 104:3414-9.

R Development Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria, 2009. Available from: <URL:<http://www.R-project.org>> [2011 jan 12]

Reis EM, Nakaya HI, Louro R, et al. Antisense intronic non-coding RNA levels correlate to the degree of tumor differentiation in prostate cancer. **Oncogene** 2004; 23:6684-92.

Ren B, Yu YP, Jing L, et al. Geneexpression analysis of human soft tissue leiomyosarcomas. **Hum Pathol** 2003; 34:549-58.

Retterspitz MF, Mönig SP, Schreckenber S, et al. Expression of {beta}-catenin, MUC1 and c-met in diffuse-type gastric carcinomas: correlations with tumour progression and prognosis. **Anticancer Res** 2010; 30:4635-41.

Rosenberg AE. Malignant fibrous histiocytoma: past, present, and future. **Skeletal Radiol** 2003; 32:613-8.

Saeed AI, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis. **Biotechniques** 2003; 34:374-8.

Scapolan M, Perin T, Wassermann B, et al. Expression profiles in malignant fibrous histiocytomas: clues for differentiating 'spindle cell' and 'pleomorphic' subtypes. **Eur J Cancer** 2008; 44:298-309.

Schaeffer DF, Assi K, Chan K, et al. Tumor expression of integrin-linked kinase (ILK) correlates with the expression of the E-cadherin repressor snail: an immunohistochemical study in ductal pancreatic adenocarcinoma. **Virchows Arch** 2010; 456:261-8.

Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. **Science** 1995; 270:467-70.

Schmitt T, Kasper B. New medical treatment options and strategies to assess clinical outcome in soft-tissue sarcoma. **Expert Rev Anticancer Ther** 2009; 9:1159-67.

Schürch W, Bégin LR, Seemayer TA, et al. Pleomorphic soft tissue myogenic sarcomas of adulthood. A reappraisal in the mid-1990s. **Am J Surg Pathol** 1996; 20:131-47.

Scoggins CR, Pisters PW. Diagnosis and management of soft tissue sarcomas. **Adv Surg** 2008; 42:219-28.

Segal NH, Pavlidis P, Antonescu CR, et al. Classification and subtype prediction of adult soft tissue sarcoma by functional genomics. **Am J Pathol** 2003; 163:691-700.

Shmulevich I, Hunt K, El-Naggar A, et al. Tumor specific gene expression profiles in human leiomyosarcoma: an evaluation of intratumor heterogeneity. **Cancer** 2002; 94:2069-75.

Simons A, Schepens M, Jeuken J et al. Frequent loss of 9p21 (p16(INK4A)) and other genomic imbalances in human malignant fibrous histiocytoma. **Cancer Genet Cytogenet** 2000; 118:89-98.

Skubitz KM, D'Adamo DR. Sarcoma. **Mayo Clin Proc.** 2007; 82:1409-32.

Skubitz KM, Pambuccian S, Manivel JC, Skubitz AP. Identification of heterogeneity among soft tissue sarcomas by gene expression profiles from different tumors. **J Transl Med** 2008; 6:23.

Skubitz KM, Skubitz AP. Characterization of sarcomas by means of gene expression. **J Lab Clin Med** 2004; 144:78-91.

Skubitz KM, Skubitz AP. Differential gene expression in leiomyosarcoma. **Cancer** 2003; 98:1029-38.

Slater M, Cooper M, Murphy CR. The cytoskeletal proteins alpha-actinin, Ezrin and talin are De-expressed in endometriosis and endometrioid carcinoma compared with normal uterine epithelium. **Appl Immunohistochem Mol Morphol** 2007; 15:170-4.

Snover DC, Sumner HW, Dehner LP. Variability of histologic pattern in recurrent soft tissue sarcomas originally diagnosed as liposarcoma. **Cancer** 1982; 49:1005-15.

Solyom S, Pylkäs K, Winqvist R. Screening for large genomic rearrangements of the BRIP1 and CHK1 genes in Finnish breast cancer families. **Fam Cancer** 2010; 9:537-40.

Stolf BS, Santos MM, Simao DF, et al. Class distinction between follicular adenomas and follicular carcinomas of the thyroid gland on the basis of their signature expression. **Cancer** 2006; 106:1891-900.

Svarvar C, Böhling T, Berlin O, et al. Clinical course of nonvisceral soft tissue leiomyosarcoma in 225 patients from the Scandinavian Sarcoma Group. **Cancer** 2007; 109:282-91.

Terry J, Saito T, Subramanian S, et al. TLE1 as a diagnostic immunohistochemical marker for synovial sarcoma emerging from gene expression profiling studies. **Am J Surg Pathol** 2007; 31:240-6.

Thway K. Pathology of soft tissue sarcomas. **Clin Oncol** 2009; 21:695-705.

Toguchida J, Nakayama T. Molecular genetics of sarcomas: applications to diagnoses and therapy. **Cancer Sci** 2009; 100:1573-80

Tonic I, Yu WN, Park Y, Chen CC, Hay N. Akt activation emulates Chk1 inhibition and Bcl2 overexpression and abrogates G2 cell cycle checkpoint by inhibiting BRCA1 foci. **J Biol Chem** 2010; 285:23790-821.

Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CD, Devesa SS. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978-2001: An analysis of 26,758 cases. **Int J Cancer** 2006; 119:2922-30.

Trent JC, Lazar AJ, Zhang W. Molecular approaches to resolve diagnostic dilemmas: the case of gastrointestinal stromal tumor and leiomyosarcoma. **Future Oncol** 2007; 3:629-37.

Tschoep K, Kohlmann A, Schlemmer M, Haferlach T, Issels RD. Gene expression profiling in sarcomas. **Crit Rev Oncol Hematol** 2007; 63:111-24.

Van de Rijn M, Fletcher JA. Genetics of soft tissue tumors. **Annu Rev Pathol** 2006; 1:435-66.

Van der Loop FT, Schaart G, Timmer ED, Ramaekers FC, van Eys GJ. Smoothelin, a novel cytoskeletal protein specific for smooth muscle cells. **J Cell Biol** 1996; 134:401-11.

Vandesompele J, De Preter K, Pattyn F, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. **Genome Biol** 2002; 3:1-12.

Verweij J, Baker LH. Future treatment of soft tissue sarcomas will be driven by histological subtype and molecular aberrations. **Eur J Cancer** 2010; 46:863-8.

Waalkes S, Atschekzei F, Kramer MW, et al. Fibronectin 1 mRNA expression correlates with advanced disease in renal cancer. **BMC Cancer** 2010; 10:503.

Wang R, Lu YJ, Fisher C, Bridge JA, Shipley J. Characterization of chromosome aberrations associated with soft-tissue leiomyosarcomas by twenty-four-color karyotyping and comparative genomic hybridization analysis. **Genes Chromosomes Cancer** 2001; 31:54-64.

Wang R, Titley JC, Lu YJ, Summersgill BM, Bridge JA, Fisher C, Shipley J. Loss of 13q14-q21 and gain of 5p14-pter in the progression of leiomyosarcoma. **Mod Pathol** 2003; 16:778-85.

Wang Z, Wang DZ, Pipes GC, Olson EN. Myocardin is a master regulator of smooth muscle gene expression. **Proc Natl Acad Sci USA** 2003; 100:7129-34.

Wang Z, Wang DZ, Hockemeyer D, McAnally J, Nordheim A, Olson EN. Myocardin and ternary complex factors compete for SRF to control smooth muscle gene expression. **Nature** 2004; 428:185-9.

Wardelmann E, Schildhaus HU, Merkelbach-Bruse S et al. Soft tissue sarcoma: from molecular diagnosis to selection of treatment. Pathological diagnosis of soft tissue sarcoma amid molecular biology and targeted therapies. **Ann Oncol** 2010; 21:265-269.

Weigelt B, Baehner FL, Reis-Filho JS. The contribution of gene expression profiling to breast cancer classification, prognostication and prediction: a retrospective of the last decade. **J Pathol** 2010; 220:263-80.

Weiss SW, Enzinger FM. Malignant fibrous histiocytoma: an analysis of 200 cases. **Cancer** 1978; 41:2250-66.

Weitz J, Antonescu CR, Brennan MF. Localized extremity soft tissue sarcoma: improved knowledge with unchanged survival over time. **J Clin Oncol** 2003; 21:2719-25.

Weng WH, Ahlén J, Lui WO, et al. Gain of 17q in malignant fibrous histiocytoma is associated with a longer disease-free survival and a low risk of developing distant metastasis. **Br J Cancer** 2003; 89:720-6.

West RB, Corless CL, Chen X, et al. The novel marker, DOG1, is expressed ubiquitously in gastrointestinal stromal tumors irrespective of KIT or PDGFRA mutation status. **Am J Pathol** 2004a; 165:107-13.

West RB, Harvell J, Linn SC, et al. Apo D in soft tissue tumors: a novel marker for dermatofibrosarcoma protuberans. **Am J Surg Pathol** 2004b; 28:1063-9.

West RB. Expression profiling in soft tissue sarcomas with emphasis on synovial sarcoma, gastrointestinal stromal tumor, and leiomyosarcoma. **Adv Anat Pathol** 2010; 17:366-73.

Wunder JS, Nielsen TO, Maki RG, O'Sullivan B, Alman BA. Opportunities for improving the therapeutic ratio for patients with sarcoma. **Lancet Oncol** 2007; 8:513-24.

Yan S, Shun-Chang J, Li C, Jie L, Ya-Li L, Ling-Xiong W. Topoisomerase II alpha expression and the benefit of adjuvant chemotherapy for postoperative patients with non-small cell lung cancer. **BMC Cancer** 2010; 10:621.

Yang D, Ylipaa A, Yang J, et al. An integrated study of aberrant gene copy number and gene expression in GIST and LMS. **Technol Cancer Res Treat** 2010; 9:171-8.

Yang J, Du X, Chen K, et al. Genetic aberrations in soft tissue leiomyosarcoma. **Cancer Lett** 2009; 275:1-8.

Ylipää A, Hunt KK, Yang J, et al. Integrative genomic characterization and a genomic staging system for gastrointestinal stromal tumors. **Cancer** 2011; 117:380-9.

Zhu Z, Xu X, Yu Y, et al. Silencing heat shock protein 27 decreases metastatic behavior of human head and neck squamous cell cancer cells in vitro. **Mol Pharm** 2010; 7:1283-90.

Anexo 1 - Lista de genes diferencialmente expressos obtidos por análise supervisionada de todos os 44 casos de sarcomas.

Genes	Expressão	Fold Change
BI026064	Diminuição	1.5089854
SLC39A7	Diminuição	1.3028402
METRNL	Diminuição	1.5422198
HMOX1	Diminuição	1.6509058
BG118529	Diminuição	1.9717922
C9orf140	Diminuição	1.3659436
THC2668815	Diminuição	1.6543422
DUXAP10	Diminuição	1.9624641
CA337741	Diminuição	1.2113332
MAGEA2B	Diminuição	1.60263
HMX1	Diminuição	1.3951862
A_32_P152696	Diminuição	1.4071614
SDF2L1	Diminuição	1.4031483
PYCR1	Diminuição	1.5974721
TIMM13	Diminuição	1.2709258
AF152351	Diminuição	1.1915805
IL17B	Aumento	0.5443541
ITGB1BP2	Aumento	0.5386763
SORBS1	Aumento	0.53601027
CD359823	Aumento	0.6691924
A2M	Aumento	0.6008792
CX3CL1	Aumento	0.6534483
GKAP1	Aumento	0.640447
C1orf133	Aumento	0.50003284
SHROOM3	Aumento	0.47280774
TPM2	Aumento	0.5782486
C12orf54	Aumento	0.5393912
CAP2	Aumento	0.529457
SCRG1	Aumento	0.4385267
FHL1	Aumento	0.549048
DSTN	Aumento	0.68825006
LOC51149	Aumento	0.81603056
CSRP1	Aumento	0.5963849
C10orf116	Aumento	0.53997135
RAI2	Aumento	0.5463539
C2orf40	Aumento	0.6077632
PDLIM5	Aumento	0.7065852
LAMA5	Aumento	0.69825655
MYL9	Aumento	0.61344564
MYLK	Aumento	0.5115786
BCHE	Aumento	0.45901492
SPOP	Aumento	0.82347924
ADHFE1	Aumento	0.6527549
DMN	Aumento	0.54789126
AK124702	Aumento	0.59502953
SLC25A4	Aumento	0.6362881
KCNMB1	Aumento	0.7211379
NENF	Aumento	0.8171195
FBXL22	Aumento	0.6529534
KARCA1	Aumento	0.613887
LMOD1	Aumento	0.5985782
RBPMS	Aumento	0.575111
LOC55908	Aumento	0.7134697

A_32_P214565	Aumento	0.5772532
BF675806	Aumento	0.59065264
SLC20A2	Aumento	0.7136668
SEMA6A	Aumento	0.7463925
HSPB7	Aumento	0.5765764
RAMP1	Aumento	0.49138224
ILK	Aumento	0.78457975
BFSP1	Aumento	0.7299431
AK130878	Aumento	0.5753566
CDC42EP3	Aumento	0.63050723
REEP2	Aumento	0.6123399
TCEAL1	Aumento	0.640547
FRY	Aumento	0.6016561
TAGLN	Aumento	0.6339436
TPD52L1	Aumento	0.5172491
PDLIM3	Aumento	0.44182304
SELM	Aumento	0.80656356
THC2661433	Aumento	0.5707166
FAM89B	Aumento	0.8683924
ME3	Aumento	0.750355
ENST00000377548	Aumento	0.69726783
PBXIP1	Aumento	0.78893644
THC2699924	Aumento	0.6249751
TCEAL4	Aumento	0.70358354
TCEAL2	Aumento	0.5787822
LDOC1	Aumento	0.64615285
MYCBPAP	Aumento	0.55140626
CBX7	Aumento	0.7033727
AMIGO2	Aumento	0.6116415
GADD45G	Aumento	0.6782379
CNN1	Aumento	0.5202636
BOC	Aumento	0.6603748
AF339771	Aumento	0.70347065
CKMT2	Aumento	0.5857985
A_24_P315256	Aumento	0.7336124
LOC387763	Aumento	0.64427185
ST5	Aumento	0.8157673
MAOB	Aumento	0.54720277
THC2660651	Aumento	0.625371
GPRASP1	Aumento	0.6026453
RYR3	Aumento	0.6422392
BQ019626	Aumento	0.70467037
C1orf142	Aumento	0.86465865
C5orf4	Aumento	0.6627757
C1orf175	Aumento	0.7136889
CLU	Aumento	0.5670722
AK096580	Aumento	0.67691344
SDHALP2	Aumento	0.6281722
THC2557134	Aumento	0.8038006
AK123912	Aumento	0.70212
LOC653364	Aumento	0.76572317
AK057923	Aumento	0.60981756
A_24_P76210	Aumento	0.75303155
AK130911	Aumento	0.6836029
SALL2	Aumento	0.6364833
THC2673161	Aumento	0.57986474
ENST00000367169	Aumento	0.6033278
ING4	Aumento	0.844885

MSRA	Aumento	0.8741925
ACTC1	Aumento	0.64602077
RGN	Aumento	0.58600354
C9orf18	Aumento	0.616519
TGFB1I1	Aumento	0.7280391
THC2739236	Aumento	0.631808
THC2705989	Aumento	0.63001925
THC2544977	Aumento	0.73354405
MYL6	Aumento	0.8094318
FLJ35220	Aumento	0.726401
JPH2	Aumento	0.6778685
WBSCR17	Aumento	0.62292916
FGL2	Aumento	0.62282693
HCCA2	Aumento	0.8079061
A_24_P600036	Aumento	0.53628105
SUOX	Aumento	0.8232429
FLJ41603	Aumento	0.6690802
KIAA0649	Aumento	0.756276
BX117479	Aumento	0.57032
AK124281	Aumento	0.65388095
AIG1	Aumento	0.7279365
THC2669541	Aumento	0.5828298
CA3	Aumento	0.58289635
AL133090	Aumento	0.5725773
TRIM6	Aumento	0.6813001
THC2529452	Aumento	0.7616536
RNF180	Aumento	0.7336222
CASQ2	Aumento	0.56808
MRVI1	Aumento	0.6794553
THC2505770	Aumento	0.77730143
FNBP1	Aumento	0.7730555
SCHIP1	Aumento	0.6835969
ALDH1B1	Aumento	0.69110525
ACTA2	Aumento	0.67180085
BQ777622	Aumento	0.7034156
NOXA1	Aumento	0.7595172
LOC201229	Aumento	0.72773707
SGCA	Aumento	0.68979406
THC2646628	Aumento	0.6202969
RDH5	Aumento	0.6977325
LOC286254	Aumento	0.74837303
ANG	Aumento	0.73924905
BZRAP1	Aumento	0.6446928
SRC	Aumento	0.80072105
AJ710526	Aumento	0.65849376
ENST00000354349	Aumento	0.72749734
DKFZp686D0972	Aumento	0.6363338
ENST00000372045	Aumento	0.42012072
NME3	Aumento	0.83680135
C21orf6	Aumento	0.7902041
ATP1B1	Aumento	0.7050132
ENST00000371207	Aumento	0.6831574
EPHX2	Aumento	0.61734766
THC2652817	Aumento	0.60857064
THC2655298	Aumento	0.72031575
HOP	Aumento	0.69161993
SVIL	Aumento	0.681278
LDB3	Aumento	0.7061043

CA437634	Aumento	0.66633815
FAM80B	Aumento	0.7946352
PHYHIP	Aumento	0.6838513
PAK3	Aumento	0.6803995
PDLIM7	Aumento	0.8260334
TMEM47	Aumento	0.6433277
MAP1B	Aumento	0.76682836
ENST00000375855	Aumento	0.60068744
RUSC2	Aumento	0.82600623
ABTB1	Aumento	0.8288828
AES	Aumento	0.8140003
VPS28	Aumento	0.88417345
WFDC2	Aumento	0.6401867
THC2645562	Aumento	0.65224403
LOC400027	Aumento	0.82252413
LEFTY2	Aumento	0.7086016
A_24_P177634	Aumento	0.8524742
CALM1	Aumento	0.8051721
CHRD1	Aumento	0.5656857
CRELD1	Aumento	0.8160403
TP53INP1	Aumento	0.79386705
PLXNB1	Aumento	0.6966736
PPP1R14A	Aumento	0.66651237
FLJ45803	Aumento	0.6741438
NT5DC1	Aumento	0.7679209
FLJ30313	Aumento	0.71398276
TMEM106C	Aumento	0.7711011
BC038432	Aumento	0.7117454
C11orf2	Aumento	0.8657844
LOC649003	Aumento	0.8107418
HSPB8	Aumento	0.5823413
SPG3A	Aumento	0.7126839
CYHR1	Aumento	0.8150104
COX7C	Aumento	0.85295635
COX7A2	Aumento	0.8371606
MICAL1	Aumento	0.7462828
DDIT4L	Aumento	0.5919679
RBP1	Aumento	0.6870627
NHS	Aumento	0.62405926
AK055112	Aumento	0.70960605
AK055214	Aumento	0.71415454
MYOCD	Aumento	0.73374933
FHL2	Aumento	0.7482641
AK055915	Aumento	0.68895787
HSPA4L	Aumento	0.6427596
CCDC102A	Aumento	0.7826694
DAAM2	Aumento	0.7343606
A_32_P182070	Aumento	0.7456113
C6orf60	Aumento	0.7039166
BX337332	Aumento	0.718285
A_32_P82580	Aumento	0.6671679
LGI4	Aumento	0.69289196
COX6C	Aumento	0.82130516
LOC651275	Aumento	0.7923452
TRAPPC2	Aumento	0.75542367
RTDR1	Aumento	0.675822
APBB1	Aumento	0.77978015
NBLA00301	Aumento	0.68995917

PYGM	Aumento	0.7284357
TCEAL5	Aumento	0.6900214
CALCOCO1	Aumento	0.8795992
GPRASP2	Aumento	0.7245423
MAPK10	Aumento	0.6275811
DGKG	Aumento	0.6549444
ZFP62	Aumento	0.8310188
GABBR1	Aumento	0.7891534
SMCR7	Aumento	0.820728
MAMDC2	Aumento	0.6215212
EFEMP2	Aumento	0.8545279
BOLA1	Aumento	0.8332979
ZMAT1	Aumento	0.7402712
DIXDC1	Aumento	0.67953223
AK123627	Aumento	0.7565436
CYB5R1	Aumento	0.8463264
R3HCC1	Aumento	0.8327166
ACTA1	Aumento	0.65006465
THYN1	Aumento	0.83935803
PRELP	Aumento	0.64059186
USP51	Aumento	0.7817232
SEPW1	Aumento	0.7780178
AI042308	Aumento	0.6302344
LOC63920	Aumento	0.76432365
AI659667	Aumento	0.69538945
MSRB2	Aumento	0.8121436
KIAA0753	Aumento	0.8300848
BAI3	Aumento	0.6341091
C9orf103	Aumento	0.7887293
C1orf165	Aumento	0.70889175
CCR10	Aumento	0.80854726
ENST00000285206	Aumento	0.74667895
FILIP1L	Aumento	0.76450217
ZSCAN16	Aumento	0.79851675
FOXJ2	Aumento	0.8582174
ZBTB47	Aumento	0.8275452
MBNL1	Aumento	0.80501556
SH3BGRL	Aumento	0.6805633
FYCO1	Aumento	0.7451499
THC2537367	Aumento	0.6918664
ITPKB	Aumento	0.79512286
SLC2A8	Aumento	0.86779535
C21orf2	Aumento	0.74573016
AI219192	Aumento	0.73515636
HSPB1	Aumento	0.8136381
C1QTNF2	Aumento	0.75187486
THC2685393	Aumento	0.7309575
CCDC81	Aumento	0.734462
BX419129	Aumento	0.81107444
AF220235	Aumento	0.68974686
LOC572558	Aumento	0.6647004
MYOM2	Aumento	0.73326117
C12orf57	Aumento	0.8542169
EPHB6	Aumento	0.74149305
CD9	Aumento	0.6957043
ADAMTS8	Aumento	0.71695113
PHACS	Aumento	0.800246
MANBAL	Aumento	0.822737

AOC3	Aumento	0.63030624
ENST00000357802	Aumento	0.8251387
BE780682	Aumento	0.8021337
MEF2C	Aumento	0.8048486
KCNJ8	Aumento	0.6838771
NDRG2	Aumento	0.7392628
LONRF2	Aumento	0.5753972
TCEAL6	Aumento	0.7470701
PCP4	Aumento	0.56492436
DNAJB5	Aumento	0.76961285
UQCRB	Aumento	0.84969
AW014342	Aumento	0.7277166
LOC285382	Aumento	0.68370116
ENST00000381800	Aumento	0.7603476
TCEAL3	Aumento	0.74890673
LOC152573	Aumento	0.64295244
TSPAN2	Aumento	0.66002464
ACTN1	Aumento	0.739447
TMOD1	Aumento	0.6426155
PRKCDBP	Aumento	0.8168565
CRIP3	Aumento	0.774221
DKFZp667E0512	Aumento	0.73741055
ITFG3	Aumento	0.8770067
GAS6	Aumento	0.6617351
CD671494	Aumento	0.6344415
AF087999	Aumento	0.78185034
ISOC1	Aumento	0.7047401
THC2650514	Aumento	0.83308727
KIAA1737	Aumento	0.8367092
ACTG2	Aumento	0.59997165
EHBP1L1	Aumento	0.8129305
RNF150	Aumento	0.6463567
BAI2	Aumento	0.7371463
ITIH3	Aumento	0.68477774
THC2538568	Aumento	0.7247711
TACC1	Aumento	0.84309787
LRCH2	Aumento	0.6868722
BC038512	Aumento	0.7625011
ENST00000284767	Aumento	0.61563694
NPNT	Aumento	0.72353905
THC2674891	Aumento	0.76092213
KHDRBS3	Aumento	0.73971945
ZNF16	Aumento	0.84393686
THRA	Aumento	0.799446
KLHDC6	Aumento	0.7150993
AA495894	Aumento	0.75055146
A_24_P928804	Aumento	0.7919217
LOC400236	Aumento	0.805287
CLDND1	Aumento	0.8880195
BC048201	Aumento	0.62331367
SRF	Aumento	0.8467525
FGF13	Aumento	0.6631036
NUMA1	Aumento	0.8123401
RSPO3	Aumento	0.56999975
TDRD7	Aumento	0.85024065
XK	Aumento	0.6380176
RNASE4	Aumento	0.7186326
PRTFDC1	Aumento	0.76571155

LMO1	Aumento	0.6741699
MAGI2	Aumento	0.73629457
BAD	Aumento	0.8585133
AK000774	Aumento	0.69446224
HSPA2	Aumento	0.7810501
BCL2L2	Aumento	0.85645765
NT5M	Aumento	0.76735795
GPR161	Aumento	0.783815
MUC1	Aumento	0.7149594
HAPLN2	Aumento	0.7648761
PPP1R3C	Aumento	0.63931775
THC2538534	Aumento	0.7947591
NTF3	Aumento	0.6247419
LOC729970	Aumento	0.7309076
ATP1A2	Aumento	0.58366317
ZNF250	Aumento	0.8400718
CABC1	Aumento	0.81062734
SDCCAG10	Aumento	0.8621676
CSRP2	Aumento	0.6341851
SLC24A3	Aumento	0.65920204
ENST00000325151	Aumento	0.8072539
ChGn	Aumento	0.7279812
MTA3	Aumento	0.86487585
AK022030	Aumento	0.88821447
SMTN	Aumento	0.7662147
AF086032	Aumento	0.7395103
CACNA1C	Aumento	0.73129576
LRRC49	Aumento	0.818289
A_23_P124177	Aumento	0.8271934
HTATIP	Aumento	0.85993207
BBS1	Aumento	0.84273195
PRPH	Aumento	0.7005601
HDAC5	Aumento	0.85237354
LOC283152	Aumento	0.82773453
ITIH4	Aumento	0.78961766
PPIE	Aumento	0.8742883
FLNA	Aumento	0.73866224
AF086548	Aumento	0.81216264
MOAP1	Aumento	0.8104091
AA233733	Aumento	0.70971256
THC2520542	Aumento	0.8022569
ZNF671	Aumento	0.77056503
CALD1	Aumento	0.7411432
BI597240	Aumento	0.79779595
LOC651154	Aumento	0.8118258
RBBP7	Aumento	0.8110815
AK129724	Aumento	0.73680705

Anexo 2 - Lista de genes diferencialmente expostos obtidos por análise supervisionada dos casos (42) após exclusão dos casos com a síndrome de Li-Fraumeni.

Genes	Expressão	Fold Change
IL17B	Aumento	0.64243263
ITGB1BP2	Aumento	0.5797615
CD359823	Aumento	0.6845056
A2M	Aumento	0.5359316
SORBS1	Aumento	0.48574615
CX3CL1	Aumento	0.73016816
CSRP1	Aumento	0.55356425
C1orf133	Aumento	0.60456735
LOC51149	Aumento	0.5581434
ILK	Aumento	0.5429021
C12orf54	Aumento	0.60814255
SCRG1	Aumento	0.5071694
RAI2	Aumento	0.7049499
GKAP1	Aumento	0.58548594
TPM2	Aumento	0.54111516
DSTN	Aumento	0.649062
C10orf116	Aumento	0.58762515
SHROOM3	Aumento	0.6448901
ME3	Aumento	0.8134018
FHL1	Aumento	0.6976099
C2orf40	Aumento	0.60484
RBPM5	Aumento	0.5638098
CAP2	Aumento	0.53837204
MYL9	Aumento	0.59560984
MYLK	Aumento	0.65121245
PDLIM5	Aumento	0.6374582
LMOD1	Aumento	0.8281738
DMN	Aumento	0.47133252
FBXL22	Aumento	0.72670776
AK124702	Aumento	0.5794723
ADHFE1	Aumento	0.8003698
NENF	Aumento	0.593423
LAMA5	Aumento	0.8213107
KARCA1	Aumento	0.8642143
AK130878	Aumento	0.6161366
THC2673161	Aumento	0.80569863
BF675806	Aumento	0.7003171
AK130911	Aumento	0.5259064
SLC25A4	Aumento	0.7174967
SPOP	Aumento	0.6404766
BCHE	Aumento	0.7498149
KCNMB1	Aumento	0.63164926
A_32_P214565	Aumento	0.5834803
SELM	Aumento	0.6585966
FRY	Aumento	0.6929176
NME3	Aumento	0.63912517
FAM89B	Aumento	0.7212881
THC2699924	Aumento	0.65134394
ST5	Aumento	0.5132219
C1orf175	Aumento	0.5126755
MAOB	Aumento	0.71001965
SLC20A2	Aumento	0.715799

LDOC1	Aumento	0.4982479
SEMA6A	Aumento	0.68812984
TAGLN	Aumento	0.80435157
HSPB7	Aumento	0.5779445
TRIM6	Aumento	0.8584479
ENST00000377548	Aumento	0.7020908
TCEAL1	Aumento	0.5741222
LOC55908	Aumento	0.60766643
C5orf4	Aumento	0.70176625
TPD52L1	Aumento	0.73539346
CNN1	Aumento	0.6350671
RNF180	Aumento	0.79856163
TGFB1I1	Aumento	0.8374752
RAMP1	Aumento	0.72915256
AK123912	Aumento	0.55043715
MYL6	Aumento	0.7601063
TCEAL2	Aumento	0.6433792
C1orf142	Aumento	0.64176214
TCEAL4	Aumento	0.5971833
THC2661433	Aumento	0.65695894
WBSCR17	Aumento	0.66061807
AF339771	Aumento	0.6246675
BFSP1	Aumento	0.6215164
RYR3	Aumento	0.70481056
HCCA2	Aumento	0.61123127
ING4	Aumento	0.8011321
A_24_P315256	Aumento	0.5983396
BX117479	Aumento	0.5639771
LOC653364	Aumento	0.7167069
CDC42EP3	Aumento	0.703526
LOC387763	Aumento	0.81552875
ENST00000367169	Aumento	0.3861905
BOC	Aumento	0.7944596
ENST00000371207	Aumento	0.55636984
REEP2	Aumento	0.684902
SDHALP2	Aumento	0.76510316
CBX7	Aumento	0.4582945
AMIGO2	Aumento	0.7489024
AES	Aumento	0.6293223
GPRASP1	Aumento	0.6296303
CLU	Aumento	0.75666696
FLJ35220	Aumento	0.7296029
BQ019626	Aumento	0.6114435
SUOX	Aumento	0.6749488
ENST00000372045	Aumento	0.7633008
PBXIP1	Aumento	0.6706456
MYCBPAP	Aumento	0.625187
GADD45G	Aumento	0.5685611
FNBP1	Aumento	0.6170555
PDLIM3	Aumento	0.8763455
NOXA1	Aumento	0.53583497
THC2660651	Aumento	0.57763344
THC2739236	Aumento	0.66686505
A_24_P76210	Aumento	0.8155735
THC2544977	Aumento	0.6668277
AK057923	Aumento	0.5398423
AK096580	Aumento	0.6828296
CCDC102A	Aumento	0.71808225

MRV11	Aumento	0.6160916
C9orf18	Aumento	0.8085668
AL133090	Aumento	0.6098699
FGL2	Aumento	0.8173485
MSRA	Aumento	0.7477263
A_24_P600036	Aumento	0.66606295
THC2669541	Aumento	0.5449994
FLJ41603	Aumento	0.5395503
PDLIM7	Aumento	0.6737834
SVIL	Aumento	0.5847463
CHRD1	Aumento	0.5392912
HOP	Aumento	0.63888794
ENST00000354349	Aumento	0.5735511
THC2646628	Aumento	0.48833066
THC2557134	Aumento	0.8066296
CKMT2	Aumento	0.7672161
RUSC2	Aumento	0.5357702
LOC286254	Aumento	0.42626223
NBLA00301	Aumento	0.5342367

Anexo 3 - Lista de genes diferencialmente expostos obtidos por análise supervisionada dos casos (40) após exclusão dos casos com a metástase ao diagnóstico.

Genes	Expressão	Fold Change
IL17B	Aumento	0.52829134
ITGB1BP2	Aumento	0.5262058
SORBS1	Aumento	0.5294845
CD359823	Aumento	0.66939086
TPM2	Aumento	0.57127565
RAI2	Aumento	0.5102326
A2M	Aumento	0.5897533
C2orf40	Aumento	0.5778234
CSRP1	Aumento	0.5833231
PDLIM5	Aumento	0.69427955
ADHFE1	Aumento	0.63303614
MYLK	Aumento	0.49919108
CX3CL1	Aumento	0.6491574
SCRG1	Aumento	0.44122088
MYL9	Aumento	0.5973626
FHL1	Aumento	0.5394514
LOC51149	Aumento	0.8119823
C1orf133	Aumento	0.50435275
GKAP1	Aumento	0.64494663
NHS	Aumento	0.582001
SHROOM3	Aumento	0.4909152
AMIGO2	Aumento	0.6001711
DMN	Aumento	0.5286144
C10orf116	Aumento	0.5396682
DSTN	Aumento	0.69107395
LMOD1	Aumento	0.5840383
LAMA5	Aumento	0.7073947
TPD52L1	Aumento	0.513486
CAP2	Aumento	0.53772646
GPRASP1	Aumento	0.56033224
AK124702	Aumento	0.5954822
SLC25A4	Aumento	0.63296205
A_32_P214565	Aumento	0.5688925
TCEAL1	Aumento	0.61406845
FBXL22	Aumento	0.6416993
NENF	Aumento	0.80983347
SPOP	Aumento	0.8208456
TAGLN	Aumento	0.6161334
ILK	Aumento	0.7750482
C12orf54	Aumento	0.5643051
LOC387763	Aumento	0.6181905
TCEAL4	Aumento	0.6835936
RBPM5	Aumento	0.57002884
LOC55908	Aumento	0.70450366
KARCA1	Aumento	0.6145105
A_24_P600036	Aumento	0.5195343
CASQ2	Aumento	0.52589923
TCEAL2	Aumento	0.56485707
BCHE	Aumento	0.4668209
THC2661433	Aumento	0.55060303
BF675806	Aumento	0.5921353
C1orf142	Aumento	0.852296

AK130878	Aumento	0.5667748
THC2699924	Aumento	0.6136477
SDHALP2	Aumento	0.6124682
KCNMB1	Aumento	0.7239076
A_24_P315256	Aumento	0.72058845
HSPB7	Aumento	0.5979103
PDLIM3	Aumento	0.42725316
SELM	Aumento	0.80511665
SEMA6A	Aumento	0.75625104
CBX7	Aumento	0.6870967
RAMP1	Aumento	0.50804365
BZRAP1	Aumento	0.619352
ST5	Aumento	0.8061579
PLXNB1	Aumento	0.6653035
KHDRBS3	Aumento	0.7014207
SALL2	Aumento	0.6140344
AK057923	Aumento	0.5860663
A_24_P76210	Aumento	0.74847484
THC2705989	Aumento	0.61596197
CDC42EP3	Aumento	0.62701476
LOC400027	Aumento	0.7960133
THC2529452	Aumento	0.73855966
PBXIP1	Aumento	0.78267735
FRY	Aumento	0.60488176
REEP2	Aumento	0.62017095
AK096580	Aumento	0.6722961
C5orf4	Aumento	0.65403455
CKMT2	Aumento	0.5775553
BQ019626	Aumento	0.6916879
MYL6	Aumento	0.80734223
CLU	Aumento	0.5535101

Anexo 4 - Lista de genes diferencialmente expostos obtidos por análise supervisionada dos casos (38) após exclusão dos casos com a síndrome de Li-Fraumeni e metástase ao diagnóstico.

Genes	Expressão	Fold Change
IL17B	Aumento	0.7175479
RAI2	Aumento	0.56011206
ITGB1BP2	Aumento	0.5826821
RNF180	Aumento	0.8740823
C2orf40	Aumento	0.616881
A2M	Aumento	0.7712361
TPM2	Aumento	0.73583275
SORBS1	Aumento	0.6741091
FBXL22	Aumento	0.5662341
CD359823	Aumento	0.8212037
SUOX	Aumento	0.64101744
CX3CL1	Aumento	0.6572193
CYB5R1	Aumento	0.6752834
MYL9	Aumento	0.6757423
SCRG1	Aumento	0.5741783
CSRP1	Aumento	0.6312721
RBP1	Aumento	0.61479086
SHROOM3	Aumento	0.601613
AMIGO2	Aumento	0.68601584
ING4	Aumento	0.63091123
LOC51149	Aumento	0.8045823
C1orf133	Aumento	0.72951615
FHL1	Aumento	0.6975559
LAMA5	Aumento	0.80609924
C10orf116	Aumento	0.7240457
NENF	Aumento	0.64649385
KHDRBS3	Aumento	0.75110215
MYLK	Aumento	0.8284405
GPRASP1	Aumento	0.798063
FAM89B	Aumento	0.61009985
NHS	Aumento	0.7451602
THC2505770	Aumento	0.8250861
PDLIM5	Aumento	0.8210643
BOC	Aumento	0.71022666
A_24_P600036	Aumento	0.7907136
ABTB1	Aumento	0.88059366
KARCA1	Aumento	0.5846261
TAGLN	Aumento	0.59313864
SLC25A4	Aumento	0.82856834
ADHFE1	Aumento	0.8280697
MAOB	Aumento	0.70005554
LMOD1	Aumento	0.65743154
FGL2	Aumento	0.71072227
A_32_P214565	Aumento	0.561742
GKAP1	Aumento	0.807496
TCEAL1	Aumento	0.84761935
CRELD1	Aumento	0.8091193
LOC55908	Aumento	0.81709605
PLXNB1	Aumento	0.74417853
ILK	Aumento	0.7154825
DSTN	Aumento	0.7932265
AK124702	Aumento	0.7633166

THC2529452	Aumento	0.8111982
THTPA	Aumento	0.78461134
PBXIP1	Aumento	0.66793585
RBPMS	Aumento	0.72191364
A_24_P315256	Aumento	0.7352267
RUSC2	Aumento	0.5937147
ST5	Aumento	0.55272293
BC038432	Aumento	0.7901628
PHACS	Aumento	0.86864936
MSRA	Aumento	0.6652074
CAP2	Aumento	0.7314267
DMN	Aumento	0.47354397
PDLIM7	Aumento	0.62093914
SELM	Aumento	0.81647223
AK096580	Aumento	0.60185766
REEP2	Aumento	0.76521
TCEAL4	Aumento	0.7725035
TPD52L1	Aumento	0.6860192
SRC	Aumento	0.62962383
EFEMP2	Aumento	0.7784885
SMCR7	Aumento	0.7319106
HCCA2	Aumento	0.6519302
THC2544977	Aumento	0.5941361
CBX7	Aumento	0.8649759
ITPKB	Aumento	0.6978587
KIAA0649	Aumento	0.8533435
THC2557134	Aumento	0.7591825
THRA	Aumento	0.7416263
LOC387763	Aumento	0.6951875
BQ019626	Aumento	0.7043865
ENST00000377548	Aumento	0.632682
CASQ2	Aumento	0.73607516
RAMP1	Aumento	0.6883384
LOC653364	Aumento	0.7139313
C1orf142	Aumento	0.8405259
CR616845	Aumento	0.666827
C21orf2	Aumento	0.7432841
PDLIM3	Aumento	0.59946966
BAI3	Aumento	0.65638924
CYHR1	Aumento	0.8340226
THC2661433	Aumento	0.73245364
KCNMB1	Aumento	0.7897392
NANOS3	Aumento	0.7628023
MRVI1	Aumento	0.6780298
TCEAL2	Aumento	0.8548663
SEPW1	Aumento	0.7903461
BAI2	Aumento	0.5352575
SALL2	Aumento	0.6071831
THC2669541	Aumento	0.70540994
SLC2A8	Aumento	0.89127636
SLC40A1	Aumento	0.6956632
SPOP	Aumento	0.6407332
TGFB1I1	Aumento	0.6630997
FLJ35220	Aumento	0.7335404
ACTA2	Aumento	0.85274297
BE795374	Aumento	0.7108128
AK057923	Aumento	0.85705876
AK123912	Aumento	0.7166503

LGI4	Aumento	0.60710764
GPRASP2	Aumento	0.8578293
R3HCC1	Aumento	0.58335316
THC2705989	Aumento	0.7134877
C1orf175	Aumento	0.80174994
MYCBPAP	Aumento	0.8409696
HSPB7	Aumento	0.7693326
MANBAL	Aumento	0.78963697
LOC201229	Aumento	0.8620571
ENST00000325151	Aumento	0.71936345
BHLHB9	Aumento	0.6419099
CDC42EP3	Aumento	0.84746087
ZSCAN16	Aumento	0.81247836
FBLIM1	Aumento	0.76789796
BCHE	Aumento	0.6510056
CLU	Aumento	0.7994675
FLJ45803	Aumento	0.7947179
VPS28	Aumento	0.7642864
C5orf4	Aumento	0.87695956
AK130878	Aumento	0.785417
THC2739236	Aumento	0.64348006
DAAM2	Aumento	0.7488625
BAG1	Aumento	0.8532974
SCHIP1	Aumento	0.8452572
BAD	Aumento	0.8243945
AK130911	Aumento	0.78388035
SCG5	Aumento	0.6818071
PURA	Aumento	0.7215664
ATP1A2	Aumento	0.85090214
LDOC1	Aumento	0.55891913
APBB1	Aumento	0.51653767
THYN1	Aumento	0.56357425
BFSP1	Aumento	0.69212997
CCDC102A	Aumento	0.5970814
RRAGA	Aumento	0.62289387
ZSCAN18	Aumento	0.60811055
CKMT2	Aumento	0.5775814
HDAC5	Aumento	0.6486838
LOC651154	Aumento	0.7176419
ACTN1	Aumento	0.810794
FRY	Aumento	0.686741
ME3	Aumento	0.8147798
C1orf203	Aumento	0.6320271
CORO6	Aumento	0.4835627
C1orf122	Aumento	0.63482887
USP51	Aumento	0.6521279
C12orf54	Aumento	0.5430893
ANG	Aumento	0.6367467
COX7A2	Aumento	0.83906245
MYL6	Aumento	0.828699
LOC400027	Aumento	0.540171
NOXA1	Aumento	0.5727523
AK124281	Aumento	0.7345339
SPG3A	Aumento	0.57945013
LOC145783	Aumento	0.83271176

Anexo 5 - Lista de genes diferencialmente expostos obtidos por análise supervisionada dos casos (24) de SP e LM mais característicos.

Genes	Expressão	Fold Change
MAGEA12	Diminuição	1.6354046
TIMM13	Diminuição	1.2551867
MAGEA6	Diminuição	1.6330463
SLC38A6	Diminuição	1.2992396
HSD17B7P2	Diminuição	1.3654788
CA337741	Diminuição	1.2690853
BI026064	Diminuição	1.6720037
MAPK13	Diminuição	1.4157853
MAGEA2B	Diminuição	2.0274642
AMIGO2	Aumento	0.5187305
IL17B	Aumento	0.48998165
PDLIM5	Aumento	0.69176745
CAP2	Aumento	0.4636492
SHROOM3	Aumento	0.36736372
BCHE	Aumento	0.3727083
ITGB1BP2	Aumento	0.53024006
TPM2	Aumento	0.5375153
SORBS1	Aumento	0.49549055
A2M	Aumento	0.5835521
THC2661433	Aumento	0.45728374
ADHFE1	Aumento	0.6520849
GKAP1	Aumento	0.61036676
FHL1	Aumento	0.4764148
PDLIM3	Aumento	0.36803383
CNN1	Aumento	0.39934647
LAMA5	Aumento	0.66892296
RAI2	Aumento	0.5201944
LOC387763	Aumento	0.55775553
DSTN	Aumento	0.6756567
C2orf40	Aumento	0.5535742
HSPB7	Aumento	0.49801984
FTCD	Aumento	0.7101013
C14orf65	Aumento	0.7734543
LOC572558	Aumento	0.5232195
BBS1	Aumento	0.79441196
SALL2	Aumento	0.5667837
BZRAP1	Aumento	0.54221463
SCRG1	Aumento	0.3921736
THC2655298	Aumento	0.616334
ENST00000285206	Aumento	0.64303243
PTPRF	Aumento	0.6715788
BI597240	Aumento	0.78902775
FILIP1L	Aumento	0.69030416
CX3CL1	Aumento	0.6199934
LOC285382	Aumento	0.5614077
TOX4	Aumento	0.8352144
LOC55908	Aumento	0.6797485
DMN	Aumento	0.5166792
KIAA0649	Aumento	0.69378686
ZSCAN18	Aumento	0.60278594
TMEM27	Aumento	0.6170113
ACTG2	Aumento	0.40334535
CDC42EP3	Aumento	0.54165226

BTG1	Aumento	0.79207087
FRY	Aumento	0.54710805
AOC3	Aumento	0.5784464
AF086261	Aumento	0.63314664
XPA	Aumento	0.83006763
SEMA6A	Aumento	0.69932294
FBXO21	Aumento	0.8331037
CD359823	Aumento	0.6810989
ZBED3	Aumento	0.6789654
FBXL5	Aumento	0.831334
TXLNB	Aumento	0.5731356
SDHALP2	Aumento	0.58098876
KIAA0753	Aumento	0.81592214
EIF1B	Aumento	0.85833454
ENST00000284767	Aumento	0.5462365
TCEAL4	Aumento	0.66908085
KIAA1166	Aumento	0.65221083
DYNLT1	Aumento	0.84775376
NPNT	Aumento	0.6185359
MPP2	Aumento	0.6412057
BX098326	Aumento	0.7227933
ACTC1	Aumento	0.5804747
LRCH2	Aumento	0.63307744
AIG1	Aumento	0.6775717
KIAA1505	Aumento	0.7238019
AK130878	Aumento	0.5159576
MAP1B	Aumento	0.7325248
ALPK3	Aumento	0.64779544
HSPB8	Aumento	0.48557797
FLJ41603	Aumento	0.6024465
BF675806	Aumento	0.55015314
KIAA0182	Aumento	0.8034877
THC2625435	Aumento	0.6875667
TPD52L1	Aumento	0.46838558
FXYD6	Aumento	0.6205543
PTBP2	Aumento	0.76552165
ZNF250	Aumento	0.8329746
THC2652817	Aumento	0.5443773
CCNI	Aumento	0.8700359
SMTN	Aumento	0.6629929
KCNJ8	Aumento	0.6783637
THC2645562	Aumento	0.5525209
TSPYL4	Aumento	0.8347787
DKFZp686D0972	Aumento	0.53514636
DKFZP761M1511	Aumento	0.70051074
KIAA1737	Aumento	0.81269
PBX1	Aumento	0.62762374
WDR6	Aumento	0.82074475
NEXN	Aumento	0.6307305
WBSCR17	Aumento	0.58953816
AL133090	Aumento	0.51491416
AK055112	Aumento	0.7291986
AK123627	Aumento	0.75555104
AK124281	Aumento	0.5943903
THC2674306	Aumento	0.56543994
MYCBPAP	Aumento	0.51364666
MAPK10	Aumento	0.5392644
RTDR1	Aumento	0.6232891

SUSD4	Aumento	0.632888
ATBF1	Aumento	0.7595955
ENST00000375855	Aumento	0.55284846
DDIT4L	Aumento	0.5625303
BX117479	Aumento	0.48400107
SLC40A1	Aumento	0.69495565
MYL9	Aumento	0.5970693
MYLK	Aumento	0.49330935
THC2683448	Aumento	0.7610516
PFKP	Aumento	0.69881046
FBXL22	Aumento	0.6332472
SLC24A3	Aumento	0.55331117
THC2557134	Aumento	0.7797053
A_24_P76210	Aumento	0.7115803
RBPM5	Aumento	0.5698599
DKFZp586C0721	Aumento	0.6889572