

**Universidade Estadual Paulista “Júlio de Mesquita Filho”
Instituto de Biociências**

PRISCILA DANIELE RAMOS CIRILO

**ALTERAÇÕES GENÔMICAS E PERFIL DE EXPRESSÃO
GÊNICA GLOBAL EM LEIOMIOMAS UTERINOS**

Tese apresentada ao Instituto de Biociências de Botucatu, Universidade Estadual Paulista “Júlio de Mesquita Filho” – UNESP, como requisito á obtenção do título de Doutor em Ciências Biológicas – Área de concentração: Genética.

Orientadora: Profa. Dra. Silvia Regina Rogatto

**Botucatu
2011**

Dedico este trabalho,

Ao meu Deus, que sempre cuidou de mim em todas as etapas da minha vida.

Obrigada Senhor, por mais esta conquista.

“Porque dele e por ele, e para ele, são todas as coisas.”

Romanos 11:36

AGRADECIMENTOS

À minha orientadora Profa. Dra. Silvia Regina Rogatto,

Obrigada pela oportunidade de trabalhar em sua equipe, pelo carinho, confiança e ensinamentos que me acompanharão por toda a vida. Sempre serei grata por toda a atenção e tempo dedicados a mim para realização deste trabalho. Muito Obrigada!

À Dra. Anaglória Pontes,

À Dra. Maria Aparecida Custódio Domingues,

À Dra. Cláudia Aparecida Rainho,

Agradeço todas vocês pelas colaborações essenciais neste trabalho.

À Biblioteca e Pós-graduação - IBB, UNESP, Botucatu,

Obrigada por toda ajuda e atenção durante estes anos.

À todos do Centro Internacional de Pesquisa e Ensino (CIPE),

Obrigada por nos receber e pela contribuição necessária para execução deste trabalho.

À FAPESP, CAPES e CNPq pelo apoio financeiro.

À todos integrantes e ex-integrantes

do Laboratório NeoGene,

Ana, Anderson, André, Ariane, Carlos, Carol, Cássia, Deborah, Eliane, Eloísa, Fábio, Fabíola, Fernanda, Francine, Georgia, Graziela, Gustavo, Haydée, Hellen, Juan, Julia, Lívia, Luciana, Márcia, Mateus, Miriam, Renata Bueno, Renata Canevari, Renée, Rodrigo, Rogério, Sabrina, Sandra, Sandra Regina e Sara.

e do CIPE - Hospital A. C. Camargo,

Alex, Carol, Eloísa, Felipe, Gustavo, Rodrigo, Nair, Vera e Yukie.

Cada um, de alguma forma, marcou a minha vida. Obrigada pelos momentos felizes, tristes, pelas ajudas nos experimentos, correções de relatório, tese e etc....enfim, pela história construída durante este tempo. Obrigada!

**“E amigos são prá sempre como eterno é nosso Deus. Como amigos nós diremos
até breve, nunca adeus.”**

Á minha família,

meus pais Jair e Gumercinda e minha irmã Rafaela, pelo amor, carinho e por sempre acreditarem nos meus sonhos. Vocês sabem que não foi fácil, mas com vocês ao meu lado, tudo se tornou mais simples. Amo vocês. Obrigada.

À minha avó Cida,

mesmo com diagnóstico de Mal de Alzheimer há sete anos, sou muito grata por ainda se lembrar de mim. Saiba que todas as vezes que eu ouço da sra “A vó sempre ora pedindo prá Deus te guardar lá em São Paulo” isso me dá forças prá continuar.

Ao Marcelo,

por sua paciência e compreensão durante a realização deste trabalho. Sua presença, amor, carinho e senso de humor peculiar me ajudaram muito! Te amo!

Aos meus amigos Fábio, Lucas e Virginei,

por todas as vezes que eu volto para Araçatuba e me fazerem sentir como se eu nunca tivesse saído de lá! Mesmo longe, obrigada por existirem em minha vida.

Á minha amiga Bruna,

por sempre me fazer acreditar que a vida “é mais do que isso”. Obrigada pelas risadas e pelos conselhos!

À todos da Igreja Metodista em Araçatuba e Botucatu

pelas orações e amizade construída durante todo este tempo. Deus abençoe vocês!

RESUMO

Os Leiomiomas Uterinos (LU) são os tumores benignos mais comuns do trato genital feminino afetando entre 25-30% das mulheres em idade reprodutiva. Embora com etiologia pouca conhecida, estes tumores constituem um importante problema de saúde pública, sendo a principal indicação para histerectomia. Aproximadamente 40-50% dos LU apresentam anormalidades citogenéticas não randômicas, portanto, a outra metade pode conter alterações submicroscópicas, como as alterações no número de cópias de DNA (CNVs). Estudos de expressão gênica globais têm revelado o envolvimento de genes que participam das vias de proliferação e ciclo celular, do ácido retinóico, sinalização TGF-beta e IGF-1 em resposta ao estrógeno e a progesterona. Entretanto, poucos genes mapeados em regiões de CNVs foram diretamente associados com o desenvolvimento dos LU. O objetivo deste estudo foi investigar alterações no número de cópias de DNA em pacientes com LU múltiplos e únicos. Utilizando a técnica de CGH array, foram avaliadas 80 amostras de LU obtidas de 56 pacientes. O perfil de expressão gênica global foi investigado em 51 amostras com o objetivo de identificar genes diferencialmente expressos em relação ao miométrio adjacente (MM). Posteriormente, os dados genômicos e transcriptômicos foram integrados com o objetivo de identificar CNVs funcionais que possam estar associadas ao fenótipo tumoral, sendo esta análise importante para a descrição de novos marcadores moleculares e possíveis alvos terapêuticos. Amostras de DNA e RNA foram obtidas dos LU e MM de pacientes nas fases proliferativa e secretora do ciclo menstrual. As 80 amostras de DNA tumoral e DNA referência universal foram digeridas, marcadas e hibridadas em lâminas *Agilent Human 4x44K CGH Microarrays*. Os dados foram analisados pelo programa *Nexus v5.0* com o algoritmo *Rank Segmentation* e limiar de significância de $1,00E^{-4}$. A classificação de CNVs não casuais ($P \leq 0,05$) foi baseada na presença em mais de 10% dos casos e exclusão de CNVs presentes em mais de 5% da população controle. Cinquenta e uma amostras de RNA foram convertidas em cRNA marcados e foram hibridados em lâminas *Agilent Whole Human Genome 4x44K*. Foram aplicados os testes SAM, $FDR < 5\%$, 1000 permutações e agrupamento hierárquico não-supervisionado (HCL - correlação de *Pearson*) pelo programa *TMeV v.4.5*, para identificar os genes diferencialmente expressos em relação ao MM. A classificação dos genes com aumento ou diminuição de expressão foi baseada em valores de *fold-change* com *cut-off* de 1.5. A integração dos dados foi realizada pelo algoritmo CONEXIC, com utilização da análise de enriquecimento de genes (GSEA) para a classificação de genes moduladores. Foram identificadas 45 CNVs significativas em 77/80 amostras (96%), sendo 25% dos casos com CNVs não casuais; 29 CNVs (64%) ainda não haviam sido descritas. Os ganhos genômicos mais frequentes foram em 2q35, 6p21.32, 7q22.1 e 11q12.3-q13.2 e perdas em 4q28.3, 7q22.1-q22.2 e 7q31.33. A análise de agrupamento hierárquico (teste exato de Fisher) revelou uma tendência de separação de dois grupos de amostras com tumores múltiplos (Grupo 1) e não múltiplos (Grupo 2). Foram observados ganhos em 2q35, 7q22.1, 19p13.12-p13.11 e 19q13.32-q13.33 ($P < 0,01$, tumores múltiplos) e perdas em 7q22.1-q22.2 ($P < 0,01$, tumores não-múltiplos). A análise funcional de redes e vias dos genes mapeados nas regiões significativas realizadas pelo sistema *Ingenuity Pathway Analysis* (IPA) revelou funções associadas

com proliferação celular, processos fibróides e expressão gênica. A análise funcional revelou que genes do Grupo 1 (ganhos) foram associados com processos de transcrição e proliferação celular e genes do Grupo 2 (perdas), com processos de organização e ciclo celular. A análise de expressão gênica global revelou 206 genes diferencialmente expressos em relação ao MM, no qual 34% apresentaram expressão aumentada (incluindo *COL3A1*, *COL4A1*, *COL17A1*, *IGF1*) e 66% diminuída (incluindo *CDKN2A*, *JUN*, *FOS*, *PCNA*). Genes com expressão aumentada atuam principalmente na proliferação e matriz extracelular e genes com expressão diminuída participam dos processos de recombinação cromossômica e reparo a danos no DNA. A análise integrada dos dados revelou 75 genes moduladores, entre os quais 26 tiveram correlação positiva e 3 correlação negativa. Trinta e um genes com correlação inversa apresentaram regulação por miRNAs. A análise GSEA revelou 32/75 genes associados a módulos de genes de câncer. Entre estes genes, *FGFR1*, *DBN1*, *NUPR1*, *COL3A1*, *AOC3*, *IFITM1*, *CALCRL*, *NOVA2*, *MVP*, *MFAP5*, *DIP2C* apresentaram correlação positiva e *CENPF*, *RHOH* e *BICD1* apresentaram correlação negativa. A análise funcional dos 75 genes moduladores revelou que eles atuam principalmente no ciclo celular e proliferação e crescimento celular, confirmando os achados independentes de CGH array e microarrays de expressão. Os genes *FGFR1* e *IGFBP5* foram associados com vias de proliferação e crescimento celular, enquanto que o *COL3A1* foi associado com organização celular. Estes genes estavam localizados na periferia das redes e ligados indiretamente á fatores de transcrição, indicando que podem atuar indiretamente na transcrição de genes alvos. Em adição, os genes *COL3A1*, *IGFBP5* e *FGFR1* foram associados à via canônica de proliferação da fibrose hepática e podem ser potenciais alvos terapêuticos. Os achados deste estudo permitiram descrever novas regiões de CNVs, além de confirmar o envolvimento de outras já descritas em LU. Estas CNVs envolvem genes associados principalmente com mecanismos transcricionais e ciclo celular. A análise de expressão gênica global revelou a diminuição da atividade de genes que controlam os processos de recombinação cromossômica e reparo a danos no DNA e genes com atividade aumentada associados com estímulos de proliferação e acúmulo da matriz extracelular. A nova estratégia de integração dos dados genômicos e transcriptômicos revelou novos marcadores moleculares com uso potencial para o tratamento do Leiomiomas Uterinos, uma doença frequente e com etiologia pouco conhecida.

ABSTRACT

Uterine leiomyomas (UL) are the most common benign tumors affecting between 25-30% of women in reproductive age. Although little is known about its etiology, these tumors represent a major problem in public health being the main indication for hysterectomy. About 40-50% have nonrandom cytogenetic abnormalities, thus half of these tumors may have submicroscopic alterations, as well as copy number variations (CNVs). Global gene expression studies have shown genes that act in proliferation and cell cycle process, retinoic acid, TGF-beta and IGF-1 signaling in response to estrogen and progesterone. However, few genes mapped at CNVs regions were directly associated with the UL development. The aim of this study was to investigate DNA copy number alterations in patients with multiple and unique UL. For this, the CGH array technique was employed in 80 samples obtained from 56 patients. The global gene expression profile was investigated in 51 samples to identify differentially expressed genes in relation to adjacent myometrium (MM). In order to identify functional CNVs that might be associated with tumor phenotype, the genomic and transcriptomic data were evaluated using integrative analysis, which is crucial to describe new molecular markers and putative therapeutic targets. DNA and RNA samples were obtained from UL and MM of patients in the proliferative and secretory menstrual cycle phases. Eighty DNA samples and reference DNA were digested, labeled and hybridized in Agilent Human CGH Microarrays 4x44K slides. Data were analyzed by Nexus v5.0 with the Rank Segmentation algorithm and 1.00×10^{-4} threshold. The classification of nonrandom CNVs ($P \leq 0.05$) was the presence in more than 10% cases and exclusion of CNVs detected in more than 5% of the population control. Fifty-one samples of RNA were converted into labeled cRNA and hybridized in Agilent Human Whole Genome 4x44K slides. Statistical tests SAM, FDR <5% with 1000 permutations and unsupervised hierarchical clustering (HCL - *Pearson* correlation) were applied using TMEV v4.5 software to identify genes differentially expressed in LU in comparison with MM. The classification of down-regulated or up-regulated genes was based on cut-off of 1.5 fold-change. The integrative analysis was performed by the algorithm CONEXIC, using Gene Set Enrichment Analysis (GSEA) for classification of genes modulators. Forty-five significant CNVs were identified in 77/80 samples (96%) whereas 25% of samples had nonrandom CNVs; 29 (64%) of them were CNVs no described previously in UL. The most frequent gains were at 2q35, 6p21.32, 7q22.1 and 11q12.3-q13.2 and losses at 4q28.3, 7q22.1-q22.2 and 7q31.33. The hierarchical clustering analysis (Fisher's exact test) showed a marginally significant association between samples with multiple tumors (cluster 1) and non-multiple tumors (cluster 2). It was verified gains on 2q35, 7q22.1, 19p13.12-p13.11 and 19q13.32-q13.33 ($P < 0.01$, multiple tumors) and losses on 7q22.1-q22.2 ($P < 0.01$, non-multiple tumors). Networks and pathways functional analysis by Ingenuity Pathway Analysis system (IPA) of genes mapped at significant regions revealed association with cell proliferation, fibrosis processes, and gene expression function. Functional analysis revealed that genes at cluster 1 (gains) were associated with transcriptional and cell proliferation functions and cluster 2 genes (losses) were associated with organization and cell cycle. The global gene expression analysis revealed 206 genes differentially expressed in comparison with MM, in which 34% were up-regulated (including *COL3A1*, *COL4A1*, *COL17A1*, *IGF1*) and 66% down-regulated (including *CDKN2A*, *JUN*, *FOS*, *PCNA*). Up-regulated genes were involved in cellular proliferation and extracellular matrix process, while down-regulated genes participate in chromosomal recombination and DNA damage repair. The integrative analysis revealed 75 genes modulators, whereas 26 had positive correlation and 3 negative correlations. Thirty-one genes that show an inverse correlation were regulated by miRNAs. The GSEA analysis demonstrated 32/75 genes associated with

cancer genes modules. Among them, *FGFR1*, *DBN1*, *NUPR1*, *COL3A1*, *AOC3*, *IFITM1*, *CALCRL*, *NOVA2*, *MVP*, *MFAP5*, *DIP2C* had positive correlation and *CENPF*, *RHOH* and *BICD1* had negative correlation. Functional analysis of 75 genes modulators revealed that they act mainly on cell cycle and cell growth and proliferation, confirming the findings of independent CGH array and expression microarrays data. *IGFBP5* and *FGFR1* genes were associated with proliferation and cell growth process, while *COL3A1* was associated with cellular organization. These genes were located on the networks periphery, connected to transcription factors, indicating that may act indirectly in the transcription. In addition, *COL3A1*, *IGFBP5* and *FGFR1* genes have been associated with the canonical pathway of proliferation of hepatic fibrosis and may be potential therapeutic targets. The findings of this study allowed to describe new CNVs regions, and also confirmed the involvement of others CNVs already described in LU. Genes mapped at CNVs regions were associated with transcriptional mechanisms and cell cycle. The global gene expression analysis revealed decreased activity of genes that control the processes of chromosomal recombination and DNA damage repair and increased activity of genes associated with stimulation of proliferation and accumulation of extracellular matrix. This new strategy of integrative genomic and transcriptomic data reveals new potential molecular markers that could be used on treatment of Uterine Leiomyoma, a common and poorly understood disease.

LISTA DE FIGURAS

- Figura 1** A – Frequência das alterações número de cópias identificadas entre as 80 amostras de Leiomiomas Uterinos; PH = perdas homozigotas; ANG = alto nível de ganhos. É possível observar uma pequena prevalência das perdas sobre ganhos genômicos. B – Número de alterações no número de cópias identificadas ao longo dos cromossomos. O cromossomo 1 apresentou um maior número de perdas genômicas e o cromossomo 11 teve um maior número de ganhos 45
- Figura 2** Ideograma representando o perfil geral das alterações genômicas detectadas nas 80 amostras de LU avaliadas. As perdas genômicas estão representadas ao lado esquerdo dos cromossomos (vermelho) e os ganhos genômicos ao lado direito (verde) (*Nexus v.5.0 -Biodiscovery*) 46
- Figura 3** Rearranjos cromossômicos resultando em um possível cromossomo em anel nos casos 084T: ?r(1)(p36.11q42.12) e 759T: ?r(1)(p31.1q32.2). Os pontos azuis representam os sinais de hibridação. Ao lado direito do ideograma está representada a região envolvida em ganhos (verde) e ao lado esquerdo, as perdas (vermelho) (*Nexus v.5.0 -Biodiscovery*) 47
- Figura 4** Ganhos genômicos detectados em 19p, com região mínima comum em (p13.3p12) (região selecionada em retângulo vermelho). Ao lado direito do eixo observa-se a frequência das alterações. O tamanho de cada alteração está representado acima de cada reta, no qual representa o número da amostra afetada (300T_A, 301T_B, 301T_C, 304T, 317T_C, 629T_A, 631T_B, 643T_A, 733T_A, 733T_B) (*Nexus v.5.0 -Biodiscovery*) 48
- Figura 5** Perfil genômico global onde podem ser observadas as alterações significativas ($P \leq 0.05$) indicada pelos asteriscos. Estas regiões foram detectadas em $\geq 10\%$ dos casos. Foram excluídas as CNVs presentes em $\geq 5\%$ da população de referência. Os cromossomos sexuais foram excluídos da análise (*Nexus v.5.0 -Biodiscovery*) 50

Figura 6	Regiões significativas mais frequentemente observadas entre as 80 amostras de LU, como ganhos em 2q35, 6p21.32, 7q22.1 e 11q12.3-q13.2 e perdas em 4q28.3, 7q22.1-q22.2 e 7q31.33 (região selecionada com o retângulo). Ao lado direito do ideograma está representada a região envolvida em ganhos (verde) e ao lado esquerdo, as perdas (vermelho). A figura foi gerada a partir dos casos que apresentaram tais alterações (<i>Nexus v.5.0 -Biodiscovery</i>).....	56
Figura 7	Representação esquemática das regiões identificadas entre diferentes amostras de tumores múltiplos. A – As amostras 845T_A e 845T_B apresentaram 13 regiões em comum (seleção). B – As amostras 683T_A e 683T_B não apresentaram regiões em comum (<i>Nexus v.5.0 - Biodiscovery</i>)	65
Figura 8	Perdas genômicas identificadas nas amostras de pacientes com história de LU familiar. As linhas vermelhas ao centro da imagem representam as perdas entre os casos (em azul). A – Perda em 1p36.33 nas amostras 046T_A, 084T, 629T_B e 633T, onde estão mapeados genes como <i>GNB1</i> , <i>CALML6</i> , <i>TMEM52</i> , <i>KIAA1751</i> , <i>GABRD</i> e <i>PRKCZ</i> . B – Perda em 14q24.3 nas amostras 046T_A, 629T_B e 842T envolvendo vários genes, entre eles <i>DNAL1</i> , <i>PTGR2</i> , <i>VSX2</i> e <i>ABCD4</i> (<i>Nexus v.5.0 -Biodiscovery</i>).....	67
Figura 9	Representação esquemática da investigação das sondas gênicas nos casos que apresentaram as deleções mais frequentemente observadas. As flechas indicam os genes (círculo vermelho) e as sondas envolvidas. A - perda em 1p36.33 no caso 084T, mostrando a deleção da sonda A_14_P131252 que mapeia o gene <i>PRKCZ</i> envolvida em perda genômica nos casos 046T_A, 629T_B e 633T. B- perda em 14q24.3 no caso 842T, mostrando a sonda A_14_P200645 que mapeia o gene <i>VSX2</i> , envolvida em perda genômica comum nos casos 046T_A e 629T_B (<i>Nexus v.5.0 - Biodiscovery</i>)	68
Figura 10	Representação esquemática da cobertura das sondas correspondentes aos genes candidatos envolvidos em perdas genômicas entre as amostras de LU familiar. A - perda em 1p36.33, mostrando a deleção da	

	sonda A_14_P131252 que mapeia o gene <i>PRKCZ</i> com deleção de 65pb do íntron 4. B- perda em 14q24.3, mostrando a sonda A_14_P200645 que mapeia o gene <i>VSX2</i> , com deleção de 4pb do íntron 3 (<i>Nexus v.5.0 - Biodiscovery</i>)	69
Figura 11	A - Análise de agrupamento hierárquico das CNVs identificadas gerou dois grandes grupos de amostras: Grupo 1 (n=42) e Grupo 2 (n=38). B - Ganhos em 2q35, 7q22.1, 19p13.12- 13.11 e 19q13.32-q13.33 foram observados exclusivamente no Grupo 1 ($P < 0,01$). Perdas em 7q22.1-q22.2 foram presentes exclusivamente no Grupo 2 ($P < 0,01$) (setas) (<i>Nexus v.5.0 - Biodiscovery</i>).	71
Figura 12	Regiões de ganhos e perdas exclusivas aos grupos identificados. Grupo 1: amostras que apresentaram ganhos em 2q35, 7q22.1, 2q35, 7q22.1, 19p13.12-p13.11 e 19q13.32-q13.33. Grupo 2: amostras com perda em 7q22.1-q22.2 (<i>Nexus v.5.0 - Biodiscovery</i>)	72
Figura 13	Grupos 1 e 2 estatisticamente associados com tumores múltiplos e tumores não múltiplos, respectivamente (<i>Nexus v.5.0 - Biodiscovery</i>)	73
Figura 14	Análise citogenética para a sonda RP11-46J20 da região 7q22.3-q31.1. A- Metáfase obtida após cultura temporária de linfócitos de sangue periférico, confirmando o mapeamento da sonda em 7q (setas). B - Exemplo de núcleo interfásico de LU mostrando dois sinais da sonda teste (verde) e dois sinais do controle centrômero do cromossomo 7 (Dados obtidos de Perez, 2004).....	75
Figura 15	Os 1323 genes alterados quanto ao número de cópias nos LU estavam associados a várias funções, incluindo processo metabólico de esteróides, regulação da transcrição da RNA polimerase II, processo metabólico de lipídeos, mecanismos de <i>splicing</i> alternativo e principalmente genes associados à regulação da transcrição (<i>Nexus v.5.0 - Biodiscovery</i>).....	79

Figura 16	Funções biológicas associadas com as moléculas identificadas nas redes a partir dos genes mapeados em regiões de CNVs significativas. Os valores são denotados em $-\log(P \text{ value})$. A. Doenças; B. Funções Moleculares (IPA) 83	83
Figura 17	Vias canônicas geradas a partir da razão entre os genes mapeados em regiões de CNVs significativas e as vias depositadas no banco de dados IPA (IPA) 85	85
Figura 18	Representação gráfica da via de Sinalização IGF-1. Esta via têm sido associada aos mecanismos de proliferação celular em Leiomiomas Uterinos. Em verde, moléculas envolvidas em perdas genômicas e em vermelho, moléculas envolvidas em ganhos (IPA) 86	86
Figura 19	Representação gráfica da via de Sinalização TGF-beta. Estudos prévios associaram genes desta via com a característica fibróide dos Leiomiomas Uterinos. Em verde, moléculas envolvidas em perdas genômicas e em vermelho, moléculas envolvidas em ganhos (IPA)..... 87	87
Figura 20	Representação gráfica da via de Sinalização do Receptor de Estrógeno. Os estrógenos podem desencadear estímulos de proliferação celular pela via Ras e induzir a transcrição de genes alvos. Em vermelho, moléculas envolvidas em ganhos genômicos (IPA)..... 88	88
Figura 21	Análise comparativa entre as moléculas do Grupo 1 e Grupo 2. A e B – Doenças e funções moleculares similares e exclusivas associadas aos genes de cada grupo, respectivamente (IPA)..... 90	90
Figura 22	Análise de sobreposição de redes gênicas. As Redes 3 a 6 do Grupo 2 apresentaram sobreposição de funções entre as Redes 1 a 17 do Grupo 1. Também foram observadas geração de redes órfãs (Redes 18 a 25 - Grupo1 e Redes 1 e 2 - Grupo 2) (IPA)..... 91	91
Figura 23	Análise de agrupamento hierárquico não supervisionado entre as 51 amostras de LU e miométrio adjacente. A – Perfil de expressão gênica	

resultante da análise do SAM (Tusher et al., 2001). B – Dois subgrupos foram detectados: os miométrios adjacentes (MM) separaram-se dos Leiomiomas Uterinos (LU). Porém, não foi observado um agrupamento significativo das amostras em relação ao período menstrual, presença de tumores múltiplos ou únicos e história de LU na família (*TMeV v.4.5*) .. 95

- Figura 24** Redes geradas a partir das interações entre os genes diferencialmente expressos e o banco de dados IPA. A – Rede 1: genes associados com ciclo celular, câncer e doença do sistema reprodutivo; B - Rede 2: genes envolvidos com ciclo celular, doença hematológica e metabolismo de carboidratos. Em verde, genes com expressão diminuída e em vermelho, genes com expressão aumentada (IPA).....112
- Figura 25** As funções moleculares exercidas pelos genes diferencialmente expressos, em valores de $-\log(P \text{ valor})$. Observam-se as funções associadas com ciclo celular e proliferação celular com maiores valores (IPA)113
- Figura 26** Doenças associadas aos genes diferencialmente expressos, em valores de $-\log(P \text{ valor})$. Câncer, doença gastrointestinal e doença genética aparecem com valores próximos ao P valor máximo observado, seguidas por doença do sistema reprodutivo e doenças musculares e esqueléticas (IPA)114
- Figura 27** As vias canônicas associadas aos genes diferencialmente expressos estão descritas como valores de razão em relação à frequência das moléculas presentes no banco de dados. As principais vias desempenham funções associadas com reparo a erros de pareamento (*Mismatch Repair*) em eucariotos, regulação do *checkpoint* G2/M do ciclo celular em resposta a danos no DNA, regulação do *checkpoint* G1/S do ciclo celular, sinalização da apoptose mediada por MYC e regulação do ciclo celular e ciclinas (IPA).....115
- Figura 28** Via da regulação do *checkpoint* do ciclo celular em G1/S. A estimulação pelo TGF-beta, em resposta a danos no DNA, ativa a cascata de

	sinalização de moléculas envolvidas no <i>checkpoint</i> do ciclo celular e na transcrição de genes alvos. Em verde, moléculas codificadas por genes com expressão diminuída e em vermelho, moléculas codificadas por genes com expressão aumentada (IPA).....	116
Figura 29	Via Sinalização IGF-1. O IGF-1, que atua como fator de crescimento, ativa a cascata de sinalização Ras, PIK3 e JAK/STAT, que induz a transcrição de genes envolvidos com proliferação e crescimento celular. Em verde, moléculas codificadas por genes com expressão diminuída e em vermelho, moléculas codificadas por genes com expressão aumentada (IPA)	117
Figura 30	Etapas da análise dos dados de CGH array (genômicos) e de microarray de expressão (transcriptômicos) para obtenção dos genes significantes.....	118
Figura 31	Análise de segmentação realizada pelo <i>DNA Copy</i> . No eixo x, os valores de 0 a 40.000 representam o número de marcadores do algoritmo que representam todos os cromossomos. As cores alternativas representam os cromossomos (1 a 22, X e Y), no qual o primeiro grupo de pontos verdes indica o cromossomo 1. No eixo y, os valores de <i>log2</i> mostram as variações abaixo de zero (linha vermelha no eixo central), indicando regiões envolvidas em perdas e as variações acima de zero, indicam as regiões envolvidas em ganhos genômicos. Regiões sem oscilação do eixo central não apresentam alterações no número de cópias de DNA (<i>DNA Copy</i>)	119
Figura 32	Funções biológicas dos 3325 genes identificados pela análise de microarray. A categorização dos transcritos foi baseada em <i>log2ratio</i> com o <i>fold-change</i> ≤ -1.0 e ≥ 1.0 , indicando diminuição ou aumento de expressão, respectivamente. Quase metade dos genes com expressão aumentada (44,9%) possuem funções associadas com processos celulares e aproximadamente 38% dos genes com expressão diminuída, apresentam esta mesma categorial funcional. 74,4% dos genes com expressão diminuída estão relacionados a funções associadas com ciclo celular (<i>FUNNET</i>)	121

Figura 33	Agrupamento hierárquico não-supervisionado das amostras, considerando o perfil genômico (A) e transcriptômico (B) dos 75 genes moduladores. A Figura mostra a categorização das amostras considerando fase do ciclo menstrual (proliferativa e secretora), número de amostras avaliadas (2 ou mais amostras ou uma amostra) e diagnóstico de tumores múltiplos ou únicos (<i>TMeV v.4.5</i>).....	135
Figura 34	Genes moduladores identificados pelo CONEXIC. A - Padrão de ganhos e perdas dos 75 genes identificados por CGH array. B - Representação do perfil de expressão gênica dos mesmos 75 moduladores obtidos por análises globais de expressão gênica. Nas colunas, tanto em A como em B, as amostras estão dispostas na mesma ordem e nas linhas, estão os genes. Observar o perfil genômico e transcriptômico do mesmo gene na mesma amostra.....	136
Figura 35	Frequência de localização de genes ao longo dos cromossomos. A maioria dos genes (~60%) está mapeada em regiões intercromossômicas e em regiões teloméricas (30%). Em torno de 6% estão mapeados em regiões próximas aos telômeros	137
Figura 36	Redes geradas a partir das interações entre os 75 genes moduladores e o banco de dados IPA. A - Rede 1: genes associados com sinalização celular, metabolismo de ácidos nucleicos e bioquímica de pequenas moléculas; B - Rede 4: genes envolvidos com ciclo celular, doença hematológica e metabolismo de carboidratos. Em verde, genes com expressão diminuída e em vermelho, genes com expressão aumentada (IPA)	147
Figura 37	Doenças e funções associadas aos genes moduladores, em valores de $-\log(P \text{ valor})$. A- Câncer, doença gastrointestinal e do sistema reprodutivo aparecem com valores próximos ao P valor máximo observado. B - Ciclo celular, morfologia celular e proliferação e crescimento celular foram as funções com maior $-\log(P \text{ valor})$ (IPA)	148

Figura 38	Vias canônicas associadas com os genes moduladores. As cinco principais vias descritas foram sinalização do câncer de mama hereditário, via de ativação da protrombina, proteínas CHK no controle do <i>checkpoint</i> do ciclo celular, regulação do <i>checkpoint</i> G2/M do ciclo celular em resposta a danos no DNA e fibrose hepática (valores da razão sobre $-\log P$ valor) (IPA).....	149
Figura 39	Sinalização da fibrose hepática semelhante aos processos proliferativos e fibróticos dos LU. O aumento da atividade tirosina quinase de FGFR1, juntamente com o aumento de IGFBP5 em resposta a fatores de crescimento poderia desencadear a proliferação celular identificada nestes tumores, porém, o aumento da atividade de COL3A1, favoreceria o acúmulo de matriz extracelular e deposição de colágeno nas células do miométrio (IPA). Em vermelho, moléculas com expressão aumentada (IPA)	150

LISTA DE TABELAS

Quadro 1	Relação dos relatos em literatura que utilizaram análises de CGH array em Leiomiomas Uterinos, as diferentes plataformas utilizadas e os principais resultados.....	20
Tabela 1	Dados clínicos obtidos das 56 pacientes com Leiomiomas Uterinos	43
Tabela 2	Regiões significativas de perdas e ganhos genômicos detectadas por CGH array entre as amostras de Leiomiomas Uterinos	52
Tabela 3	Frequência de similaridade das alterações genômicas detectadas entre as diferentes amostras de tumores múltiplos de uma mesma paciente.....	58
Tabela 4	Descrição das alterações genômicas, genes e miRNAs em comum entre diferentes amostras de tumores múltiplos	59
Tabela 5	Regiões mínimas em comum significativas distribuídas entre as amostras de tumores múltiplos	64
Tabela 6	Deleções no braço longo do cromossomo 7 identificadas por CGH array e “validação” por FISH entre as amostras de Leiomiomas Uterinos.....	76
Tabela 7	Redes gênicas com os 10 <i>scores</i> mais altos, geradas a partir dos genes mapeados nas regiões de CNVs significativas geradas pelo IPA - <i>Ingenuity Pathway Analysis</i>	81
Tabela 8	Genes diferencialmente expressos nos Leiomiomas Uterinos em relação ao miométrio adjacente.....	97
Tabela 9	Redes e principais funções geradas a partir da associação entre os genes diferencialmente expressos analisados pelo sistema IPA (<i>Ingenuity Pathway Analysis</i>)	110
Tabela 10	Genes moduladores resultantes da análise integrada dos dados genômicos e transcriptômicos.....	123

Tabela 11	Genes identificados na análise integrada baseada nos achados de correlação inversa (ganho genômico/diminuição de expressão e perda genômica/ aumento de expressão gênica) e seus miRNAs preditos.....	128
Tabela 12	miRNAs mapeados em regiões de ganhos e perdas genômicas e o padrão de expressão de seus genes preditos.....	131
Tabela 13	Meta-análise dos genes moduladores detectados em estudos prévios de microarray em Leiomiomas Uterinos depositados no banco de dados GEO - <i>Genomic Expression Omnibus</i> (http://www.ncbi.nlm.nih.gov/geo/).....	139
Tabela 14	Genes identificados nos módulos de enriquecimento gênico.....	142
Tabela 15	Análise de enriquecimento de genes moduladores realizada pelo <i>GSEA - Gene Set Enrichment Analysis</i> (http://www.broadinstitute.org/gsea/index.jsp)	143
Tabela 16	Redes gênicas com os scores gerados a partir da associação dos 75 genes moduladores com o banco de dados <i>Ingenuity Pathway Analysis</i> (IPA).	146

LISTA DE SIGLAS E ABREVIATURAS

ANG	<i>Alto Nível de Ganhos</i>
CNV*	<i>Variações no Número de Cópias, do inglês Copy Number Variation</i>
CGH*	<i>Hibridação Genômica Comparativa, do inglês Comparative Genomic Hybridization</i>
CGH array*	<i>Comparative Genomic Hybridization based on Microarrays</i>
Cy3*	<i>Cyanine 3</i>
Cy5*	<i>Cyanine 5</i>
DNA*	<i>Deoxyribonucleic Acid</i>
dNTP*	<i>Deoxynucleoside Triphosphate</i>
DGV*	<i>Database Genomic Variation</i>
FISH*	<i>Hibridação in situ Fluorescente, do inglês Fluorescent in situ hybridization</i>
HCL*	<i>Hairy Cell Leukemia</i>
HLRCC*	<i>Hereditary Leiomyomatosis and Renal Cell Cancer</i>
HR-CGH*	<i>Hibridação Genômica Comparativa de Alta Resolução, do inglês High-Resolution Comparative Genomic Hybridization</i>
IHC	<i>Imunohistoquímica</i>
IMC	<i>Índice de Massa Corpórea</i>
IPA*	<i>Ingenuity Pathway Analysis</i>
LU	<i>Leiomiomas Uterinos</i>
LOH*	<i>Perda de Heterozigose, do inglês Loss of Heterozygosity</i>
LMS	<i>Leiomiosarcomas</i>
MCUL*	<i>Multiple cutaneous and uterine leiomyomatosis</i>
miRNA*	<i>Micro RNA</i>
MM	<i>Miométrios Adjacentes</i>
MEC	<i>Matriz Extracelular</i>
NCBI*	<i>National Center for Biotechnology Information</i>
NHEJ*	<i>Non-Homologous End-Joining</i>
PCR*	<i>Polymerase Chain Reaction</i>
PH	<i>Perda Homozigota</i>
qPCR*	<i>Quantitative Polymerase Chain Reaction</i>
qRT-PCR*	<i>Quantitative Reverse Transcriptase Polymerase Chain Reaction</i>

POMA*	<i>Paraneoplastic Opsoclonus Myoclonus Ataxia</i>
RNA*	<i>Ribonucleic Acid</i>
RT-PCR*	<i>Reverse Transcriptase Polymerase Chain Reaction</i>
SERMS*	<i>Selective Estrogen Receptor Modulator</i>
SPRMS*	<i>Selective Progesterone Receptor Modulator</i>
SNP*	<i>Single Nucleotide Polymorphisms</i>
STUMP*	<i>Smooth Muscle Tumors with Uncertain Malignant Potential</i>
TE	<i>Tampão Tris-EDTA</i>

*siglas ou abreviaturas derivadas do inglês.

SUMÁRIO

1	Introdução	1
1.1	Leiomiomas Uterinos	4
1.1.1	Epidemiologia.....	6
1.1.2	Suscetibilidade Genética	7
1.1.3	Fatores relacionados ao desenvolvimento dos Leiomiomas Uterinos	9
1.2	Aspectos Genéticos dos Leiomiomas Uterinos	12
1.2.1	Hibridação Genômica Comparativa (CGH).....	17
1.2.2	Hibridação Genômica Comparativa baseada em Microarrays (CGH array)	18
1.3	Perfil de Expressão Gênica Global.....	23
1.4	Integração de Dados Genômicos e Transcriptômicos.....	27
2	Objetivos.....	30
3	Material e Métodos	31
3.1	Casuística.....	31
3.1.1	Características clínicas e histopatológicas	31
3.2	Extração de DNA.....	32
3.3	Extração de RNA.....	32
3.4	CGH array – Marcação, hibridação e análise.....	33
3.5	Hibridação <i>in situ</i> fluorescente (FISH).....	35
3.6	Microarrays de Expressão - Marcação, hibridação e análise	36
3.7	Integração de dados genômicos e transcriptômicos.....	38
3.8	Análise funcional - IPA.....	40
4	Resultados	42
4.1	Características clínicas.....	42
4.2	Caracterização geral das alterações no número de cópias de DNA.....	44
4.2.1	Caracterização das regiões significativas.....	49
4.2.2	Tumores múltiplos	57
4.2.3	História familiar de Leiomiomas Uterinos	66
4.2.4	Agrupamento hierárquico	70
4.2.5	Investigação por FISH	74
4.3	Análise funcional dos genes mapeados em regiões significativas.....	79
4.3.1	Análise funcional dos genes mapeados em 2q, 7q, 19p e 19q.....	89
4.4	Perfil de expressão gênica global	92
4.4.1	Genes diferencialmente expressos	95
4.4.2	Análise funcional dos genes diferencialmente expressos.....	108
4.5	Análise integrada dos dados genômicos e transcriptômicos.....	118
4.5.1	CGH array	119
4.5.2	Perfil de expressão gênica global	120
4.5.3	Integração dos dados – CONEXIC.....	122

4.5.3.1	Predição de regulação gênica por miRNA.....	127
4.5.3.2	miRNAs mapeados em regiões de ganhos e perdas genômicas.....	130
4.5.3.3	Agrupamento hierárquico	134
4.5.3.4	Frequência de localização cromossômica.....	137
4.5.3.5	Meta-análise com dados públicos - GEO.....	138
4.5.3.6	Análise de enriquecimento de genes - GSEA.....	141
4.5.4	Análise funcional dos genes moduladores	144
5	Discussão.....	151
5.1	Alterações no número de cópias de DNA.....	152
5.2	Perfil de expressão gênica global	179
5.3	Integração dos dados genômicos e transcriptômicos.....	189
6	Conclusões	206
7	Referências Bibliográficas	208

ANEXOS

Anexo 1	Estudos prévios de expressão gênica em larga escala em Leiomiomas Uterinos.
Anexo 2	Dados clínicos e histopatológicos das pacientes com Leiomiomas Uterinos utilizadas neste estudo.
Anexo 3	Regiões envolvidas em alto nível de ganhos e perdas homozigotas distribuídas entre as amostras de Leiomiomas Uterinos.
Anexo 4	Genes e miRNAs mapeados nas regiões significativas de ganhos e perdas genômicos nas amostras de Leiomiomas Uterinos.
Anexo 5	Genes identificados na análise integrada. Os genes estão separados de acordo com o padrão de ganhos e perdas genômicas e expressão aumentada ou diminuída entre as amostras.

ARTIGO

Artigo 1	An integrative genomic and transcriptomic analysis reveals potential targets associated with cell proliferation in Uterine Leiomyomas.
-----------------	--

1. Introdução

A finalização do projeto de sequenciamento do genoma humano trouxe uma nova visão para os pesquisadores, ampliando a possibilidade de caracterizar as variações comuns no genoma e alterações associadas às doenças humanas (para revisão Pinto et al., 2007). No entanto, mesmo após dez anos da conclusão do projeto genoma, há ainda muitos mistérios a serem desvendados.

O genoma em geral, especialmente o humano, pode apresentar desde grandes rearranjos cromossômicos, visíveis ao microscópio, até alterações de nucleotídeos únicos. Nos últimos anos, vários estudos revelaram uma abundância de variação do número de cópias de segmentos de DNA submicroscópicas, desde kilobases (Kb) até megabases (Mb). As alterações complexas, coletivamente denominadas variações do número de cópias (CNVs) do DNA, são encontradas em todos os indivíduos e em outras espécies de mamíferos. A CNV, por definição, é um segmento de DNA maior ou igual a 1 Kb presente em número de cópias variável em comparação a um genoma de referência. As CNVs podem envolver pequenas modificações no genoma, como duplicações em tandem, ou alterações complexas, como deleções, inserções e duplicações (para revisão Redon et al., 2006). No entanto, as novas tecnologias de sequenciamento de alto desempenho têm evidenciado regiões de CNVs envolvendo segmentos de 50pb, que podem, ou não, refletir em uma alteração fenotípica (Mills et al., 2011). Os resultados do projeto *1000Genomes* (<http://www.1000genomes.org/>), que tem como objetivo sequenciar o genoma completo de aproximadamente 1200 indivíduos, tem mostrado resultados parciais demonstrando a presença de variações moleculares originadas, por exemplo, por duplicação gênica, que ocorrem na ausência de enfermidades (Wheeler et al., 2008). Porém, até o momento, uma pequena porcentagem das CNVs identificadas parecem ter efeitos fenotípicos.

A primeira associação de uma CNV com um fenótipo foi descrita 70 anos atrás, com a duplicação do gene *Bar* em *Drosophila melanogaster* associada com o fenótipo “olho Bar”

(Bridges, 1936). Acredita-se que as CNVs possam influenciar a expressão de genes levando a variações fenotípicas (Inoue e Lupski, 2002; Shaw-Smith et al., 2004) e podem conferir risco para a doenças complexas, como o câncer.

Muitos oncogenes e genes supressores de tumor são afetados por CNVs, como por exemplo, *LPP*, *MLLT3*, *MEN1*, *APC*, *TNFRSF25* e *BCAS1*, os quais estão envolvidos com alterações no número de cópias de DNA nos seguintes tumores, lipoma, leucemia, neoplasia endócrina múltipla tipo1, polipose adenomatosa familiar, mieloma múltiplo e câncer de mama, respectivamente (Conrad et al., 2006; de Smith et al., 2007; Wong et al., 2007). Estudos para determinar as consequências das CNVs para susceptibilidade ao câncer estão em andamento, mas já temos pelo menos um importante exemplo de uma associação. Uma CNV (envolvendo a deleção do gene *UGT2B17*), encontrada em cerca de 11-12% de indivíduos saudáveis, apresenta uma associação significativa com o risco de câncer de próstata em caucasianos (Park et al., 2006). Acredita-se que a deleção do gene envolvido no metabolismo do andrógeno poderia levar ao aumento dos níveis séricos de testosterona, permitindo o desenvolvimento do câncer. O potencial clínico das CNVs é imenso, entretanto ainda permanecem os desafios para a compreensão da relação entre alterações genômicas e fenótipos (Lee e Scherer, 2010).

O desenvolvimento de tecnologias que permitem a avaliação global de CNVs, como por exemplo, a metodologia de CGH array, descrita em 1998 (Pinkel et al., 1998) associadas às tão consagradas técnicas de microrray de expressão, que quantificam os níveis de expressão de milhares de genes ao mesmo tempo, tem despertado considerável interesse na comunidade científica, inclusive nas pesquisas sobre o câncer (Carter, 2007; Akavia et al., 2010).

Os Leiomiomas Uterinos (LU) representam um modelo interessante para a investigação da associação de CNVs com modificações nos níveis de expressão gênica. Estes tumores benignos apresentam capacidade proliferativa, porém não invasiva. Outra característica é que geralmente são encontrados, em média, seis tumores por paciente e os

dados em literatura indicam que cada um deles surge independentemente no útero. Além disso, o tecido normal adjacente (miométrio) pode ser usado para análises comparativas de expressão gênica (Leppert et al., 2006; Hodge et al., 2009). Mesmo com estas particularidades, há um número limitado sobre a etiologia dos Leiomiomas Uterinos. Estudos citogenéticos têm revelado que apenas um subgrupo destes tumores apresenta alterações cromossômicas numéricas e estruturais (para revisão Canevari e Rogatto, 2007). Vários relatos que utilizaram microarrays de expressão evidenciaram classes de genes envolvidos em vias biológicas particulares (para revisão Dimitrova et al., 2009). Contudo, são escassos os estudos de associação entre CNVs e os níveis transcricionais entre os Leiomiomas Uterinos (Vanharanta et al., 2005; Hodge et al., 2009; Zavadil et al., 2010).

1.1 Leiomiomas Uterinos

Leiomiomas Uterinos são os tumores pélvicos benignos mais frequentes do trato genital feminino. Carl Von Rokitansky em 1860 introduziu o termo fibroma e Virchow em meados de 1880 modificou o termo para leiomioma, por ser um tumor com derivação da musculatura lisa miometrial (FEBRASGO, 2004). Apesar de o termo *leiomioma* ser mais utilizado entre os clínicos, estes tumores também são chamados de fibróides, fibromiomas, leiomiomas, miofibromas e coloquialmente de miomas (Zaloudek e Norris, 1981; Cramer e Patel, 1990). Constituem a neoplasia mais comum do útero, acometendo cerca de 25 a 30% das mulheres em idade reprodutiva, sendo diagnosticado principalmente em mulheres em idade superior a 30 anos (Zydan et al., 2007).

Os leiomiomas são tumores de origem mesenquimal e já foram relatados em outras regiões fora do útero, como leiomiomas de ovários, de rim, de estômago, de cérebro, de esôfago, de pele e angioleiomiomas da região selar (Kandalaf e Esteban, 1992; Garcia-Torres et al., 2000; Prenzel et al., 2006; Xu et al., 2010; Xiao et al., 2010; Choi et al., 2011; Lehtonen, 2011).

Além das células do músculo liso, a matriz extracelular dos LU consiste principalmente em colágeno, fibronectina e proteoglicanos (Parker, 2007). O colágeno é a principal proteína constitutiva desta matriz e os colágenos tipo 1 e 3 parecem ser os mais comuns nestes tumores (Stewart et al., 1994; Sozen e Arici 2002). Os leiomiomas são bem vascularizados, irrigados por uma complexa rede de vasos sanguíneos, similares a muitos outros tumores, incluindo as neoplasias malignas dependentes do desenvolvimento de um suprimento de sangue para a sua sobrevivência (Nowak et al., 2001).

Em torno de 80% das mulheres com LU são clinicamente assintomáticas, geralmente por apresentarem tumores pequenos, de crescimento lento e com localização subserosa, ou seja, aquele mais externo a parede do útero (Laser et al., 2010). Os LU sintomáticos estão associados com alterações clínicas como sangramento uterino anormal, dores abdominais,

infertilidade, aborto espontâneo, perdas prematuras e disfunção urinária (Coronado et al., 2000; FEBRASGO, 2004). A incidência e a gravidade dos sintomas variam em relação ao tamanho, ao número e a localização destes tumores. Tumores gigantes e múltiplos parecem estar associados com a gravidade dos sintomas (Stovall, 2001; Peng et al., 2008) e freqüentemente são encontrados tumores múltiplos no mesmo útero, sendo cada tumor clonalmente distinto (Canevari et al., 2005a; Viswanathan et al., 2007).

Os LU são classificados morfológicamente de acordo com a localização no útero. Os LU *submucosos* localizam-se na região submucosa, abaixo da mucosa uterina, adjacente à cavidade uterina; o tipo *intramural* fica completamente dentro da parede uterina, o subseroso, localiza-se abaixo da serosa, distorcendo o contorno da superfície externa do útero (Cotran et al., 1989) e o tipo *peduncular* encontra-se ligado ao útero por um pêndulo (FEBRASGO, 2004). O subtipo intramural é o mais comum e é responsável pelo aumento irregular do fundo do útero. Como esses tumores aumentam em tamanho, podem, eventualmente, expandir-se na cavidade abdominal, através da superfície do útero, tornando-se subserosos. Os tumores submucosos por estarem localizados dentro da cavidade uterina e exatamente abaixo do endométrio são freqüentemente associados a muitos sintomas típicos, como por exemplo, dor pélvica e sangramento uterino excessivo (Cook et al., 2010).

A classificação histológica dos leiomiomas depende do número e morfologia das células musculares. Os LU podem ser divididos em: *típico* (ou usual), que apresentam células uniformes de músculo liso com núcleo alargado mostrando uma cromatina fina dispersa e pequeno nucléolo e raras mitoses. Os subtipos *epitelióide*, *mixóide*, *vascular*, *schwanoide*, *lipoleiomioma* são tumores benignos que se assemelham ao leiomioma típico exceto por apresentarem células adiposas presentes juntamente com as células de músculo liso do miométrio, vascular e formas mais raras, com células hematopoiéticas (FEBRASGO, 2004).

As neoplasias mesenquimais malignas acometem aproximadamente 3% das doenças malignas do útero e aproximadamente um terço são leiomiosarcomas (LMS) (para revisão Fadare, 2011). Em um estudo com uma série de 1332 mulheres que sofreram histerectomia

ou miomectomia para LU sintomáticos, a frequência de sarcomas uterinos (LMS, sarcoma do estroma endometrial e tumor mesenquimal mixóide) foi de 0,23% (Parker et al., 1994). Parece existir um consenso de que os LU raramente progridem para fenótipos malignos. Apenas 0,1% dos casos de LU progridem para LMS em mulheres acima de 40 anos e maioria destes tumores surge *de novo* em mulheres nesta faixa etária, o que parece não estar relacionado à deposição de estrógeno no organismo (para revisão Lee et al., 2005). Existe um número limitado de dados sobre estas duas entidades. Entretanto, várias evidências clínico-patológicas, citogenéticas e moleculares indicam que os LU e LMS são entidades biológicas diferentes. Existem evidências mais atuais indicando que os LMS podem surgir de áreas benignas (Mittal et al., 2000; Mittal et al., 2007). Estudos moleculares recentes evidenciaram que LMS que apresentam áreas *leiomioma-like* compartilham com os LU uma variedade de alterações citogenéticas, o que sugere que LMS podem ter origem destas áreas “benignas” (Mittal et al., 2009). Os LU nem sempre são facilmente distinguíveis dos LMS e ambos podem estar presentes manifestando os mesmos sintomas e características clínicas, sendo definidos apenas após um diagnóstico diferencial, pela associação de dados de anamnese, exames ginecológicos e principalmente, baseado no índice mitótico (Parker, 2007).

1.1.1 Epidemiologia

Os Leiomiomas Uterinos, embora de natureza benigna, são considerados como um importante problema de saúde pública, sendo a principal indicação para histerectomia. A histerectomia é considerada a segunda cirurgia mais frequente entre as mulheres em idade reprodutiva (Farquhar e Steiner, 2002; Zydan et al., 2007; Jacoby et al., 2010). O governo Americano despende anualmente em torno de U\$ 2,1 bilhões com os procedimentos de histerectomias e miomectomias (Hodge et al., 2008). Em 2005, o governo Brasileiro gastou cerca de R\$ 67,5 milhões realizando 112.200 histerectomias e até o mês de março de 2011, já

foram despendido cerca de R\$ 3 milhões com os procedimentos totais de histerectomias (<http://www2.datasus.gov.br/DATASUS/index.php> - código SUS 0409060135).

Estudos epidemiológicos indicam como os fatores de risco associados ao desenvolvimento destes tumores idade jovem (35-54 anos), obesidade, raça africana, menarca em idade precoce, gravidez tardia, nuliparidade ou número pequeno de paridade, história familiar de LU e o uso de hormônios exógenos, como contraceptivos orais e terapia de reposição hormonal (para revisão Peddada et al., 2008). Entre as mulheres com LU acima de 35 anos e que nunca engravidaram, mais de 60% são afro-americanas e 40% são brancas (Day-Bird et al., 2003). Em relação à etnia, mulheres afro-americanas têm risco 3x maior em desenvolver LU em relação às caucasianas e asiáticas, porém, a causa desta disparidade permanece desconhecida (Jacoby et al., 2010). O consumo de cigarro parece causar um efeito “protetor” para o desenvolvimento de LU, pela redução do estrógeno circulante, levando a diminuição da exposição do miométrio aos estrógenos (Moorehead e Conard, 2001). Apesar do impacto na saúde pública, da incidência extremamente alta e do conhecimento dos fatores de risco, a etiologia e a patogênese dos LU não são totalmente conhecidas.

1.1.2 Suscetibilidade Genética

Além dos fatores ambientais, estudos epidemiológicos têm sugerido uma predisposição herdada para o desenvolvimento dos Leiomiomas Uterinos. Vários estudos indicam a existência de uma forte influência genética para o desenvolvimento destes tumores, como predisposição racial, agregação familiar e maior frequência de tumores entre gêmeas monozigóticas (para revisão Canevari e Rogatto, 2007).

Estudos de agregação familiar mostram que mulheres com história de LU na família possuem o risco de 1,5 a 3,5 vezes maior de desenvolver Leiomiomas Uterinos (Vikhlyeva et al., 1995). Em um estudo de 2005, Okolo et al. compararam as características clínicas de mulheres com LU esporádicos e mulheres com LU familiar. As mulheres com história familiar

de LU apresentaram piores manifestações clínicas, como maior aumento do volume abdominal, dismenorréia, menorragia e história de outros tumores na família.

A manifestação dos LU em pacientes com síndromes hereditárias raras reforça a presença de um componente hereditário nestes tumores. O gene *FH* (hidratase fumarato), mapeado em 1q42.1, codifica uma proteína componente do complexo enzimático do ciclo de Krebs. Tem sido relatada uma associação entre a presença de mutações germinativas no gene *FH* e predisposição aumentada a Síndrome de leiomiomas uterinos e cutâneos múltiplos (MCUL) ou a Síndrome de leiomiomatose hereditária e câncer de células renais (HLRCC). A Síndrome MCUL é uma doença autossômica dominante que afeta tanto homens como mulheres. Os homens portadores da mutação apresentam leiomiomas cutâneos e as mulheres, além dos leiomiomas cutâneos, apresentam também Leiomiomas Uterinos. Na Síndrome HLRCC, as mulheres com LU possuem uma predisposição aumentada para o desenvolvimento de carcinoma papilífero de células renais tipo II (Martinez-Mir et al., 2003; Lehtonen et al., 2004; Stewart et al., 2008).

De uma maneira geral, a existência de um fator herdado na etiologia dos LU é sugerida pela presença de tumores múltiplos no mesmo útero, que pode indicar que anormalidades genéticas herdadas predispoem ao desenvolvimento dos LU em algumas mulheres (Mao et al., 1999). Considerando que multiplicidade tumoral é uma característica comum de síndromes de predisposição genética ao câncer, pacientes com LU múltiplos poderiam portar um alelo mutante de um determinado gene que conferiria uma predisposição ao desenvolvimento destes tumores. No entanto, até o momento, o gene *FH* foi o único descrito em literatura como associado aos LU de origem familiar.

Recentemente, Pei-Chieng et al. (2011) realizaram um estudo *genome-wide* avaliando 457 mil SNPs (*Single Nucleotide Polimorphisms*) em sangue periférico de 1.607 pacientes que foram diagnosticadas com LU e 1.428 mulheres controle. Em adição, selecionaram um grupo muito maior de pacientes para validar os achados (3.466 casos adicionais e 3.245 controles). Os autores identificaram três *loci* significativamente associados com os genes *SLK* e *OBFC1*

(10q24.33), *BET1L*, *RIC8A*, *ODF3* e *SIRT3* (11p15.5) e *TNRC6B* (22q13.1), revelando novos genes de suscetibilidade aos Leiomiomas Uterinos.

A identificação de novos genes que possam estar envolvidos na etiologia e ou predisposição ao desenvolvimento destes tumores pode ser útil no aconselhamento genético e no delineamento de novas estratégias terapêuticas e diagnósticas.

1.1.3 Fatores relacionados ao desenvolvimento dos Leiomiomas Uterinos

Embora os principais fatores que levam ao desenvolvimento dos LU ainda serem desconhecidos, existem indícios que apontam o estrógeno e a progesterona como fatores promotores da doença, estimulando o crescimento destes tumores (Rein et al., 1995; Yoshida et al., 2010). Entre as várias funções destes esteróides, sua ação no útero é responsável pela regulação da proliferação das células do miométrio e assim, podem influenciar a suscetibilidade destas células a adquirir alterações genéticas (Barbieri e Andersen, 1992; Swartz et al., 2005).

A resposta dos LU ao estrógeno é bem estabelecida, contudo, o impacto de sua ação e contribuição para o desenvolvimento destes tumores ainda permanece desconhecido (Bahamondes et al., 2008). O crescimento rápido desses tumores durante a gestação, a regressão dos mesmos após a menopausa, além da diminuição de crescimento sob ação de drogas como análogos de GnRH, danazol e antiprogestínicos, enfatizam o papel dos estrógenos e progesterona no desenvolvimentos dos Leiomiomas Uterinos (Schwartz, 2001; Bozzini et al., 2003). Estes tumores também apresentam expressão aumentada da enzima aromatase P450, que permite que as células sintetizem seu próprio estradiol endógeno (Shozu et al., 2002).

Outra evidência que apóia a dependência estrogênica dos LU está no fato de que quando os níveis dos receptores de estrógeno e progesterona estão elevados nos tumores comparados ao miométrio adjacente, estes são mais responsivos aos hormônios (Bakas et al.,

2008). Entretanto, em LU de mulheres pós-menopáusicas, a expressão do gene receptor de progesterona foi menor nos LU quando comparados ao miométrio adjacente (Strissel et al., 2007). Um estudo de nosso grupo mostrou que em LU de mulheres pré-menopáusicas, houve associação entre o aumento da expressão gênica e protéica dos receptores de estrógeno alfa (*ESR1*) e progesterona (*PGR*) (Cirilo et al., manuscrito em preparação). Recentemente Shaik et al. (2011) mostraram que polimorfismos nos genes *ESR1* e *MTCYB* foram associados com expressão aumentada dos respectivos transcritos em LU de mulheres pré-menopáusicas, contribuindo para a tumorigênese destes tumores. Elementos regulatórios da expressão do gene receptor de estrógeno alfa (*ESR1*) podem contribuir para a responsividade dos estrógenos em Leiomiomas Uterinos. Wei et al. (2011) avaliaram a expressão do gene *HDAC6* e *ESR1* em LU comparados ao miométrio adjacente e revelaram correlação de expressão positiva entre os dois genes. Além disso, os autores realizaram estudos *in vivo* e demonstraram que o silenciamento de *HDAC6* leva a diminuição de expressão de *ESR1*, indicando que *HDAC6* pode ser um alvo terapêutico para o controle do crescimento dos Leiomiomas Uterinos.

O papel da progesterona na etiologia dos LU ainda é controverso. Alguns estudos sugerem que a progesterona pode inibir a formação e o crescimento dos LU (para revisão Moorehead e Conard, 2001), porém, outros estudos sugerem que a progesterona estimula o crescimento tumoral (para revisão Walker e Stewart, 2005). O aumento da taxa mitótica em LU durante a fase secretora do ciclo menstrual (quando existem picos nos níveis de progesterona) juntamente com dados clínicos (que indicam que as progestinas inibem o efeito terapêutico de agonistas de GnRH) apoiam a hipótese de que a progesterona é o principal hormônio associado com o crescimento dos Leiomiomas Uterinos (Kawaguchi et al., 1985; Rein et al., 1995). Ishikawa et al. (2010) estabeleceram um modelo xenográfico de LU com características similares de humanos. Os animais desenvolveram LU em resposta ao estradiol (estrógeno) e progesterona, através da proliferação celular e aumento de volume tumoral pela deposição de componentes da matriz extracelular. Contudo, o volume do tumor

somente foi reduzido após tratamento com o RU486, um anti-progestágeno e não respondeu á tratamentos com anti-estrogênicos. Os autores mostraram, pela primeira vez, que os estrógenos não são os agentes principais associados com o crescimento dos LU, contudo, ele induz a expressão de *PGR*, que permite a ação da progesterona, sendo responsável pelo aumento do volume dos Leiomiomas Uterinos.

O desenvolvimento de novas drogas para o tratamento dos LU, tem se baseado, principalmente, na ação dos estrógenos e progesterona. Atualmente, existem os moduladores seletivos dos receptores de estrógeno e progesterona (*SERMS* e *SPRMS*, respectivamente) também utilizados para o tratamento de câncer de mama e endometriose (Wu et al., 2007; Yoshida et al., 2010). O sucesso terapêutico destes moduladores tem sido alcançado principalmente pelos *SPRMS*, indicando que tratamentos direcionados aos receptores de progesterona possuem maior eficácia.

Durante os anos reprodutivos o miométrio é ciclicamente exposto a níveis hormonais relativamente altos, entretanto, a maioria das mulheres não desenvolve LU, o que sugere haver influência de outros fatores associados a gênese destes tumores. Experimentos utilizando linhagens de células de músculo liso mostram que os efeitos desses hormônios esteróides na proliferação e diferenciação celular são mediados pela produção de citocinas e fatores de crescimento, que exercem efeitos autócrinos ou parácrinos no útero. Dentre esses mediadores, encontram-se o fator de crescimento epidermal (EGF), fator de crescimento insulina-like (IGF), fator de crescimento transformante beta (TGF-beta), fator de crescimento de ligação a heparina (HB-EGF), fator de crescimento derivado de plaquetas (PDGF), fator de crescimento de fibroblastos (FGF), interleucinas e intérferons (para revisão Duhan e Sirohiwal, 2010). Estes fatores de crescimento não somente regulam a proliferação, diferenciação, apoptose e a produção de matriz extracelular das células de músculo liso como também regulam a proliferação e migração das células endoteliais vasculares (Nowak et al., 2001). Novos tratamentos que atuam nestes fatores de crescimento podem gerar resultados

satisfatórios levando a regressão e ou diminuição da proliferação celular em Leiomiomas Uterinos.

1.2 Aspectos Genéticos dos Leiomiomas Uterinos

O crescimento celular normal é garantido pelo equilíbrio entre a proliferação e morte celular. A idéia de que a formação dos tumores é resultante de um genoma alterado foi descrita inicialmente em 1914, quando o citologista alemão Boveri propôs pela primeira vez que alterações cromossômicas poderiam levar uma célula normal a apresentar uma proliferação anormal (para revisão, Harris, 2008). Alterações envolvendo oncogenes ou genes supressores de tumor podem levar a produção de proteínas alteradas, em quantidade ou estrutura, resultando, entre outros mecanismos, em aumento de sua atividade e mesmo perda de sua função.

Estudos de clonalidade indicam que os LU se originam de alterações somáticas que ocorrem em células do miométrio resultando em uma perda progressiva da regulação do crescimento e subsequentemente expansão clonal. Contudo, as regiões críticas responsáveis pela transformação neoplásica da musculatura uterina ainda permanecem indeterminadas (Canevari et al., 2005a; Cai et al., 2007).

Baseados em evidências citogenéticas, Hu e Surti (1991) relataram que devido a grande heterogeneidade genética observada entre os LU, estes tumores poderiam ser classificados em oito diferentes subgrupos: 1) LU portadores de rearranjos estruturais envolvendo os cromossomos 12 e 14, 2) rearranjos envolvendo o 12 e outros cromossomos, quando o 14 está normal, 3) alterações no 14 e outros cromossomos, sem o envolvimento do 12, 4) deleções intersticiais no cromossomo 7, com ou sem outros rearranjos, 5) rearranjos entre os cromossomos 1 e 2, 6) outros rearranjos não-randômicos, 7) rearranjos numéricos, incluindo trissomias e 8) tumores apresentado cariótipos normais.

Segundo a mais recente atualização do *Mitelman Database* até o momento, já foram cariotipados 468 casos de Leiomiomas Uterinos (<http://cgap.nci.nih.gov/Chromosomes/Mitelman>). Estas análises citogenéticas revelaram achados consistentes entre os tumores. Estima-se que aproximadamente 40% dos LU possuem anormalidades cromossômicas que variam entre as pacientes e os diferentes tumores analisados (Ligon e Morton, 2000). Alguns estudos têm mostrado que a presença de rearranjos citogenéticos correlaciona-se, por exemplo, com LU gigantes e a localização anatômica do tumor (Rein et al., 1998; Brosens et al., 1998). Segundo Hennig et al. (1999), tumores com rearranjos envolvendo o cromossomo 12 são geralmente maiores que aqueles que apresentam deleções no cromossomo 7 mas são equivalentes em tamanho aos tumores cariotipicamente normais. Aparentemente, a presença de algumas características histológicas peculiares podem estar associadas a subtipos específicos de LU com alto índice mitótico. Deleções do braço curto do cromossomo 1 [del(1)(p11-12-pter)] foram observadas em 9 casos entre mais 800 amostras de LU. Estes tumores apresentavam características histológicas atípicas, incluindo hiperplasia e/ou atipia celular (Christacos et al., 2006). Entre LU plexiformes, foram descritos cariótipos complexos, evento que geralmente não é observado em LU típicos. A análise do perfil de expressão gênica global revelou a presença de um agrupamento independente dos LU plexiformes quando avaliados LU e LMS (Hodge et al., 2008).

Os tipos de rearranjos cromossômicos, como translocações, trissomias e deleções, encontradas entre os LU sugerem múltiplos mecanismos genéticos associados com o crescimento do tumor. As translocações cromossômicas podem levar a alterações da região promotora de genes direcionadores ou gerar novos genes quiméricos (a partir de fusões gênicas) que contribuem para o desenvolvimento e ou manutenção dos tumores (para revisão MacConaill e Garraway, 2010). As trissomias geralmente aumentam os níveis dos transcritos pelo aumento da dose do gene e as deleções cromossômicas mais freqüentemente resultam na perda da função do gene (para revisão Gross e Morton, 2001). Baseados nos

estudos citogenéticos, vários genes candidatos têm sido propostos como envolvidos ao desenvolvimento dos Leiomiomas Uterinos.

A primeira anormalidade citogenética associada especificamente com LU e que está presente em aproximadamente 20% desses tumores é a translocação envolvendo o braço longo dos cromossomos 12 e 14: t(12;14)(q14-q15; q23-q24) (Heim et al., 1988). O gene *HMGA2*, membro do grupo de genes que codifica proteínas de alta mobilidade (HMG), tem sido o alvo molecular central deste rearranjo. Os pontos de quebra envolvem o intron 3 do gene *HMGA2* (>140 kb) e resulta em fusão gênica levando a alteração de sua expressão. Como resultado deste rearranjo os transcritos codificados são quiméricos contendo os primeiros três exons do gene *HMGA2* em conjunto com a sequência ectópica de outros genes (Kazmierczak et al., 1995; Kools e Van de Ven., 1996). O gene *RAD51L1* mapeado em 14q23 (membro da família RAD51 associada a reparo a danos no DNA) está preferencialmente envolvido na fusão com o gene *HMGA2* em Leiomiomas Uterinos (Takahashi et al., 2001). Além disso, estudos moleculares detectaram expressão aumentada do *HMGA2* em LU comparados com miométrio normal adjacente (Arslan et al., 2005; Ingraham et al., 2006). Outro gene mapeado no cromossomo 14 e que pode fusionar-se ao *HMGA2* é o *ESR2* (14q22). O gene *ESR2* não é expresso constitutivamente em quantidades significativas no útero, porém, este tipo de translocação pode levar a ativação deste gene em LU (Nilbert and Heim, 1990; Ligon and Morton, 2001). Outros tumores associados ao tecido mesenquimal também apresentam rearranjos cromossômicos envolvendo a região 12q13-15, como lipomas, hamartomas condróides pulmonares, fibroadenomas de mama, pólipos endometriais, adenomas pleomórficos de glândula salivar e angiomixoma de vulva (para revisão Fusco e Fedele, 2007).

Deleções intersticiais do braço longo do cromossomo 7 estão presentes em aproximadamente 17% dos LU com cariótipos anormais e provavelmente, é uma das anormalidades cromossômicas mais investigadas nestes tumores (Pandis et al., 1990; Hodge et al., 2009). Os pontos de quebra envolvem 7q21-7q36 com região mínima comum em 7q22

(Sargent et al., 1994). Este rearranjo pode ser o único presente nas análises cromossômicas ou ser simultaneamente observado com a t(12;14) (Ozisik et al., 1993b). Os estudos para a identificação de gene(s) críticos em 7q envolvidos no desenvolvimento de LU são dificultados pela heterogeneidade do tamanho das deleções e também da região ser rica em genes, incluindo aqueles envolvidos no processo de desenvolvimento (*DLX5*, *DLX6*), no metabolismo do colágeno (*COL1A2*, *PCOLCE*), no reparo a danos do DNA (*PMS218*), no ciclo celular (*ORC5L*) e membro das proteínas HMG (*LHFPL3*), além de outros igualmente prováveis (para revisão Gross e Morton, 2001; Sell et al., 2005; Hodge et al., 2009). Alguns desses genes são conhecidos por influenciarem o crescimento celular e, conseqüentemente, são avaliados como genes candidatos posicionais em Leiomiomas Uterinos. Embora subgrupos de LU com deleções do cromossomo 7 mostrem LOH (*Loss of heterozygosity*) e ou expressão reduzida de *CUTL1*, *ORC5L*, *SERPINE1* ou *PCOLCE*, não foram identificadas deleções gênicas consistentes pelas técnicas utilizadas em citogenética clássica e molecular ou alterações na expressão de transcritos específicos (Sourla et al., 1996; Zeng et al., 1997; Quintana et al., 1998). Mais recentemente, Hodge et al. (2009) através estudos de associação de dados genômicos e transcriptômicos em subgrupos de LU com deleções citogenéticas em 7q, revelaram o envolvimento dos genes supressores de tumor *HPB1* (que codifica um inibidor da proliferação celular) e o *RINT1* (que codifica uma proteína que atua na manutenção da integridade da mitose).

Alterações envolvendo o braço curto do cromossomo 6 também foram relatadas em menos de 10% dos LU cariotipicamente anormais (para revisão Takahashi et al., 2002). Em 6p21 está mapeado o gene *HMGA1*, membro da família HMG, (Kazmierczak et al., 1996; Williams et al., 1997). A expressão aumentada deste gene tem sido detectada em LU com rearranjos em 6p21 quando comparada com o miométrio normal adjacente (Arslan et al., 2005; Nezhad et al., 2010). Rearranjos envolvendo 6p21 também têm sido identificados em outros tumores mesenquimais (Reeves, 2000; Tallini et al., 2000).

Em torno de 5% dos LU apresentam rearranjos envolvendo o braço longo do cromossomo 10, caracterizando outro subgrupo destes tumores. A maioria dos rearranjos é translocação balanceada entre 10q22 e uma variedade de outros cromossomos, incluindo 4, 6, 12 e 17 e menos frequentemente a *del(10)(q22q24)* (Ozisik et al., 1993a). Um candidato putativo mapeado entre estes pontos de quebra no cromossomo 10 é o gene *MORF*. Este gene codifica um membro da família MYST de acetiltransferases de histonas que geralmente encontra-se envolvidos em pontos de quebra em alguns sub-tipos de leucemia mielóide aguda (Moore et al., 2004).

Aproximadamente de 3% dos LU, apresentam monossomia do cromossomo 10, como também deleções especialmente no sítio frágil 10q23 (Ozisik et al., 1994). Nesta região, está mapeado o gene supressor tumoral *PTEN* (Steck et al., 1997). Para nosso conhecimento, não encontramos estudos associando esta deleção e atividade de *PTEN*, contudo, a literatura relata a ação deste gene em Leiomiomas Uterinos. Em um estudo avaliando a expressão protéica de *PTEN* em tumores do músculo liso, amostras de LU típicos e atípicos, STUMP (tumores do músculo liso com potencial incerto) e LMS, a marcação de *PTEN* não foi estatisticamente significativa nestes tumores (Gökaslan et al., 2005), no entanto, outro estudo já revelou marcação fraca de *PTEN* em amostras de LU comparadas ao miométrio adjacente (Kayisli et al. 2007). Mais recentemente, Jeong et al. (2010) avaliaram a expressão *PTEN* e *PPAR* γ em linhagens de LU sob o estímulo de 17beta-estradiol e demonstraram expressão aumentada de ambas proteínas nos LU em relação ao miométrio adjacente. Os autores concluíram que os estrógenos podem levar a expressão aumentada de *PTEN* e *PPAR* γ , contribuindo para a proliferação celular em Leiomiomas Uterinos.

Uma variedade de outras anormalidades citogenéticas foi relatada em menor freqüência. Estas incluem rearranjos dos cromossomos X, 1, 3 e 13 e trissomia do cromossomo 12 (para revisão Hodge et al., 2007). Os casos envolvendo o cromossomo X, preferencialmente a região Xp11-p22, incluem deleções, translocações com os cromossomos 5 e 12 e inversões entre as bandas Xp22q13 (Ozisik et al., 1992). Rearranjos do cromossomo

1 geralmente envolvem formação de cromossomos em anel e, com frequência, ocorrem concomitantemente com outras alterações cromossômicas (Nilbert, et al., 1988). Um estudo revelou a presença de um cromossomo 1 em anel em LU, que levou a perda da região 1p36.1. Nesta região, está mapeado outro gene do grupo de alta mobilidade, o *HMG2/HMG17*. Polito et al. (1999) utilizando FISH (hibridação fluorescente *in situ*) demonstram que LU com esta deleção não apresentaram sequências do gene HMG17, sugerindo que este gene não desempenha um importante papel para o desenvolvimento dos Leiomiomas Uterinos.

Há relatos de LU que apresentam deleção de quase todo o braço curto do cromossomo 1 [$del(1)(p11p36)$], juntamente com perdas do cromossomo 19 e ou 22. Porém, estes tumores possuem características histológicas não-usuais, como hiperplasticidade e atipia nuclear, como já relatado acima (Christacos et al., 2006). Outra anormalidade envolvendo o cromossomo 1 é a $t(1;2)(p36;p24)$. Esta translocação recíproca envolve os genes *AJAP1* e *NPHP4* localizados no cromossomo 1 e os genes *ITSN2* e *NCOA1* no cromossomo 2, que desempenham importantes papéis em tumores humanos (van Rijk et al., 2009).

Vários rearranjos do cromossomo 3, incluindo inserções, deleções e translocações com o cromossomo 7 foram relatadas em Leiomiomas Uterinos. Já, as deleções do braço longo do cromossomo 13 podem ocorrer sozinhas ou juntamente com outros rearranjos (Meloni et al., 1991). Aparentemente, muitas destas alterações relatadas em baixa frequência são eventos secundários em Leiomiomas Uterinos.

1.2.1 Hibridação Genômica Comparativa (CGH)

Na década de 90, foi descrito um dos métodos mais promissores para o estudo dos cromossomos humanos, a Hibridação Genômica Comparativa (CGH). Esta técnica descrita por Kallioniemi et al. (1992) permite a detecção da razão entre ganhos e perdas de sequências de DNA no genoma teste em relação a um genoma de referência normal. A CGH é um método de triagem com base nos mesmos princípios da FISH, para produzir um mapa linear de

sequências de DNA com uma função de localização da banda cromossômica no genoma. A capacidade da CGH em detectar ganhos e perdas em regiões pequenas de DNA depende do grau de condensação das metáfases e da frequência do número de cópias presentes na amostra tumoral. Mesmo assim, esta técnica revela vantagens sobre a citogenética convencional, pois possui uma maior resolução e permite identificar rearranjos não-balanceados com maior precisão (Bayani e Squire, 2005).

Para nosso conhecimento, existem quatro relatos em literatura (incluindo um estudo do nosso grupo) que utilizaram a técnica da CGH para investigar ganhos e perdas genômicos em Leiomiomas Uterinos. Packenham et al. (1997) analisaram 14 casos de LU e dois casos apresentaram alterações genéticas incluindo ganhos nos cromossomos 14q11.1-q21 e 14q24-qter e 19p12-pter e 19q12-13.1 e perdas dos cromossomos 1 e 4. Posteriormente, Rikala et al. (1998) relataram perfis genômicos normais em cinco amostras de LU. Levy et al. (2000) em 12 amostras de LU relataram ganhos em 9q43 e no cromossomo 19 e perdas em 7q e 12q. Nosso grupo encontrou uma proporção significativa de tumores com alterações genômicas (82%) e uma frequência significativa de perdas cromossômicas não relatadas anteriormente incluindo as regiões 3q26.29, 7p22.15, 11p15, 11q23.25 e 15q25.26 (Canevari e Rogatto, 2007). As perdas genômicas nestes *loci* foram investigadas usando análises de LOH e marcadores polimórficos confirmando o seu envolvimento em Leiomiomas Uterinos (Canevari et al., 2005b).

1.2.2 Hibridação Genômica Comparativa baseada em Microarrays (CGH array)

Inicialmente, os microarrays foram desenvolvidos para avaliar a expressão gênica global utilizando populações de RNA ou cDNA. Este método foi adaptado para permitir a análise de alterações no número de cópias de DNA. Uma vantagem da CGH array sobre o cDNA array é a maior uniformidade na hibridação e subsequente fidelidade do sinal, pois os alvos de DNA têm uma maior complexidade, contendo seqüências intrônicas e outras não

transcritas (Beheshti et al., 2002). A técnica de CGH array é extremamente útil para identificar marcadores cromossômicos e permite definir a localização precisa das bandas envolvidas entre as aberrações estruturais, independentemente da combinação com análises cariotípicas (Singh et al., 2011). O princípio metodológico é semelhante ao da CGH, entretanto, supera suas limitações de resolução pelo uso de clones imobilizados em uma lâmina de vidro em posições bem definidas, ao invés de metáfases normais. Alvos de DNA para microarray podem ser derivados de clones genômicos incluindo YACs (0,2 a 2 Mb em tamanho), BACs (300 Kb), P1s (~70-100 Kb), PACs (~130-150 Kb), cosmídeos (~30-45 Kb), assim como oligonucleotídeos sendo, portanto, de várias ordens de magnitude. Atualmente, as plataformas de oligonucleotídeos têm sido amplamente utilizadas na avaliação do número de cópias de DNA, com resoluções que permitem a cobertura quase completa do genoma humano (Przybytkowski et al., 2011). Esta diminuição no tamanho do alvo aumenta o poder de detecção de alterações genômicas em comparação a CGH cromossômica. Quantidades equimolares de DNA teste e referência, marcados diferencialmente, são hibridados em lâminas contendo fragmentos genômicos clonados e fixados em posições definidas formando um *chip*. Diferenças na razão da intensidade dos sinais podem ser interpretadas como diferenças no número de cópias entre os genomas teste e referência. Devido a sua maior resolução, a CGH array permite a identificação de genes que atuam nas diferentes etapas da carcinogênese revelando seqüências que podem estar envolvidas na gênese ou progressão tumoral. Entretanto, esta técnica não é capaz de identificar mosaicismos em baixa frequência e rearranjos equilibrados, assim como na CGH cromossômica (para revisão Rogatto, 2000).

Para o nosso conhecimento, existem cinco relatos em literatura utilizando a metodologia de CGH array (Quadro 1) e dois estudos utilizando SNP array para avaliar o perfil genômico dos Leiomiomas Uterinos.

Quadro 1 – Relação dos relatos em literatura que utilizaram análises de CGH array em Leiomiomas Uterinos, as diferentes plataformas utilizadas e os principais resultados.

Referência	Amostras	Plataforma	Principais Resultados	Validação
Cho et al (2005)	4LU + 7LMS	BACs, 1440 sequências	LU: normal LMS: ganhos 7q36.3, 7q33-q35, 12q13-12q15 e 12q23.3 perdas 1p21.1, 2p22.2, 6p11.2, 9p21.1, 9p21.3, 9p22.1, 14q32.33 e 14q32.33qter	FISH
Vanharanta et al. (2005)	10LU com LOH em 7q	BAC , 3452 sequências	perdas 7q22.3-q31.1	Microarray expressão qRT-PCR
Vanharanta et al. (2007)	12LU	7q <i>Tiling-Path MicroCGH array</i>	perdas 7q22.3-31.1, 7q34	LOH
Hodge et al. (2009)	6LU com del FISH 7q	<i>Agilent Human 244K CGH Microarrays</i>	perdas 7q22.1-q31.1	Microarray expressão IPA
Zavadil et al. (2010)	6LU	<i>Agilent Human 4x44K CGH Microarrays</i>	perdas 1p36.33-p36.23, 1p36.33-p34.3, 3p11.2-p11.1, 3p12.3-p11.1, 3q13.31-q21.2, 3q26.1-q29, 3q22.2-q27.2, 6p12.3-p11.1, 6p25.3-p22.2, 6q13-q24.3, 13q12.12-q33.2, 22q11.1-q13.33.	RT-PCR modelo <i>in vivo</i>

Cho et al (2005) avaliaram quatro casos de LU e sete casos de LMS. Os autores utilizaram uma plataforma de BACs contendo 1440 seqüências genômicas humanas (Macrogen, Korea, <http://www.macrogen.co.kr>) com 2,08 Mb de distância entre os BACs. Os autores não identificaram alterações em LU, contudo descreveram várias alterações nas amostras de LMS como ganhos em 7q36.3, 7q33-q35, 12q13-12q15 e 12q23.3 e perdas homozigotas em 1p21.1, 2p22.2, 6p11.2, 9p21.1, 9p21.3, 9p22.1, 14q32.33 e 14q32.33qter. A provável explicação para a ausência de alterações nos LU poderia ser o pequeno número de seqüências presentes na plataforma e/ou a grande distância entre elas.

Vanharanta et al. (2005) avaliaram 10 amostras de LU que apresentaram previamente LOH em 7q, para confirmar as deleções identificadas nessa região. Os autores utilizaram uma plataforma de BAC contendo 3452 seqüências com 1Mb de espaçamento (Douglas et al., 2004). Os autores definiram a região mínima comum em 7q22.3-q31.1 envolvida nestas

amostras. Em adição, os autores realizaram microarrays de expressão em amostras de LU que apresentavam a deleção 7q22.3-q31.1 com amostras que não apresentaram nenhuma deleção em 7q, contudo, não identificaram expressão gênica diferencial entre estes dois grupos de amostras.

Vanharanta et al. (2007) avaliaram 12 amostras de LU de mulheres de diferentes etnias utilizando a plataforma *Tiling-Path MicroCGH array* para realizar um *screening* das alterações em 7q. Os autores identificaram uma baixa frequência de deleções em 7q22.3-31.1 (6% dos casos) e uma nova deleção em 7q34. Os genes *SPRK2*, *PUS7*, *RINTI*, *SYPL*, *PBEFI*, *PIK3CG*, *PRKAR2B*, *HBPI*, *COG5*, *GPR22* e *DLD* envolvidos nestas regiões de perda foram avaliados por análise de LOH. Dentre os 11 genes, oito tiveram a região de perda confirmada. Segundo os autores, podem existir diferentes populações com suscetibilidade para o desenvolvimento dos LU na via de alterações em 7q.

Hodge et al. (2009) também focaram seus estudos em amostras de LU com deleções em 7q. Os autores utilizaram a técnica de CGH array (*Agilent Human 244K CGH Microarrays*) em seis casos de LU com deleções em 7q previamente identificadas por citogenética clássica ou pela FISH. A região mínima comumente deletada foi a 7q22.1-q31.1 e com o envolvimento de ~9.5Mb. Os autores também realizaram experimentos de expressão gênica global (*Affymetrix*). Após análises *in silico*, foi confirmada a diminuição de expressão de alguns genes mapeados em 7q22.1-q31.1. Entre eles, o gene inibidor de proliferação *HPB1* e o gene *RINT1*, um supressor de tumor envolvido na manutenção da integridade celular e mitose. Análises de interações gênicas utilizando o IPA (*Ingenuity Pathway Analysis*) revelou que os genes centrais participam dos processos de ubiquitinação de proteínas (*PSMA2*, *PSMB3*, *PSMC2*, *PSMC3*, *UBE2I* e *UBE2J1*). Os autores concluíram que alterações nos processos de ubiquitinação protéica podem ser primordiais no desenvolvimento dos LU com deleções em 7q.

Zavadil et al. (2010) realizaram CGH array (*Agilent Human 4x44K CGH Microarrays*) em seis amostras de LU com o objetivo de avaliar se os miRNAs com expressão alterada,

identificados previamente, estavam mapeados em regiões de instabilidade genômica. Os miRNAs são pequenos RNAs não-codificadores (20-24nt) que estão envolvidos na regulação pós-transcricional da expressão gênica em organismos multicelulares, afetando a estabilidade da tradução de mRNAs, incluindo mRNA de genes envolvidos em câncer (Huang et al., 2010). Em LU, existem vários estudos que reportam o perfil de expressão de miRNAs (para revisão Pan et al., 2009).

Os autores identificaram exclusivamente perdas genômicas em 1p36.33-p36.23, 1p36.33-p34.3, 3p11.2-p11.1, 3p12.3-p11.1, 3q13.31-q21.2, 3q26.1-q29, 3q22.2-q27.2, 6p12.3-p11.1, 6p25.3-p22.2, 6q13-q24.3, 13q12.12-q33.2, 22q11.1-q13.33. Os genes regulados pelos miRNAs mapeados em 1p36.33-p34.3, 6p12.3-p11.1 e 13q12.12-q33.2 foram validos por RT-PCR e modelos *in vivo*.

Em adição aos estudos de CGH array, Bowden et al. (2009) avaliaram 16 amostras de LU utilizando a metodologia de SNP array que apresenta maior resolução para a identificação de alterações genômicas em relação a CGH array. Os autores identificaram deleções em 1p, 1q42.13, 2q37.3, 3q25.1-q26.1, 7q21.12, 7q21.2-q31.31, 7q21.3-q31.1, 7q32.3, 11q13.1, 12q13.13, 14q23.3-qter, 15q11.2-q23, 17p-q21.31, 20q11.21, 22q12.2-q12.3, monossomia do 13 e ganhos em 13q12.13. Os genes *RHOU* (1q42.13) *MAP3K11* (11q13.1) e *WASF3* (13q12.13) foram validados por qPCR indicando que estas alterações podem contribuir para a patogênese dos Leiomiomas Uterinos.

Recentemente, Meadows et al. (2011) avaliaram 37 amostras de LU utilizando SNP arrays (*Affymetrix 100K*) para avaliar a frequência de amplificações e LOH nestas amostras. Os autores identificaram LOH em apenas 4/37 amostras que foi estatisticamente associado com tumores maiores. Duas amostras apresentaram LOH em 7q e uma amostra apresentou uma nova região de LOH em 16p13.11. Baixos níveis de amplificações foram identificadas em todos os cromossomos de uma amostra e amplificação em 3p26.3 foi detectada em 4/37 amostras. Os autores concluíram que LU possuem baixa frequência de LOH e amplificações,

evento que podem não estar diretamente associados com características clínicas e histológicas nestes tumores.

A análise de variações no número de cópias do DNA pode permitir a identificação de alterações genômicas não detectadas pelos métodos inicialmente utilizados (análise cromossômica), os quais consideraram que um número significativo de casos era cromossomicamente normal (60%) e revelar alterações funcionais que podem ser úteis para o delineamento de estratégias terapêuticas nestes tumores.

1.3 Perfil de Expressão Gênica Global

Os perfis de expressão gênica em larga escala têm sido obtidos para auxiliar no conhecimento da biologia tumoral, refinar a estratificação da doença e melhorar a precisão do diagnóstico e prognóstico (Perou et al., 2000; Alizadeh et al., 2000 e 2001; Sotiriou et al., 2003; Sorlie et al., 2003). A obtenção das “assinaturas moleculares” dos tumores tem auxiliado na conduta clínica de pacientes, principalmente aquelas afetadas por tumores de mama (Eifel et al., 2001; Gevaert et al., 2006). A análise dos dados de perfis de expressão resulta em listas de genes que são diferencialmente expressos entre um tecido normal e tumoral ou que se correlacionam com fenótipos tumorais, como por exemplo, tumor primário *versus* metastático (para revisão Goodison et al., 2010). Diferentemente do princípio da CGH ou CGH array, no qual o resultado é obtido a partir da razão de sinal entre as amostras teste e referência, nos experimentos de expressão gênica global de duas cores, a referência pode representar tanto uma amostra cujo perfil de expressão em relação à amostra teste é de interesse biológico ou simplesmente uma amostra de referência, usada em todos os experimentos a fim de controlar a variabilidade experimental (para revisão Simon, 2008).

A metodologia de microarrays de expressão tem sido aplicada em LU e permitido a identificação de genes diferencialmente expressos nos tumores em relação ao miométrio normal adjacente da mesma paciente, assim como, para obter o perfil de expressão dos LU

associados à resposta a drogas e ou tratamentos e também permitir a comparação do perfil molecular destes tumores com outras doenças. Para o nosso conhecimento, até o momento já foram descritos 27 estudos de expressão gênica em LU (Anexo 1).

Como já mencionado, os estrógenos e a progesterona são considerados os principais fatores associados ao desenvolvimento dos LU, porém, a maioria dos estudos de microarrays de expressão não tem demonstrado um perfil de expressão gênica diferencial de genes envolvidos na ação destes hormônios. Na verdade, estes estudos têm revelado uma gama de genes envolvidos na formação da matriz extracelular, colágenos, proteoglicanos e elastina (Leppert et al., 2006). O primeiro estudo avaliando a expressão gênica diferencial dos LU em relação aos miométrios adjacentes foi realizado por Tsibris et al. (2002). Os autores selecionaram cinco amostras de LU na fase proliferativa e quatro amostras na fase secretora do ciclo menstrual. Entre os genes com expressão aumentada, foram encontrados o *IGF2* (desenvolvimento e crescimento celular), *FRIZZLED* (sinalização WNT), *CRABP2* (sinalização do ácido retinóico) e *TGFB3* (diferenciação celular) e entre aqueles com expressão diminuída, *HNF1A* (fator de transcrição), *TPSAB1* (regulação da estabilidade de mRNA), *DPT* (matriz extracelular), *THBS1* (adesão celular) e *c-FOS* (fator de transcrição). As funções exercidas por estes genes representam as vias metabólicas que geralmente encontram-se alteradas nos LU, incluindo os genes associados ao remodelamento da matriz extracelular.

Os LU apresentam uma característica fibróide particular, devido à deposição contínua de colágenos. Stewart e Nowak (1998) propuseram uma hipótese para a patogênese dos LU, no qual a formação do tumor seria causada por uma resposta a um dano tecidual, similar ao que acontece na formação dos quelóides (cicatrizes). Esta lesão seria causada pelo aumento de substâncias vasoconstritoras (como prostaglandinas e vasopressinas) no miométrio durante o período menstrual, e como reação a esta lesão, haveria um aumento da produção de fator básico de crescimento de fibroblastos (bFGF), com conseqüente acúmulo de matriz extracelular (Flake et al., 2003). O TGF-beta tem sido considerado o maior regulador autócrino/parácrino da fibrose nos LU, pois além de regular sua própria expressão, regula

também os genes codificadores dos Smads, c-Fos, c-Jun, fibronectina, colágeno e PAI1 (Chegini et al., 1999 e 2002; Ding et al., 2004).

Catherino et al. (2004b) avaliaram o perfil de expressão gênica de 23 LU e MM e identificaram genes diferencialmente expressos codificadores de proteínas de matriz extracelular. Entre os genes descritos, estudos de validação (qRT-PCR) revelaram expressão diminuída de *DPT* (dermatopontina). A dermatopontina tem associação direta com a formação quelóide e os autores demonstraram a expressão reduzida desta proteína em ambos tecidos. Em adição, análises estruturais revelaram semelhanças na orientação das fibras de colágeno em LU e quelóides. Além disso, expressão diminuída de *DPT* foi estaticamente associada com aumento de *TGFB3*, um dos genes codificadores de TGF-beta, o principal fator associado com a via fibróide. Os autores concluíram que expressão diminuída de *DPT* e aumento de *TGFB3* indicam uma semalhança molecular entre LU e quelóides.

Luo et al. (2005b) investigaram a expressão gênica diferencial entre células de LU e do miométrio sob a ação do TGF-beta e identificaram genes que atuam como reguladores da transcrição, do metabolismo celular, do ciclo celular, da transdução de sinal e moléculas da matriz extracelular.

Baseados nos achados anteriores, Leppert et al. (2006) sugeriram uma “nova hipótese sobre a origem dos Leiomiomas Uterinos”. Segundo os autores, as células normais do miométrio podem se transformar em células leiomiomatosas em resposta a sinalização desordenada dos sinais da matriz extracelular. Entre estes componentes, o TGF-beta seria o principal fator, que leva a inflamação e proliferação destas células, mas que ao mesmo tempo, contribui para a deposição de colágeno e remodelamento contínuo da matriz extracelular.

Posteriormente, Luo et al. (2007) confirmaram esta hipótese. Os autores identificaram uma incrível semelhança molecular entre os LU e quelóides, demonstrando inclusive, a presença dos mesmos genes codificadores dos colágenos I e III entre as duas entidades. Estes dados revelam o importante papel do TGF-beta na formação fibróide dos Leiomiomas Uterinos.

Além de sua natureza fibróide, os LU apresentam características proliferativas. Swartz et al. (2005) avaliaram a expressão gênica diferencial entre células de LU e de miométrio normal, tratadas e não tratadas com 17 beta-estradiol. Entre os genes diferencialmente expressos, os genes *IGF-I* e *A-MYB* apresentaram expressão aumentada nos Leiomiomas Uterinos. Estes genes são conhecidos por desempenharem um papel mitogênico em resposta ao estrógeno e poderiam estar associados a proliferação das células leiomiomatosas (Yu e Rohan, 2000). O estudo mais recente (Jiang et al., 2010) teve como objetivo testar a hipótese sobre o crescimento celular anormal dos LU associado com alterações da atividade das proteínas tirosina quinases. Jiang et al. (2010) selecionaram 42 LU e MM e avaliaram o perfil de expressão gênica diferencial, utilizando uma plataforma customizada contendo sequências de 244 genes (incluindo 90 codificadores de tirosina quinases e 103 que codificam os seus ligantes). Os autores confirmaram a hipótese no qual as alterações nas atividades de proteínas tirosina quinases, provavelmente causadas pela alteração na expressão de seus ligantes, podem contribuir para o crescimento e/ou desenvolvimento dos Leiomiomas Uterinos.

Estudos de microarrays de expressão também têm sido úteis para avaliar o perfil molecular dos LU em resposta a tratamentos. No Brasil, um estudo avaliou o perfil de expressão gênica de LU tratados e não tratados com goserelina e será descrito a seguir. Na comparação entre o grupo tratado em relação ao não tratado foram identificados principalmente os genes *CYR61*, *SLUF1*, *HMGN1* e *EGR1* com expressão aumentada e o gene *WIF1* com expressão diminuída. Baseado nestes achados foi possível estabelecer quais pacientes respondem ou não ao tratamento com esta droga (Borsari et al, 2010).

Outro estudo para avaliar a resposta ao tratamento foi realizado entre LU de ratos e humanos (Crabtree et al., 2009). Este estudo revelou a sobreposição significativa de genes desregulados da via mTOR entre os tumores de ratos e humanos. Ratos produzidos sem o gene *Tsc2* desenvolveram LU por meio da regulação aumentada da via mTOR. Os ratos foram tratados com uma droga inibidora desta via (WAY-129327) e após quatro meses de

tratamento os tumores diminuíram significativamente de tamanho. Este modelo permitiu desvendar parte da fisiopatologia dos LU e identificar via mTOR como um novo alvo terapêutico nestes tumores.

De uma forma geral, os microarrays de expressão têm contribuído para o conhecimento da etiologia dos LU e revelado inúmeros candidatos a alvos terapêuticos nestes tumores. Foram descritos genes codificadores de proteínas de matriz extracelular, envolvidos nas vias de transdução de sinal e que atuam como fatores de crescimento e de transcrição. Em adição estes estudos permitiram a avaliação da assinatura molecular dos tumores em resposta a drogas. Entretanto, para o nosso conhecimento, nenhum destes estudos teve como objetivo associar os perfis de expressão gênica com dados genômicos numa forma de identificar alterações genômicas funcionais em LU. Atualmente, a metodologia de microarrays de expressão tem sido intensamente explorada em associação com estudos de variações no número de cópias de DNA para a identificação de CNVs que possam levar a efeitos fenotípicos.

1.4 Integração de Dados Genômicos e Transcriptômicos

Quando associados com estudos matemáticos e modelos funcionais, a abordagem de análise integrativa pode revelar genes candidatos associados a importantes vias metabólicas, extremamente úteis para o desenvolvimento de terapias alvo em doenças genéticas. Como já mencionado, as CNVs são eventos fundamentais para a carcinogênese humana e os estudos de CGH array têm sido particularmente úteis para identificar os genes candidatos em regiões amplificadas ou perdidas no genoma tumoral. Considerando que geralmente muitos genes estão localizados em uma região de ganhos ou perdas, os resultados de expressão permitem a classificação dos genes com base em quanto eles correspondem a um determinado fenótipo (Akavia et al., 2010).

Os resultados são puramente baseados em análises de bioinformática, portanto, os dados são obtidos de acordo com a aplicação de algoritmos empregados no estudo. Por ser uma estratégia relativamente nova, não existe um protocolo estabelecendo o algoritmo *gold standard* para a aplicação em análises integrada de dados genômicos e transcriptômicos. Na verdade, existem alguns métodos estatísticos mais robustos empregados na maioria dos estudos, como por exemplo, testes de correlação de *Pearson* e *Spearman* (Meza-Zepeda et al., 2006; Vincent-Salomon et al., 2008). Algoritmos que utilizam outros princípios matemáticos também têm sido desenvolvidos, como o *CONEXIC* (baseado em modelos Bayesianos) (Akavia et al., 2010) e o *GISTIC* (baseado no *Q value*) (Green et al., 2011). Na tentativa de facilitar o uso destes métodos estatísticos, vários programas têm sido desenvolvidos com o objetivo de permitir a realização de análises integradas, como o *CHESS*, (Lee e Kim, 2009) *ICPS* (Zhang et al., 2010) e *CNAmet* (Louhimo e Hautaniemi, 2011).

Em uma busca realizada no banco de dados públicos *PubMed* usando os termos *integrative genomic and transcriptome* o resultado gerou mais de 400 estudos (www.pubmed.com-acesso em 29/03/2011). Contudo, para o nosso conhecimento, ainda não existem relatos utilizando esta abordagem em Leiomiomas Uterinos. O primeiro estudo que associou “indiretamente” (ou seja, sem a aplicação de algoritmos) CNVs e níveis de expressão gênica em LU foi realizado em 2004. Neste estudo, entre sete amostras de LU avaliadas por citogenética clássica, cinco apresentaram cariótipo normal, uma apresentava a $\text{del}(7)(q22q32)$ e a outra a $\text{t}(12;14)(q15;q24)$. A amostra com a deleção em 7q não apresentou nenhum gene com alteração de expressão mapeado na região, já a amostra com $\text{t}(12;14)$ apresentou expressão aumentada do gene *HMGA2* (*fold-change* 55 vezes mais alto do que as amostras com cariótipos normais) (Quade et al., 2004).

Os outros dois estudos de análise integrada indireta basearam-se em amostras de LU com deleções em 7q evidenciadas por CGH array e FISH e que foram avaliadas por expressão gênica global (já apresentados acima). Independentemente, os estudos não revelaram fortes genes candidatos com expressão alterada mapeados nesta região (Vanharanta et al., 2005;

Hodge et al., 2009). Estes estudos focaram seus esforços para identificar um gene supressor tumoral mapeado em 7q

Estes dados reforçam a necessidade de novos estudos genômicos e transcriptômicos para a identificação de CNVs que possam levar a expressão alterada de genes críticos associados com Leiomiomas Uterinos.

2. Objetivos

- ✓ Caracterizar as alterações no número de cópias de DNA em amostras de Leiomiomas Uterinos múltiplos e únicos, utilizando a metodologia de CGH array;
- ✓ Obter o perfil de expressão gênica global entre os Leiomiomas Uterinos e miométrio adjacente de um subgrupo de pacientes nas fases proliferativa e secretora do ciclo menstrual;
- ✓ Integrar os dados genômicos e transcriptômicos por análises de bioinformática para descrição de genes moduladores associados com o fenótipo tumoral;
- ✓ Aplicar análises funcionais *in silico* para identificar redes e vias canônicas e caracterizar alvos moleculares associados a etiologia dos Leiomiomas Uterinos.

3. Material e Métodos

3.1 Casuística

Oitenta amostras de tecido a fresco foram obtidas de 56 pacientes com Leiomiomas Uterinos. O miométrio normal adjacente também foi coletado (n=56). As amostras foram transferidas imediatamente para nitrogênio líquido e estocadas a -80°C. Fragmentos destes tumores também foram avaliados pelo patologista responsável do Departamento de Patologia da Faculdade de Medicina de Botucatu (FMB-UNESP). Estas amostras foram obtidas de pacientes submetidas à histerectomia no Departamento de Ginecologia e Obstetrícia (FMB-UNESP) entre Outubro de 1995 e Fevereiro de 2004. As amostras fazem parte do Banco de Amostras do NeoGene Laboratório, Departamento de Urologia (FMB-UNESP), sob a responsabilidade da Dra Silvia Regina Rogatto. Este estudo foi aprovado pelo Comitê de Ética e Pesquisa da FMB-UNESP (146/2007-CEP). Todas as pacientes assinaram “Termo de Consentimento Livre e Esclarecido” antes da coleta das amostras, os quais se encontram sob a responsabilidade da Dra Anaglória Pontes, Departamento de Ginecologia e Obstetrícia (FMB-UNESP).

3.1.1 Características clínicas e histopatológicas

Treze pacientes apresentaram tumores únicos (n=13) e 43 tinham tumores múltiplos. Entre as pacientes com tumores múltiplos, 26 tiveram apenas uma amostra avaliada (n=26), dez tiveram duas amostras avaliadas (n=20) e sete tiveram três amostras avaliadas (n=21).

Todas as pacientes eram pré-menopáusicas e não fizeram uso de terapia hormonal por pelo menos três meses antes da cirurgia. Tecidos conectivos e necróticos foram removidos das amostras. Em 2008, os prontuários e achados histopatológicos foram

revisados para a atualização dos dados clínicos como, idade, idade à menarca, idade à primeira gravidez, índice de massa corpórea (IMC), fase do ciclo, número de tumores, raça, gestações, história familiar de LU e ou outros tumores e tipo histológico tumoral (Anexo 2). Todas as amostras foram classificadas como LU típicos. Para a determinação da fase do ciclo, foi avaliado o *status* endometrial de cada paciente, utilizando o critério de Noyes et al. (1950). Os valores de referência para o IMC (índice de massa corpórea) foram ≤ 18 abaixo do peso, 18,5 a 24,9 peso normal, 25 a 29,9 sobrepeso e ≥ 30 obesidade (WHO, 1997).

3.2 Extração de DNA

Amostras de DNA genômico foram obtidas das 80 amostras. Para a extração, foi utilizado o método padrão fenol/clorofórmio/álcool isoamílico. Estas amostras foram tratadas com RNase 20mg/mL (Sigma-Aldrich, St. Louis, MO) para digestão de moléculas de RNA remanescentes do processo de extração. Após a extração do DNA, as amostras foram ressuspensas em até 50 μ L de água ultra-pura estéril e armazenadas a -20°C . Em seguida, 100ng/ μ L de cada amostra foram aplicados em gel de agarose 0,8% para avaliação da integridade do DNA genômico. As amostras também foram quantificadas em espectrofotômetro (NanoDrop ND-1000 Spectrophotometer v.3.0.1, Labtrade) para avaliação de pureza e qualidade da molécula.

3.3 Extração de RNA

A extração do RNA foi realizada em um subgrupo de tumores de 34 pacientes, totalizando 51 amostras. O critério de seleção destas amostras foi baseado na informação da fase do ciclo menstrual no momento da cirurgia, onde 13 pacientes estavam na fase proliferativa e 21 encontravam-se na fase secretora.

Também foram extraídas as amostras de RNA do miométrio adjacente desses casos (n=34). O protocolo estabelecido para extração de RNA foi o *RNeasy Mini Kit* (Qiagen). A qualidade e a pureza do RNA foram avaliadas utilizando-se espectrofotômetro (NanoDrop ND-1000 Spectrophotometer v.3.0.1, Labtrade) e o *Bioanalyzer* (RNA 6000 NanoLabChip kit 2100 Agilent Technologies), respectivamente.

3.4 CGH array – Marcação, hibridação e análise

Foram utilizadas as lâminas *Agilent Human 4x44K CGH Microarrays* (Agilent Technologies). Esta é uma plataforma de oligonucleotídeos de 60pb com espaçamento médio entre as sondas de aproximadamente 43Kb. A resolução desta plataforma é de aproximadamente 76Kb, que é o resultado da razão entre o número de pares de base do genoma humano (~ 3.260Mb) sobre o espaçamento das sondas na plataforma (43Kb). Possuem aproximadamente 43.000 sondas que mapeiam genes bem caracterizados, particularmente aqueles envolvidos em câncer, incluindo sequências codificadoras e não-codificadoras (<http://www.chem.agilent.com>). DNA Genômico Universal (Promega) foi utilizado como amostra referência, que é constituído por amostras genômicas provenientes de vários doadores.

Amostras tumorais com moléculas de DNA de alto peso molecular foram submetidas à marcação pelo *Kit Bioprime CGH Labeling Module* (Invitrogen), de acordo com especificações do fabricante. Para a realização dos experimentos foram utilizados 500ng de DNA os quais foram submetidos a uma etapa de digestão enzimática randômica em reação contendo água ultra-pura estéril; 2,6µL de tampão 10x; 0,2µL de BSA; 0,5µL de cada enzima (*AluI* e *RsaI*) totalizando 26µL. As reações foram incubadas a 37°C por 2 horas, 65°C por 20min e mantidas em gelo durante o procedimento. Foram adicionados na reação 20µL de *Random Primer* (Invitrogen), e então, as amostras foram desnaturadas a 95°C por 5min e transferidas para o gelo. As amostras de referência e teste foram

marcadas com 1,5µL de Cy3 e Cy5 dCTPs, respectivamente acrescidas de 5µL do dNTP mix e 1µL da enzima *Exo-Klenow*. A reação de marcação foi realizada a 37°C por 2 horas, 65°C por 10min e mantidas em gelo durante a reação. As amostras foram purificadas em minicolunas de celulose (Microcon YM-30, Millipore) e ressuspendidas em 21µL de TE (tampão Tris-EDTA) 1x pH=8.0 (Prodinol). A eficiência de marcação das amostras foi obtida baseando-se na razão entre pmol/µL do corante sobre concentração em µg/µL. As amostras referência e teste com eficiências semelhantes foram combinadas. Adicionou-se a cada reação *Human Cot DNA*, *10X Blocking Agent* e *2X Hi-RPM Hybridization Solution*. A reação foi desnaturada a 95°C por 3min e incubada a 37°C por 30min. As amostras foram aplicadas nas lâminas e permaneceram a 65°C a 20rpm por 24hs. Após este período, as imagens de hibridação de cada lâmina foram capturadas usando o *Scanner Agilent G2565CA*. Os dados foram extraídos pelo programa *Feature Extraction v.10.1.1.1* (Agilent Technologies).

A análise dos dados foi realizada utilizando o programa *Nexus v.5.0* (Biodiscovery). Apenas os casos que apresentaram valores de qualidade de hibridação $\leq 0,3$ foram selecionados para análise. Todas as amostras preencheram este critério. As amostras foram hibridadas utilizando DNA de referência masculino pois essa estratégia, é um controle de hibridação a cada experimento, devido à competição não balanceada na hibridação das sequências dos cromossomos sexuais. Portanto, alterações envolvendo os cromossomos X e Y foram excluídas das análises.

O algoritmo *Rank Segmentation* foi utilizado para segmentar cada intervalo cromossômico e estimar o número de cópias. Este método é uma variante do algoritmo CBS, que utiliza um modelo paramétrico, em vez de testes de permutação, para determinar a significância dos agrupamentos de segmentação (Olshen et al., 2004). Os parâmetros de análise foram: limiar de significância de $1,00E^{-4}$, espaçamento máximo de 1Mb entre sondas adjacentes e a presença de três sondas consecutivas alteradas para determinação de um segmento de DNA alterado. Os limiares para a classificação de ganhos ou perdas

foram 0,2=ganhos; 0,6=alto nível de ganhos; -0,2=perdas; -1,0=perdas homozigotas, como recomendado pelo fornecedor (*Nexus version 5.0*, Biond Discovery).

Para verificar se as amostras se agrupavam de acordo com a presença de alterações genômicas específicas, foi aplicado o teste *Complete Linkage Hierarchical*. As comparações entre os diferentes grupos com alterações genômicas e dados clínicos foram realizadas pelo Teste Exato de Fisher com $P \leq 0.05$. A descrição de CNVs significativas (picos significativos) baseou-se na presença de três critérios: 1) alterações presentes em mais de 10% dos casos; 2) regiões com $P \leq 0.05$ e 3) exclusão de CNVs presentes em >5% da população controle. Esta população controle foi composta por um grupo de 82 mulheres brasileiras com a ausência de tumores e doenças genéticas hereditárias (as quais fazem parte de outro estudo da equipe). A exclusão de CNVs presentes >5% dessa população foi baseada nos critérios de McCarroll et al. (2008). Segundo esses autores, 80% das diferenças no número de cópias entre pares de indivíduos ocorrem devido a CNVs comuns com uma frequência alélica >5%. Os genes mapeados nas regiões significativas foram submetidos a análise funcional.

3.5 Hibridação *in situ* fluorescente (FISH)

Em um estudo prévio do grupo, 27 amostras de LU foram investigadas por FISH para caracterização de três regiões genômicas envolvidas em deleções no cromossomo 7, incluindo a del(7)(q22.3-q31.1) (Perez, 2004). Vinte destas amostras também foram também avaliadas por CGH array. As lâminas contendo preparações cromossômicas das amostras tumorais e cinco lâminas controle obtidas a partir de linfócitos normais estocadas a -20°C foram utilizadas na investigação por FISH. Neste estudo, a sonda teste loco-específica RP11-46J20 (localizada em 7q22.3-q31.1) foi clonada em vetores do tipo BAC (*Bacterial Artificial Chromosome*). O mapeamento físico da sonda teste mostrou o envolvimento de quatro genes: *COG5* (*component of oligomeric golgi complex 5*), *DUS4L*

(*dihydrouridine synthase 4-like*), *BCAP29* (*B cell receptor-associated protein 29*) e *SLC26A4* (*solute carrier family 26, member 4*). Procedimentos de hibridação em duas cores (*dual color*) com a sonda controle (α satélite p7t1 α 7) e a sonda teste foram realizados para cada caso. Uma média de 100 núcleos interfásicos intactos e não sobrepostos, contracolorados com 4,6-diamino-2-phenylindole (DAPI), foram analisados para determinar o número de sinais da sonda e a frequência de perdas envolvendo a região 7q22.3-q31.1. As alterações no número de cópias detectadas foram consideradas significativas quando o valor observado para cada caso era superior a média mais dois desvios-padrão das amostras do tecido normal usado como controle, calculados para cada sonda. Deleções hemizigotas foram consideradas significativas quando detectadas em mais de 30% das células.

3.6 Microarrays de Expressão - Marcação, hibridação e análise

O perfil de expressão gênica em larga escala foi obtido em 51 amostras de LU utilizando-se a plataforma de microarray *Whole Human Genome 4x44K* (Agilent Technologies), composta por 43.376 sequências biológicas e 32 sondas *Spike-in* (controles positivos) (<http://www.genomics.agilent.com>). Os procedimentos de marcação das amostras, hibridação e detecção seguiram o protocolo *Two-color microarray-based gene expression analysis* (Agilent Technologies). As amostras de LU foram hibridadas contra RNA referência universal composto por 15 linhagens celulares (Gomes et al., 2005; Boccardo, et al. 2010), assim como, as amostras de miométrios normais adjacentes segundo a fase do ciclo (*pool* de amostras). Para o protocolo de marcação com “duas cores,” foi utilizado o kit o *Quick Amp Labeling* (Agilent Technologies). A reação é iniciada com uma pré-síntese de cDNA, no qual foram utilizados 800ng de RNA (volume máximo de 8,3 μ L), 2 μ L da diluição do *Spike A* para Cy3 (teste) ou *Spike B* para Cy5 (referência) e 1,2 μ L de T7 *Promoter Primer*. Essa reação foi incubada a 65°C por 10min e colocada no gelo por 5min. A síntese de cDNA consistiu na adição de 4 μ L de *5X First Strand Buffer*, 2 μ L de DTT

0,1M, 1µL de dNTP mix, 0,5µL da enzima *RNaseOut* e 1µL da enzima *MMLV-RT*. As reações foram incubadas em um termociclador por 40°C por 2hs, 65°C por 15min e transferidas para o gelo. A transcrição de cDNA para cRNA ocorreu em um volume final de 80µL por reação contendo 15,3µL de água ultra-pura estéril, 20µL de *4X Transcription Buffer*, 6,4µL de 50% PEG, 6µL de DDT 0,1M, 8µL de NTP mix, 0,5µL de *RNaseOUT*, 0,6µL de *Inorganic Pyrophosphatase*, 2,4µL de Cy3 ou Cy5 e 0,8 µL da enzima *T7 RNA Polimerase*. As amostras foram mantidas a 40°C por 2hs em termociclador sendo em seguida purificadas utilizando o protocolo estabelecido pelo *RNeasy Mini Kit* (Qiagen). A atividade específica da marcação foi obtida para todas amostras e 825ng de cada amostra teste foi combinada à mesma concentração da amostra referência. Foram adicionados 11µL de 10X Blocking Agent, 2,2µL de 25X Fragmentation Buffer e água ultra pura estéril para um volume final de 55µL. A solução foi desnaturada a 95°C por 3min, incubada a 37°C por 30min e 100µL foram aplicados às lâminas. As lâminas foram transferidas para o forno de hibridação por 17horas a 10rpm. Após este período, as imagens de hibridação de cada lâmina foram capturadas no Scanner Agilent G2565CA (Agilent Technologies). Os dados foram extraídos pelo programa *Feature Extraction v.10.1.1.1* (Agilent Technologies).

Para a obtenção do perfil de expressão gênica foi inicialmente avaliada a razão entre as fluorescências dos corantes Cy3 e Cy5 ($\log_2 Ratio$) em cada ponto da lâmina. Foi obtida a média geométrica da intensidade de sinais das sondas para todos os casos, gerando uma lista única de genes significantes. Os dados brutos foram normalizados pela centralização mediana dos genes de cada caso e sofreram transformação logarítmica. Além disso, os genes com 30% ou mais de *score* "ausentes" foram filtrados. Os restantes, aproximadamente 16 mil genes foram analisados. Esses dados serviram de entrada para o programa *TMeV v.4.5* (<http://www.tm4.org>), onde foi utilizado o método SAM (*Significance Analysis of Microarrays*) (Tusher et al., 2001) com 1000 permutações e <5% de falso-positivos (FDR). O SAM utiliza o teste *t* modificado por meio de permutações entre as amostras, para estimar a quantidade de resultados falso-positivos. O FDR (*False*

Discovery Ratio) é definido como a proporção esperada de “falsas descobertas” entre os genes considerados significantes. Em seguida, foi aplicada a análise de agrupamento não-supervisionado pelo teste de agrupamento hierárquico (HCL *Hierarchical Clustering*), utilizando a métrica de distância pela correlação de Pearson para verificar a existência de agrupamento entre os grupos de tumores e miométrios adjacentes.

Foram também obtidos os genes diferencialmente expressos em relação aos miométrios normais adjacentes. As análises de *fold-change* foram realizadas a partir da razão entre as médias geométricas das intensidades de expressão ($\log_2\text{Ratio}$) dos tumores em relação ao tecido normal. Para selecionar os genes potencialmente envolvidos com a doença, foram utilizados os *fold-changes* ≥ 1.5 e ≤ -1.5 , que indicam aumento e diminuição da expressão de determinado gene na amostra tumoral em relação ao tecido normal, respectivamente. Os genes resultantes foram submetidos à análise funcional.

3.7 Integração de dados genômicos e transcriptômicos

Para a realização da análise integrada, foram selecionados os dados da $\log_2\text{Ratio}$ obtidos da análise de CGH array (genômicos) e de expressão gênica (transcriptômicos). As 51 amostras submetidas à análise de expressão gênica e previamente avaliadas por CGH array, foram selecionadas para a análise integrada. Para a obtenção de genes significantes mapeados em regiões com alterações no número de cópias de DNA foi realizada a análise de segmentação dos dados das 51 amostras de Leiomiomas Uterinos, utilizando o pacote *DNA Copy* com o algoritmo JISTIC (Sanchez-Garcia et al., 2010). Esta nova análise dos dados de CGH array foi realizada devido ao fato de que o algoritmo utilizado para integração dos dados não é compatível com os dados gerados pelo programa *Nexus*. O JISTIC é baseado no GISTIC (Beroukhim et al., 2007) que utiliza um método estatístico (*G-score*) que representa a frequência e amplitude de cada alteração cromossômica. O JISTIC é superior ao GISTIC por permitir a distinção entre CNVs “passageiras” (*passengers*:

alterações que não são cruciais para o desenvolvimento tumoral) das CNVs “direcionadoras” (*drivers*: alterações que têm efeito direto com o fenótipo tumoral). A aplicação deste algoritmo é baseada na fórmula representada abaixo, onde G indica ganho ou perda, i é uma função indicadora e m é o marcador correspondente ao evento genômico:

$$\sum_i G_n(m,i) = G_n(m) > G_{thres}$$

Após a obtenção dos genes significantes pelo JISTIC, os dados foram comparados aos dados brutos do perfil de expressão gênica para identificação dos genes moduladores. O algoritmo CONEXIC (Akavia et al., 2010) é um método Bayesiano baseado em redes em módulos (Segal et al., 2003), mas que engloba uma série de modificações importantes que o tornam adequado para a identificação dos genes moduladores. O CONEXIC usa um guia de pontuação para identificar a combinação de moduladores que melhor explica o comportamento de um módulo de expressão gênica em amostras que possuem maior pontuação dentro das regiões de ganhos e perdas genômicas. Portanto, os genes moduladores envolvidos em ganhos e perdas em um número significativo de amostras podem contribuir com uma vantagem adaptativa para o tumor. As análises de expressão gênica e integrada foram realizadas com a colaboração de Fábio Albuquerque Marchi, Programa de Pós-Graduação Interinstitucional em Bioestatística, Instituto de Matemática e Estatística da Universidade de São Paulo, IME, Universidade de São Paulo, SP.

Os genes moduladores cujos dados de CGH array e expressão gênica não foram concordantes (ganho/diminuição de expressão ou perda/aumento de expressão) foram avaliados quanto à regulação por miRNAs (do inglês *Micro RNA*) com a utilização de duas

ferramentas de predição de alvos biológicos de miRNAs: *Target Scan* (<http://www.targetscan.org/>) e *PicTar* (<http://pictar.mdc-berlin.de/>).

3.8 Análise funcional - IPA

Os genes significantes obtidos a partir das análises de CGH array, aqueles diferencialmente expressos em relação ao miométrio adjacente e os resultantes da análise integrada, foram submetidos à análises funcionais utilizando o *Ingenuity Pathways Analysis* (IPA) (Ingenuity Systems, <http://www.ingenuity.com>). Abaixo, está apresentada a descrição de cada etapa das análises funcionais:

- 1) **Geração de redes:** uma lista de genes e o *status* correspondente (valores de expressão ou ganhos e perdas genômicos) foi importada ao sistema. O IPA identifica cada gene e seu *status* e associa-os com as informações depositadas nos bancos de dados. A maioria dos genes é classificada como “moléculas elegíveis para rede”. Utilizando algoritmos próprios do IPA, as redes das moléculas elegíveis foram geradas baseadas em suas conexões. Os dados foram representados em valores de *scores* obtidos a partir da presença dos genes importados em relação aos genes do IPA.
- 2) **Funções biológicas:** esta análise identificou as funções biológicas e doenças que foram mais significativamente observadas. O teste exato de Fisher foi utilizado para calcular o *P* valor que determinou a probabilidade de que cada função biológica e ou doença atribuídos a esse conjunto de dados fosse resultante do acaso ou que apresentasse algum significado estatístico. Os dados foram convertidos em $-\log(P \text{ valor})$.

- 3) **Vias canônicas:** esta análise identificou as vias depositadas no IPA que foram mais significativamente observadas entre os dados importados. A significância da associação entre os dados foram calculadas de duas formas: a) a razão entre o número de moléculas dos dados importados dividida pelo número total de moléculas das vias canônicas; b) teste exato de Fisher foi utilizado para calcular o *P* valor que determinou a probabilidade da associação entre os genes importados e as vias canônicas.
- 4) **Representação gráfica:** as redes são representadas graficamente para verificar as relações entre as moléculas. As moléculas são representadas como “nós” e as relações biológicas entre dois nós são representadas pelas “linhas”. Todas as linhas são baseadas em pelo menos uma referência da literatura ou nas informações canônicas depositadas no banco de dados do IPA. Genes ortólogos são armazenados separadamente no IPA, porém, são representados como nós únicos nas redes. No caso dos dados de expressão, variações na intensidade das cores nos nós indicam o grau de regulação gênica (vermelho=expressão aumentada e verde=expressão diminuída). Para os dados de CGH array, os genes apresentaram cores “fortes” de vermelho para ganhos e verde para perdas genômicas. Os nós são representados de várias formas geométricas que indicam a classe funcional do produto gênico. As linhas são representadas em duas formas principais, indicando a relação direta e indireta entre os nós.

4. Resultados

4.1 Características clínicas

A idade média à cirurgia das 56 pacientes deste estudo foi de 45 anos; 73% delas estavam na faixa etária entre 40 e 50 anos. A média do IMC foi de 27,7; 41% das pacientes apresentavam sobrepeso. Quarenta e oito pacientes (86%) se declararam brancas e oito (14%) não-brancas (negras ou mulatas). A idade média da menarca foi de 13 anos (dados não disponíveis em três casos) e a maioria das pacientes (73%) tiveram menarca antes dos 13 anos. A idade média das pacientes na primeira gravidez foi de 21 anos e a maioria (36%) teve sua primeira gestação em idade superior a 21 anos. O número médio de gestações foi de 3 e 62% das pacientes tiveram 3 ou menos gestações. Trinta e nove por cento das pacientes estavam no período secretor à cirurgia, 34% encontravam-se no período proliferativo e 27% apresentaram endométrio atrófico, basal ou autolisado, não sendo possível a classificação histológica. As pacientes apresentaram uma frequência maior de tumores múltiplos (77%) do que tumores únicos (23%). Metade das pacientes (52%) relatou história de câncer na família, entre as quais, sete tiveram história de LU na família. A Tabela 1 apresenta a distribuição dos casos de acordo com as variáveis clínicas.

Tabela 1. Dados clínicos obtidos das 56 pacientes com Leiomiomas Uterinos.

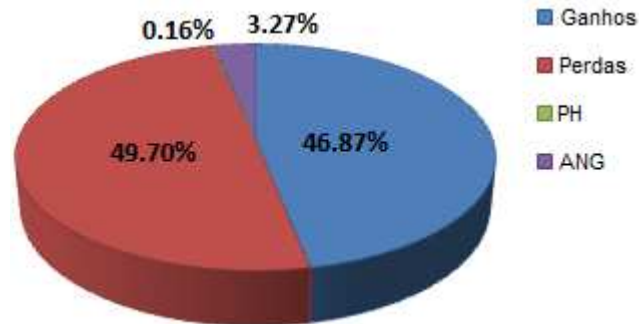
Dados	Categorias	n (%)
Idade	<40	10 (18)
	≥40 e ≤50	41 (73)
	>50	5 (9)
IMC	18,5 - 24,9	19 (34)
	25,0 - 29,9	23 (41)
	≥30,0	14 (25)
Grupo étnico	Branca	48 (86)
	Não-branca	8 (14)
Idade á menarca	≤13	41 (73)
	>13	12 (21)
	ND	3 (6)
Idade á primeira gestação	0	5 (9)
	≤21	14 (25)
	>21	20 (36)
	ND	17 (30)
Gestações	0	5 (9)
	≤3	35 (62)
	>3	15 (27)
	ND	1 (2)
Fase do ciclo	Proliferativo	19 (34)
	Secretor	22 (39)
	Outros	15 (27)
Número de tumores	Múltiplos	43 (77)
	Único	13 (23)
História familiar de câncer	Sim	29 (52)
	Não	26 (46)
	ND	1 (2)
História de LU na família	Sim	7 (12)
	Não	48 (86)
	ND	1 (2)

IMC=índice de massa corpórea; ND=não definido.

4.2 Caracterização geral das alterações no número de cópias de DNA

Foram detectadas 2.413 CNVs entre as 80 amostras de LU com média de $30,16 \pm 22,10$ CNVs por caso. Todas as amostras apresentaram CNVs, com números variáveis de três até 118 eventos por caso. A frequência de distribuição de ganhos e perdas genômicos entre os casos foi de 46,87% (1131) ganhos, 49,70% (1199) perdas, 0,16% (4) perdas homozigotas e 3,27% (79) alto nível de ganhos. O Anexo 3 mostra a distribuição das perdas homozigotas e do alto nível de ganhos entre as amostras. Foi observado um pequeno predomínio das perdas sobre os ganhos genômicos (Figura 1A). As alterações estavam distribuídas em todos os cromossomos. O cromossomo 1 apresentou maior frequência de perdas (129) e o cromossomo 11 apresentou o maior número de ganhos (95). O cromossomo 21 apresentou a menor frequência de perdas genômicas (26), enquanto que o cromossomo 13 tinha o menor número ganhos genômicos (18) (Figura 1B). O tamanho das CNVs variou entre 64Kb e 120Mb. O ideograma representativo de todas as alterações detectadas entre as 80 amostras de Leiomiomas Uterinos está apresentado na Figura 2.

A



B

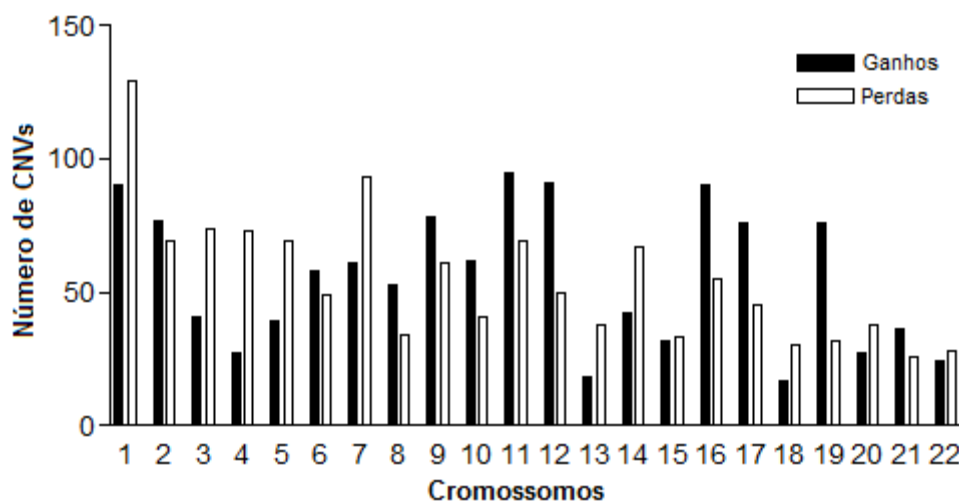


Figura 1. A – Frequência das alterações no número de cópias identificadas entre as 80 amostras de Leiomiomas Uterinos; PH = perdas homozigotas; ANG = alto nível de ganhos. É possível observar uma pequena prevalência das perdas sobre ganhos genômicos. **B** – Número de alterações no número de cópias identificadas ao longo dos cromossomos. O cromossomo 1 apresentou um maior número de perdas genômicas e o cromossomo 11 teve um maior número de ganhos.

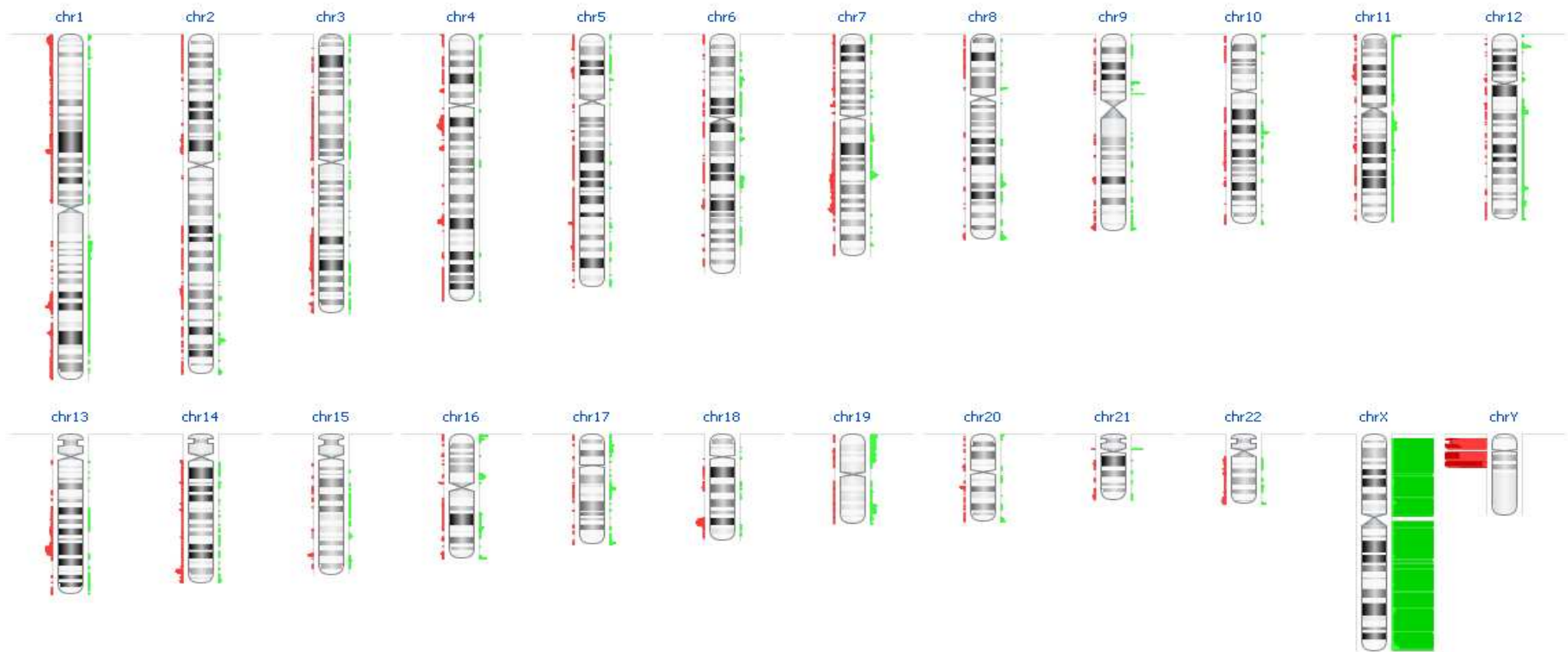


Figura 2. Ideograma representando o perfil geral das alterações genômicas detectadas nas 80 amostras de LU avaliadas. As perdas genômicas estão representadas ao lado esquerdo dos cromossomos (vermelho) e os ganhos genômicos ao lado direito (verde) (*Nexus v.5.0 -Biodiscovery*).

Este estudo revelou alterações genômicas estruturais envolvendo grandes segmentos (>10Mb), nos cromossomos 1 e 19. Em dois casos foram identificados um possível cromossomo 1 em anel. O caso 084T apresentou o rearranjo ?r(1)(p36.11q42.12) e o 759T o ?r(1)(p31.1q32.2) (Figura 3). Além disso, dois casos apresentaram a perda de todo o braço (p) curto do cromossomo 1 (046T_A e 1015T) e um caso (629T_B), apresentou perda de quase todo o braço curto do cromossomo 1: del(1) (p36.33p22.1). Dez casos (300T_A, 301T_B, 301T_C, 304T, 317T_C, 629T_A, 631T_B, 643T_A, 733T_A, 733T_B) apresentaram ganhos do braço curto do cromossomo 19: +(19)(p13.3p12) (Figura 4).

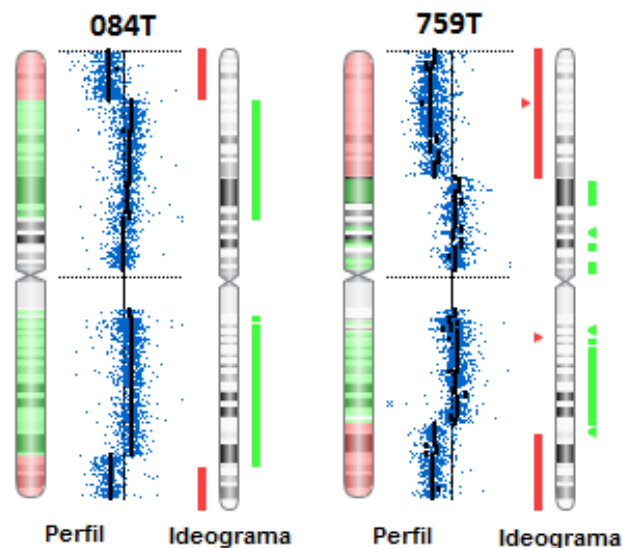


Figura 3. Rearranjos cromossômicos resultando em um possível cromossomo em anel nos casos 084T: ?r(1)(p36.11q42.12) e 759T: ?r(1)(p31.1q32.2). Os pontos azuis representam os sinais de hibridação. Ao lado direito do ideograma está representada a região envolvida em ganhos (verde) e ao lado esquerdo, as perdas (vermelho) (Nexus v.5.0 -Biodiscovery).



Figura 4. Ganhos genômicos detectados em 19p, com região mínima comum em (p13.3p12) (região selecionada em retângulo vermelho). Ao lado direito do eixo observa-se a frequência das alterações. O tamanho de cada alteração está representado acima de cada reta, no qual representa o número da amostra afetada (300T_A, 301T_B, 301T_C, 304T, 317T_C, 629T_A, 631T_B, 643T_A, 733T_A, 733T_B) (Nexus v.5.0 -Biodiscovery).

4.2.1 Caracterização das regiões significativas

As alterações detectadas em mais de 10% dos casos foram selecionadas e comparadas ao banco contendo dados genômicos de 82 mulheres brasileiras saudáveis. Regiões alteradas presentes em mais de 5% da amostra de referência foram excluídas da análise (McCarroll et al., 2008) e as CNVs restantes foram selecionadas para estudo em maiores detalhes. Esta análise resultou em 45 CNVs significativas distribuídas entre os LU avaliados. A Figura 5 representa as alterações significativas (asteriscos) distribuídas pelos cromossomos de todos os LU analisados. Segundo estes critérios, três amostras (642T_B, 683T_A, 684T) apresentaram perfis genômicos normais e a média de CNVs observada foi de $6,76 \pm 6,0$ por caso. A amostra 643T_A apresentou o maior número de CNVs significativas (25) e oito amostras (085T, 092T, 614T, 631T_A, 643T_B, 708T, 719T, 729T_A) apresentaram apenas uma CNV significativa. A frequência de CNVs significativas variou entre 10% e 25% entre as amostras estudadas.

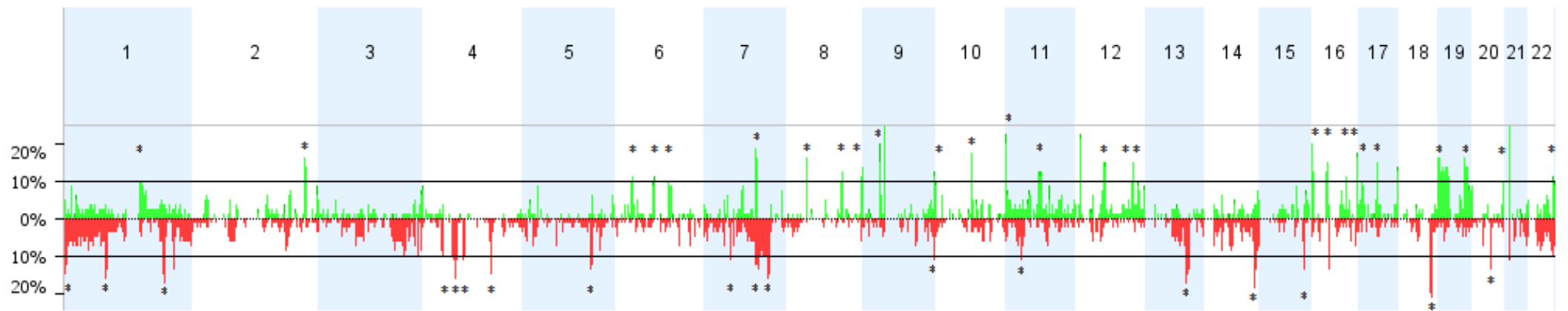


Figura 5. Perfil genômico global onde podem ser observadas as alterações significativas ($P \leq 0.05$) indicada pelos asteriscos. Estas regiões foram detectadas em $\geq 10\%$ dos casos. Foram excluídas as regiões alteradas presentes em $\geq 5\%$ da população de referência. Os cromossomos sexuais foram excluídos da análise (*Nexus v.5.0 - Biodiscovery*).

Entre as 45 alterações genômicas significativas, 27 estavam envolvidas em ganhos e 18 em perdas genômicas. Entre as 80 amostras de LU analisados, os **ganhos** genômicos foram identificados em 1q (8 amostras), 2q (15 amostras), 6p e 6q (10 e 16 amostras, respectivamente), 7q (15 amostras), 8p e 8q (13 e 22 amostras, respectivamente), 9p (17 amostras), 10p e 10q (8 e 14 amostras, respectivamente), 11p e 11q (19 e 20 amostras, respectivamente), 12q (29 amostras), 16p e 16q (20 e 23 amostras, respectivamente), 17p e 17q (8 e 12 amostras, respectivamente), 19p e 19q (9 e 15 casos, respectivamente), 20q (8 amostras) e 22q (9 casos). O tamanho destas CNVs variou entre 74Kb a 5Mb e apresentou o envolvimento de 1.189 genes e 28 miRNAs. As 18 **perdas** genômicas significativas envolveram 1p e 1q (25 e 11 amostras, respectivamente), 4p e 4q (8 e 23 amostras, respectivamente), 5q (14 amostras), 7p e 7q (10 e 19 amostras, respectivamente), 9q (9 amostras), 11p14.1 (9 amostras), 13q (14 amostras), 14q (16 amostras), 15q (11 amostras), 18q (17 amostras) e 20q (11 amostras). O tamanho das perdas genômicas variaram de 271Kb a 7Mb e envolveram 134 genes e três miRNAs. Entre estes achados, 16 regiões genômicas foram previamente relatadas em outros estudos de LU, além de 29 regiões ainda não descritas. O Anexo 4 apresenta os genes e miRNAs mapeados nas regiões significativas de ganhos e perdas genômicos. A Tabela 2 apresenta as alterações genômicas significativas detectadas entre os 80 LU analisados. A Figura 6 mostra algumas das alterações mais frequentemente observadas entre as 80 amostras de LU, como ganhos em 2q35, 6p21.32, 7q22.1 e 11q12.3-q13.2 e perdas em 4q28.3, 7q22.1-q22.2 e 7q31.33.

Tabela 2. Regiões significativas de perdas e ganhos genômicos detectadas por CGH array entre as amostras de Leiomiomas Uterinos.

Região mínima comum	Start	End	Tamanho (bp)	Evento	Genes	miRNAs	P Valor	Amostras (%)	% de CNVs
<i>Cromossomo 1</i> p36.33 - p36.31	1,921,978	5,577,281	3.655.303	Perda	32	1	0.02	046T_A, 084T, 147T_C, 303T, 629T_B, 630T_A, 630T_B, 633T, 670T, 711T, 722T_A, 728T, 759T, 845T_B, 1001T, 1015T (20%)	96.4
p31.1	82,163,689	84,655,179	2.491.490	Perda	6	0	0.02	046T_A, 046T_B, 147T_C, 629T_B, 722T_B, 722T_C, 729T_B, 729T_C, 733T_A, 845T_A, 948T, 1003T, 1015T (16,25%)	32.9
q21.3	149,836,310	150,010,690	174.380	Ganho	6	0	<0.01	084T, 301T_B, 317T_C, 631T_A, 643T_A, 759T, 854T, 1015T (10%)	0
q41	212,893,807	214,014,200	1.120.393	Perda	5	0	0.01	301T_C, 304T, 317T_C, 630T_B, 631T_B, 634T, 733T_A, 759T, 845T_B, 948T, 954T (13,75%)	4.0
<i>Cromossomo 2</i> q35	217,253,192	219,351,693	2.098.501	Ganho	26	1	<0.01	300T_B, 301T_A, 301T_B, 304T, 307T_B, 317T_B, 317T_C, 630T_B, 643T_A, 643T_B, 643T_C, 733T_A, 733T_B, 842T, 845T_B (18,75%)	1.1
<i>Cromossomo 4</i> p14	40,185,743	40,457,450	271.707	Perda	2	0	<0.01	642T_A, 719T, 759T, 761T, 790T, 855T, 901T, 1005T (10%)	78.1
q13.1 - q13.2	60,129,993	68,031,967	7.901.974	Perda	4	1	0.01	046T_B, 147T_B, 301T_B, 301T_C, 304T, 307T_A, 307T_B, 643T_A, 729T_C, 733T_A, 845T_A, 845T_B, 854T, 1003T (17,5%)	15.8
q21.21	80,461,995	82,110,583	1.648.588	Perda	7	0	0.01	046T_B, 301T_B, 303T, 304T, 643T_A, 722T_A, 729T_C, 845T_A, 948T (11,25%)	12.4
q28.3	132,220,774	135,650,424	3.429.650	Perda	2	0	<0.01	147T_B, 301T_B, 301T_C, 317T_C, 643T_A, 722T_A, 729T_B, 845T_B, 854T, 867T, 954T, 1003T (15%)	62.6
<i>Cromossomo 5</i> q31.1 - q31.2	134,274,907	135,585,827	1.310.920	Perda	19	1	<0.01	301T_C, 307T_B, 317T_C, 615T, 629T_A, 630T_A, 630T_B, 631T_B, 631T_C, 733T_A, 759T, 761T, 844T_B, 1015T (17,5%)	33.6
<i>Cromossomo 6</i> p21.32	31,962,130	32,132,073	169.943	Ganho	17	1	<0.01	300T_A, 317T_C, 631T_B, 643T_A, 643T_C, 711T, 733T_B, 790T, 844T_A, 1015T (12,25%)	91.7
q13	73,782,170	74,568,779	786.609	Ganho	13	0	<0.01	091T, 303T, 683T_B, 729T_C, 733T_A, 844T_B, 845T_B, 948T, 1001T (11,25%)	1.7
q16.3	102,041,584	103,869,362	1.827.778	Ganho	1	0	<0.01	085T, 091T, 147T_A, 644T, 670T, 769T, 901T, 1001T	35.5

(10%)

<i>Cromossomo 7</i>									
p12.2	50,138,493	50,702,500	564.007	Perda	5	0	<0.01	046T_B, 147T_B, 147T_C, 631T_B, 643T_A, 681T, 733T_B, 844T_B, 845T_A, 948T (12,5%)	12.2
q22.1	98,769,492	102,068,500	3.299.008	Ganho	98	4	<0.01	300T_A, 301T_B, 301T_C, 304T, 307T_A, 307T_B, 317T_A, 317T_C, 629T_A, 631T_B, 643T_A, 733T_B, 845T_B, 1003T, 1015T (18,75%)	73.1
q22.1 - q22.2	104,284,386	105,139,642	855.256	Perda	9	0	0.02	092T, 631T_C, 733T_A, 759T, 769T, 844T_A, 845T_A, 853T, 855T, 942T, 1005T (13,75%)	56.8
q31.33	124,268,700	125,951,548	1.682.848	Perda	2	0	<0.01	317T_C, 631T_C, 722T_B, 729T_C, 733T_A, 769T, 844T_A, 845T_A, 853T, 948T, 1002T, 1003T, 1015T (16,25%)	58.2
<i>Cromossomo 8</i>									
p12 - p11.23	37,606,006	38,685,823	1.079.817	Ganho	21	0	<0.01	046T_B, 147T_B, 147T_C, 301T_B, 301T_C, 303T, 304T, 643T_A, 722T_A, 722T_C, 845T_A, 867T, 948T (16,25%)	16.0
q22.3 - q23.1	105,831,966	107,742,335	1.910.369	Ganho	2	0	<0.01	046T_A, 147T_A, 615T, 634T, 644T, 670T, 769T, 853T, 857T, 947T (12,5%)	5.6
q24.3	143,989,891	146,027,963	2.038.072	Ganho	80	4	<0.01	300T_A, 301T_A, 301T_C, 304T, 625T, 629T_A, 631T_B, 643T_A, 729T_B, 733T_B, 845T_A, 1015T (15%)	100.0
<i>Cromossomo 9</i>									
p13.3	33,871,385	34,349,039	477.654	Ganho	8	0	<0.01	301T_C, 615T, 625T, 644T, 670T, 681T, 683T_B, 708T, 722T_B, 729T_B, 729T_C, 769T, 857T, 948T, 954T, 1001T, 1003T (21,25%)	10.0
q34.3	137,569,888	138,507,449	937.561	Perda	23	0	<0.01	084T, 630T_A, 630T_B, 633T, 670T, 722T_A, 722T_C, 729T_A, 1001T (11,25%)	69.7
<i>Cromossomo 10</i>									
p15.3	0	1,064,364	1.064.364	Ganho	8	0	<0.01	615T, 630T_A, 670T, 683T_B, 729T_C, 954T, 1001T, 1003T (10%)	41.9
q21.3	68,975,747	70,562,567	1.586.820	Ganho	20	1	<0.01	046T_B, 147T_C, 301T_B, 304T, 317T_C, 629T_A, 630T_B, 631T_B, 722T_A, 844T_B, 845T_A, 947T, 948T, 1001T (17,5%)	23.9
<i>Cromossomo 11</i>									
p15.5	0	1,034,144	1.034.144	Ganho	48	1	0.01	046T_A, 300T_A, 301T_C, 303T, 317T_C, 625T, 629T_A, 631T_B, 642T_A, 643T_A, 729T_B, 733T_A, 733T_B, 842T, 844T_A, 845T_A, 845T_B, 1005T, 1015T (23,75%)	94.3
p14.1	29,114,213	29,511,311	397.098	Perda	0	0	0.02	304T, 317T_C, 631T_B, 643T_C, 722T_B, 728T, 733T_A, 733T_B, 854T (11,25%)	33.6
q12.3 - q13.2	62,133,411	67,979,861	5.846.450	Ganho	131	3	0.02	046T_A, 301T_B, 301T_C, 317T_A, 625T, 630T_B, 629T_A, 631T_B, 633T, 642T_A, 643T_A, 643T_C, 722T_B, 733T_B, 790T, 842T, 845T_B, 954T, 1001T,	100.0

								1015T (25%)	
<i>Cromossomo 12</i>									
q13.2 - q13.3	54,386,743	56,279,267	1.892.524	Ganho	75	2	0.04	147T_C, 301T_B, 317T_C, 633T, 643T_A, 722T_B, 729T_B, 729T_C, 733T_A, 733T_B, 842T, 1001T, 1003T, 1015T (17,5%)	34.1
q24.11 - q24.12	108,897,746	110,392,895	1.495.149	Ganho	21	0	0.04	147T_C, 300T_A, 300T_B, 303T, 317T_C, 722T_A, 728T, 845T_B, 854T, 948T, 954T, 1015T (15,0)	8.7
q24.31	119,131,582	119,853,673	722.091	Ganho	18	0	0.04	147T_C, 300T_A, 304T, 317T_C, 629T_A, 643T_A, 845T_B, 1002T (10%)	7.3
<i>Cromossomo 13</i>									
q31.1	79,882,431	85,782,478	5.900.047	Perda	3	0	<0.01	300T_A, 317T_C, 643T_A, 683T_B, 722T_B, 728T, 729T_C, 733T_A, 845T_A, 845T_B, 854T, 948T, 954T, 1003T (17,5%)	34.9
<i>Cromossomo 14</i>									
q32.2	96,180,194	99,038,571	2.858.377	Perda	6	0	0.03	046T_A, 147T_B, 147T_C, 301T_B, 301T_C, 629T_B, 630T_A, 633T, 722T_B, 845T_A, 845T_B, 942T, 948T, 954T, 1001T, 1003T (20%)	11.7
<i>Cromossomo 15</i>									
q25.3	85,542,221	86,595,151	1.052.930	Perda	2	0	<0.01	046T_B, 147T_B, 147T_C, 722T_C, 729T_B, 729T_C, 733T_A, 845T_A, 948T, 1001T, 1003T (13,75%)	12.9
<i>Cromossomo 16</i>									
p13.3	2,756,969	3,081,334	324.365	Ganho	22	0	0.04	317T_B, 625T, 631T_B, 631T_C, 643T_A, 643T_C, 733T_B, 842T, 844T_A, 1015T (12,5%)	45.8
p11.2	28,184,420	31,429,424	3.245.004	Ganho	110	0	0.04	300T_A, 301T_B, 301T_C, 303T, 304T, 307T_B, 317T_C, 629T_A, 629T_B, 643T_A, 733T_B, 845T_B, 1015T (16,25%)	86.8
q22.1	65,732,841	66,551,504	818.663	Ganho	32	1	<0.01	300T_A, 300T_B, 301T_C, 304T, 317T_C, 614T, 625T, 631T_B, 733T_B, 790T, 1001T, 1015T (15%)	26.8
q24.3	88,134,943	88,526,617	391.674	Ganho	16	0	<0.01	301T_C, 303T, 304T, 307T_A, 631T_B, 643T_A, 683T_B, 722T_C, 729T_B, 729T_C, 733T_B, 845T_A, 845T_B, 854T, 948T (18,75%)	74.6
<i>Cromossomo 17</i>									
p13.1	7,431,136	7,505,810	74.674	Ganho	6	0	0.04	300T_A, 304T, 317T_C, 625T, 642T_A, 643T_A, 733T_B, 1015T (10%)	10.5
q21.31	38,105,718	39,165,518	1.059.800	Ganho	30	0	<0.01	301T_C, 304T, 631T_B, 643T_A, 681T, 722T_B, 722T_C, 729T_B, 845T_B, 948T, 1001T, 1015T (15%)	50.5
<i>Cromossomo 18</i>									
q22.1 - q22.2	60,070,073	65,021,204	4.951.131	Perda	5	0	<0.01	147T_C, 300T_A, 301T_B, 301T_C, 307T_A, 307T_B, 317T_C, 643T_A, 722T_B, 729T_B, 729T_C, 845T_A, 845T_B, 857T, 948T, 1001T, 1003T (21,25%)	36.2

<i>Cromossomo 19</i>									
p13.12 - p13.11	15,989,861	16,541,429	551.568	Ganho	14	0	<0.01	300T_A, 301T_B, 301T_C, 304T, 317T_C, 629T_A, 631T_B, 643T_A, 733T_B (11,25%)	8.0
q13.32 - q13.33	50,012,054	55,753,381	5.741.327	Ganho	315	5	0.02	300T_A, 301T_A, 301T_B, 301T_C, 303T, 304T, 307T_A, 317T_C, 625T, 629T_A, 629T_B, 631T_B, 643T_A, 733T_B, 845T_B (18,75%)	100.0
<i>Cromossomo 20</i>									
q11.23 - q12	37,049,227	38,416,306	1.367.079	Perda	2	0	<0.01	046T_B, 301T_B, 301T_C, 303T, 304T, 317T_C, 631T_B, 845T_A, 845T_B, 867T, 948T (13,75%)	17.1
q13.33	61,881,839	62,198,449	316.610	Ganho	25	5	<0.01	625T, 643T_A, 683T_B, 729T_B, 733T_B, 844T_A, 845T_A, 1015T (10%)	88.8
<i>Cromossomo 22</i>									
q13.33	49,041,686	49,691,432	649.746	Ganho	25	0	<0.01	300T_A, 301T_B, 629T_A, 631T_B, 643T_A, 643T_C, 733T_B, 842T, 845T_A (11,25%)	49.2

*Baseado no DGV - Database of Genomic Variations; pb:pares de bases; CNV = variação no número de cópias. Em negrito, alterações genômicas identificadas previamente em Leiomiomas Uterinos (Packenham et al., 1997; Levy et al., 2000; Rogatto et al., 2003; Vanharanta et al., 2007; Hodge et al., 2009).

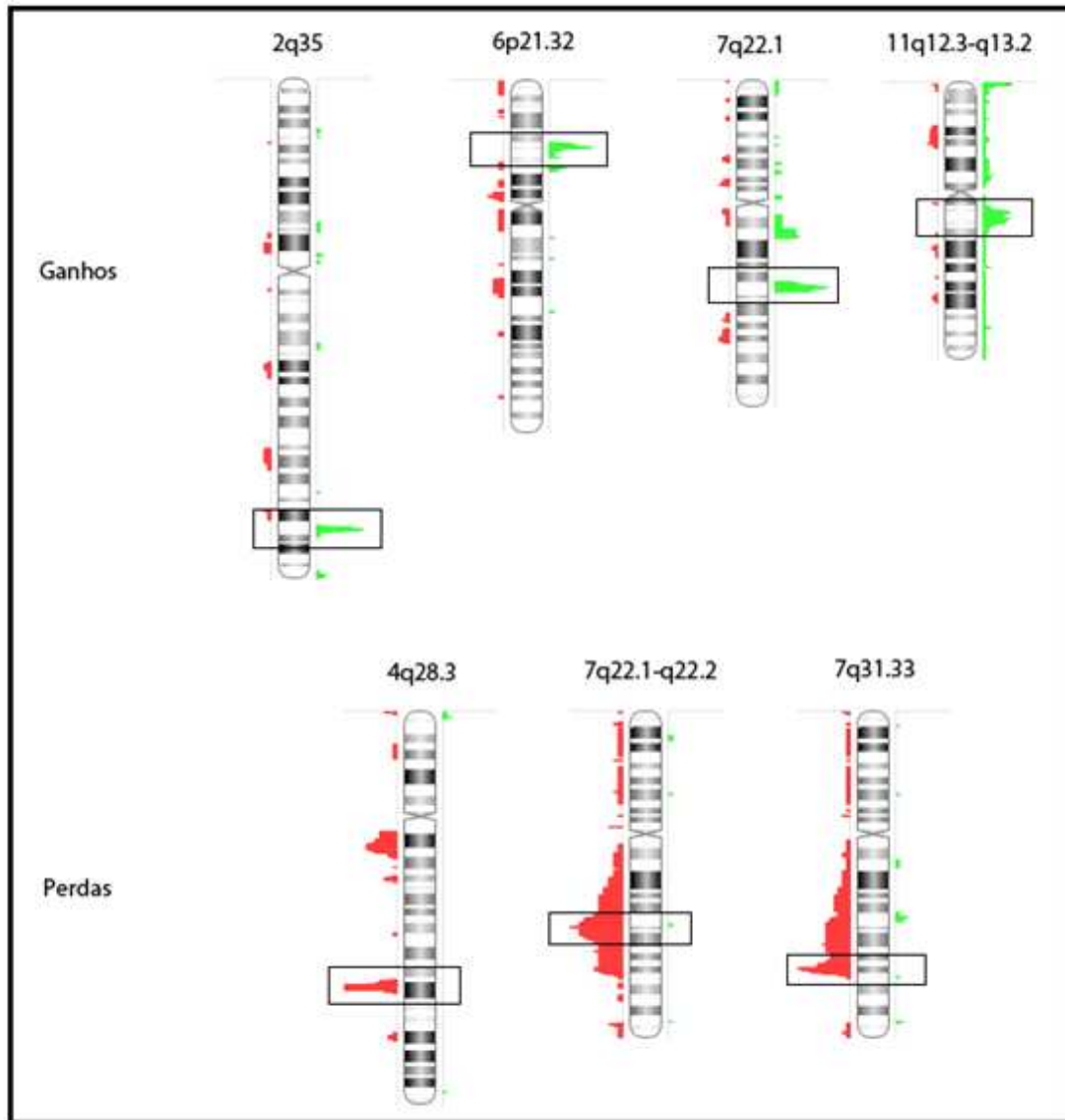


Figura 6. Regiões significativas mais frequentemente observadas entre as 80 amostras de LU, como ganhos em 2q35, 6p21.32, 7q22.1 e 11q12.3-q13.2 e perdas em 4q28.3, 7q22.1-q22.2 e 7q31.33 (região selecionada com o retângulo). Ao lado direito do ideograma está representada a região envolvida em ganhos (verde) e ao lado esquerdo, as perdas (vermelho). A figura foi gerada a partir dos casos que apresentaram tais alterações (*Nexus v.5.0 -Biodiscovery*).

4.2.2 Tumores múltiplos

Foram investigadas a frequência de similaridade de alterações entre amostras de tumores múltiplos (mais de um tumor avaliado; n=41). Inicialmente, foram consideradas apenas as regiões descritas com os critérios utilizados para a identificação de CNVs excluindo-se aquelas presentes em >5% da população controle. As regiões em comum em pelo menos duas amostras foram selecionadas e foi calculada a frequência de similaridade destas regiões entre os diferentes tumores (Tabela 3). A frequência de similaridade quanto a presença destas alterações variou entre as amostras de 0% (683T) até 13,7% (301T). Os genes e os miRNAs comuns em pelo menos duas amostras de tumores múltiplos do mesmo paciente estão descritos na Tabela 4. Ganhos em 2q35, 7q22.1, 8p12, 11q13.1-q13.2, 11p15.5, 12p13.31, 16p13.3 e perdas em 1p31.1, 3q28, 4q13.1, 5q31.1, 9q34.2-q34.3, 11p15.5, 14q32.2, 15q25.3, 16q24.3, 18q22.1-q22.2 foram as alterações observadas presentes em mais de um caso com tumores múltiplos. A Figura 7A mostra o exemplo do caso 845T, onde foi possível observar 13 regiões alteradas em comum e a Figura 7B mostra o exemplo do caso 683T, onde não foram observadas regiões alteradas em comum. Quando foram selecionadas apenas as regiões de CNVs significativas, (regiões descritas na Tabela 2), foram observados o mesmo padrão de alterações distribuídas entre as diferentes amostras de uma mesma paciente (Tabela 5). As amostras do caso 683T continuaram não apresentando similaridade de alterações, as amostras do caso 845T apresentaram sete regiões significativas em comum e as amostras do caso 301T foram as que tiveram o maior número de regiões significativas em comum, 11 regiões.

Tabela 3. Frequência de similaridade das alterações genômicas detectadas entre as diferentes amostras de tumores múltiplos de uma mesma paciente.

Casos	Amostras	*Número de Alterações	Alterações em comum	Frequência similaridade (%)
046T	046T A e B	41	3	7,3
147T	147T A, B e C	64	7	10,9
300T	300T A e B	42	2	4,8
301T	301T A, B e C	126	17	13,5
307T	307T A e B	44	5	11,4
317T	317T A, B e C	67	2	3,0
629T	629T A e B	104	2	1,9
630T	630T A e B	118	6	5,1
631T	631T A, B e C	102	6	5,9
642T	642T A e B	38	1	2,6
643T	643T A, B e C	108	5	4,6
683T	683T A e B	36	0	0,0
722T	722T A, B e C	99	7	7,1
729T	729T A, B e C	61	8	13,1
733T	733T A, B e C	104	6	5,8
844T	844T A e B	51	1	2,0
845T	845T A e B	95	13	13,7

*A descrição destas alterações foi baseada apenas nos critérios de corte para identificação de CNVs deste estudo: limiar de significância de $1,00E^{-4}$, espaçamento máximo de 1Mb entre sondas adjacentes e a presença de três sondas consecutivas alteradas para determinação de um segmento de DNA alterado e exclusão de CNVs presentes em >5% da população controle.

Tabela 4. Descrição das alterações genômicas, genes e miRNAs em comum entre diferentes amostras de tumores múltiplos.

Caso	Amostras	Posição	Região	Evento	Genes	miRNA
046T	046T_A, 046T_B	82,163,689-84,463,792	1p31.1	Perda	<i>LPHN2, TTLL7, PRKACB</i>	
		31,780,604-33,040,650	11p13	Ganho	<i>PAX6, RCN1, WT1, WIT1, EIF3M, CCDC73, PRRG4, QSER1, DEPDC7, TCP11L1</i>	
		77,044,079-77,750,925	3q14.1	Ganho	<i>RSF1, C11orf67, INTS4, KCTD14, THRSP, NDUFC2, ALG8, KCTD21, USP35, GAB2</i>	
147T	147T_A, 147T_B, 147T_C	191,649,580-193,674,663	3q28	Perda	<i>TMEM207, IL1RAP, LOC647309, SNAR-I, OSTN, UTS2D, CCDC50, PYDC2, FGF12</i>	
		50,138,493-50,750,119	7p12.2	Perda	<i>C7orf72, IKZF1, FIGNL1, DDC, GRB10</i>	
		37,606,006-38,055,296	8p12	Ganho	<i>ZNF703, ERLIN2, LOC728024, PROSC, GPR124, BRF2, RAB11FIP1, GOT1L1, ADRB3, EIF4EBP1</i>	
		7,856,996-9,042,700	12p13.31	Ganho	>10	
		67,714,013-69,095,797	12q15	Ganho	>10	hsa-mir-1279
		96,180,194-98,643,585	14q32.2	Perda	<i>VRK1, C14orf64, C14orf177</i>	
300T	300T_A, 300T_B	85,542,221-86,595,151	15q25.3	Perda	<i>NCRNA00052, NTRK3</i>	
		109,668,284-110,392,895	12q24.11-q24.12	Ganho	<i>CCDC63, MYL2, CUX2, FAM109A, SH2B3, ATXN2</i>	
301T	301T_A, 301T_B, 301T_C	65,977,069-66,528,910	16q22.1	Ganho	>10	
		217,245,072-218,517,041	2q35	Ganho	<i>IGFBP5, TNP1, DIRC3, TNS1</i>	
		72,506,438-73,756,762	3p13-p11.1	Perda	<i>RYBP, SHQ1, GXYLT2, PPP4R2, EBLN2, PDZRN3</i>	
		62,045,433-68,093,851	4q12-q13.2	Perda	<i>LPHN3, TECRL, EPHA5, CENPC1</i>	hsa-mir-1269
		134,289,919-135,342,353	7q28.3	Perda	<i>PCDH10, PABPC4L</i>	
		97,748,914-102,099,418	7q21.3-q22.1	Ganho	>10	hsa-mir-25, hsa-mir-93, hsa-mir-

		37,672,458-38,445,509	8p12	Ganho	>10	106b, hsa-mir-548o
		143,542,378-144,694,763	8q24.3	Ganho	>10	
		102,830,851-104,820,591	9q31.1	Perda	<i>LPPR1, BAAT, MRPL50, ZNF189, ALDOB, C9orf125, RNF20, PPP3R2, GRIN3A, CYLC2</i>	
		63,098,519-67,973,319	11q13.1-q13.2	Ganho	>10	hsa-mir-1237, hsa-mir-192, hsa-mir-194-2, hsa-mir-612
		53,486,204-53,956,682	14q22.2	Perda	<i>BMP4, CDKN3</i>	
		96,333,436-98,253,850	14q32.2	Perda	<i>VRK1, C14orf64, C14orf177</i>	
		27,706,350-31,427,207	16p11.2	Ganho	>10	
		35,040,885-39,111,613	18q12.2-q12.3	Perda	<i>LOC647946, KC6, PIK3C3, RIT2, SYT4</i>	
		61,568,467-64,873,406	18q22.1-q22.2	Perda	<i>CDH7, CDH19, DSEL, TMX3, CCDC102B</i>	hsa-mir-924
		11,950-19,635,503	19p13.3-p13.11	Ganho	>10	>10
		49,943,817-55,834,904	19q13.31-q13.33	Ganho	>10	hsa-mir-640, hsa-mir-330, hsa-mir-642, hsa-mir-769, hsa-mir-220c
		36,334,426-37,399,801	22q13.1	Ganho	>10	hsa-mir-658, hsa-mir-659
307T	307T_A, 307T_B	138,387,269-169,581,585	3q22.3-q26.2	Perda	>10	hsa-mir-15b, hsa-mir-16-2
		175,499,839-178,457,613	3q26.31-q26.32	Perda	>10	hsa-mir-1263, hsa-mir-720
		60,129,993-66,171,072	4q13.1	Perda	<i>LPHN3, TECRL, EPHA5</i>	
		98,882,914-102,068,500	7q22.1	Ganho	>10	hsa-mir-25, hsa-mir-93, hsa-mir-106b, hsa-mir-548o
		62,957,157-63,985,142	18q22.1	Perda	<i>DSEL</i>	
317T	317T_A, 317T_B, 317T_C	218,505,630-218,829,850	2q35	Ganho	<i>TNS1, CXCR2P1, RUFY4, CXCR2, CXCR1, ARPC2</i>	
		99,832,729-100,646,532	7q22.1	Ganho	>10	

629T	629T_A, 629T_B	27,614,937-28,749,072 50,542,897-50,694,179	1p36.11-p35.3 19q13.32	Perda Ganho	>10 <i>KLC3, ERCC2, PPP1R13L, CD3EAP, ERCC1, FOSB, RTN2, PPM1N</i>	
630T	630T_A, 630T_B	1,921,978-3,562,051 134,304,814-135,244,438	1p36.33-p36.32 5q31.1	Perda Perda	>10 <i>PCBD2, CATSPER3, PITX1, H2AFY, C5orf20, TIFAB, NEUROG1, CXCL14, LOC340074, LOC153328</i>	hsa-mir-551a
		135,413,654-138,828,862 932,004-1,885,673	9q34.2-q34.3 11p15.5	Perda Perda	>10 >10	hsa-mir-126
		104,210,820-105,455,157 86,328,975-87,812,143	14q32.33 16q24.2-q24.3	Perda Perda	>10 >10	
631T	631T_A, 631T_B, 631T_C	151,635,029-153,290,183 115,985,067-123,774,540 132,261,512-132,601,566 36,936,474-39,321,782	1q21.3 5q23.1-q23.2 5q31.1 7p14.2-p14.1	Ganho Perda Perda Perda	>10 >10 <i>AFF4, ZCCHC10, HSPA4, FSTL4</i> <i>ELMO1, GPR141, TXNDC3, SFRP4, EPDR1, STARD3NL, TARP, AMPH, FAM183B, VPS41, POU6F2</i>	hsa-mir-190b hsa-mir-1244
		51,471,512-52,465,026 2,792,750-3,116,060	13q14.3-q21.1 16p13.3	Perda Ganho	>10 >10	
642T	642T_A, 642T_B	17,149,220-19,702,891	22q11.21	Perda	>10	hsa-mir-185, hsa-mir-1306, hsa-mir-1286
643T	643T_A, 643T_B, 643T_C	217,253,192-218,913,319 66,882,158-67,570,565 1,817,234-3,028,512 14,028,233-15,460,172	2q35 11q13.1-q13.2 16p13.3 17p12	Ganho Ganho Ganho Ganho	>10 >10 >10 <i>COX10, CDRT15, MGC12916, HS3ST3B1, CDRT7, PMP22, TEKT3, CDRT4, FAM18B2,</i>	hsa-mir-1225, hsa-mir-940

					<i>CDRT1</i>	
		49,041,686-49,691,432	22q13.33	Ganho	>10	
722T	722T_A, 722T_B, 722T_C	82,229,267-84,682,403	1p31.1	Perda	<i>LPHN2, TTLL7, PRKACB, SAMD13, UOX, DNASE2B</i>	
		181,930,714-182,786,793	2q31.3-q32.1	Perda	<i>ITGA4, CERKL, NEUROD1, SSFA2, PPP1R1C, PDE1A</i>	
		37,694,718-38,685,823	8p12-p11.23	Ganho	>10	
		135,310,865-138,789,316	9q34.2-q34.3	Perda	>10	hsa-mir-126
		13,121,205-13,562,391	10p13	Ganho	<i>OPTN, MCM10, UCMA, PHYH, SEPHS1, BEND7</i>	
		7,620,638-8,615,932	12p13.31	Ganho	>10	
		38,218,690-39,051,741	17q21.31	Ganho	>10	
		46,795,328-48,579,667	22q13.31-q13.33	Perda	<i>FAM19A5, C22orf34, LOC90834, BRD1</i>	
729T	729T_A, 729T_B, 729T_C	82,163,689-84,655,179	1p31.1	Perda	<i>LPHN2, TTLL7, PRKACB, SAMD13, UOX, DNASE2B</i>	
		33,871,385-34,274,455	9p13.3	Ganho	<i>UBE2R2, SNORD121B, SNORD121A, UBAP2, DCAF12, UBAP1, KIF24</i>	
		7,530,089-8,720,997	12p13.31	Ganho	>10	
		54,386,743-55,435,586	12q13.2-q13.3	Ganho	>10	
		57,617,329-58,305,867	14q23.1	Ganho	<i>C14orf37, ACTR10, PSMA3, FLJ31306, ARID4A, TOMM20L, TIMM9, KIAA0586, DACT1</i>	
		85,542,221-86,520,726	15q25.3	Perda	<i>NCRNA00052, NTRK3</i>	
		88,329,390-88,827,254	16q24.3	Ganho	>10	
		60,070,073-65,021,204	18q22.1-q22.2	Perda	<i>CDH7, CDH19, DSEL, TMX3, CCDC102B</i>	
733T	733T_A, 733T_B	151,621,154-155,073,014	1q21.3-q23.1	Ganho	>10	hsa-mir-190b, hsa-mir-92b, hsa-mir-555, hsa-mir-9-1
		218,879,301-219,385,190	2q35	Ganho	>10	hsa-mir-26b

		279,263-1,481,453	11p15.5	Ganho	>10	hsa-mir-210
		29,009,427-29,511,311	11p14.1	Perda	0	
		55,799,155-56,130,257	12q13.3	Ganho	<i>LRP1, MIR1228, NXPH4, SHMT2, NDUFA4L2, STAC3, R3HDM2, INHBC</i>	hsa-mir-1228
		0-15,085,636	19p13.3-p13.12	Ganho	>10	>10
844T	844T_A, 844T_B	121,303,093-122,078,243	7q31.32	Perda	<i>PTPRZ1, AASS, FEZF1, CADPS2</i>	
845T	845T_A, 845T_B	193,404,142-195,096,722	1q31.3	Perda	<i>KCNT2, CFH, CFHR3, CFHR1</i>	
		79,347,620-84,788,605	2p12-p11.2	Perda	<i>CTNNA2, LRRTM1, LOC1720, FUNDC2P2, SUCLG1, DNAH6</i>	
		191,649,580-193,674,663	3q28	Perda	<i>TMEM207, IL1RAP, LOC647309, SNAR-I, OSTN, UTS2D, CCDC50, PYDC2, FGF12</i>	
		63,649,592-65,938,025	4q13.1	Perda	<i>TECRL, EPHA5</i>	
		39,883,239-40,351,828	7p14.1	Ganho	<i>CDK13, C7orf11, C7orf10</i>	
		318,929-892,124	12p13.33	Ganho	<i>KDM5A, CCDC77, B4GALNT3, NINJ2, WNK1, RAD52</i>	
		79,882,431-86,423,650	13q31.1	Perda	<i>SLITRK1, MIR548F1, SLITRK6</i>	
		96,180,194-98,956,993	14q32.2	Perda	<i>VRK1, C14orf64, C14orf177, BCL11B, SETD3</i>	
		49,705,956-51,033,907	16q12.1	Perda	<i>SALL1, TOX3</i>	
		84,729,051-85,325,767	16q24.1	Perda	<i>LOC732275, FOXF1, MTHFSD, FLJ30679, FOXC2, FOXL1</i>	
		88,247,077-88,410,341	16q24.3	Ganho	<i>CHMP1A, C16orf55, CDK10, SPATA2L, C16orf7, ZNF276, FANCA</i>	
		59,828,467-65,021,204	18q22.1-q22.2	Perda	<i>C18orf20, CDH7, CDH19, DSEL, TMX3, CCDC102B</i>	
		36,985,549-38,699,013	20q11.23-q12	Perda	<i>FAM83D, DHX35, LOC339568</i>	

Em negrito, alterações identificadas em mais de um caso.

Tabela 5. Regiões mínimas em comum significativas distribuídas entre as amostras de tumores múltiplos.

Casos	1p 31.1	1p 36.33- p36.31	2q 35	4q 13.1- q13.2	4q 28.3	5q 31.1- q31.2	6p 13.3	6p 21.32	7p 12.2	7q 22.1	8p 12- p11.23	8q 24.3	9p1 3.3	9q 34.3	11p 15.5	11p 14.1	11q 12.3- q13.2	12q 24.11- q24.12	12q 13.2- q13.3	13q 31.1	14q 32.2	15q 25.3	16p 11.2	16q 22.1	16q 24.3	17q 21.31	18q 22.1- q22.2	19p 13.12- 13.11	19q 13.32- q13.33	20q 11.23- 12	22q 13.33				
046TA																																			
046TB																																			
147TA																																			
147TB																																			
147TC																																			
300TA																																			
300TB																																			
301TA																																			
301TB																																			
301TC																																			
307TA																																			
307TB																																			
317TA																																			
317TB																																			
317TC																																			
629TA																																			
629TB																																			
630TA																																			
630TB																																			
631TA																																			
631TB																																			
631TC																																			
642TA																																			
642TB																																			
643TA																																			
643TB																																			
643TC																																			
683TA																																			
683TB																																			
722TA																																			
722TB																																			
722TC																																			
729TA																																			
729TB																																			
729TC																																			
733TA																																			
733TB																																			
844TA																																			
844TB																																			
845TA																																			
845TB																																			

As regiões em vermelho indicam a presença de alterações significativas (picos significativos com $P < 0.05$; presença $> 10\%$ dos casos; exclusão CNVs presentes $> 5\%$ população referência) entre as amostras de tumores múltiplos.

A



B



Figura 7. Representação esquemática das regiões identificadas entre diferentes amostras de tumores múltiplos. **A** – As amostras 845T_A e 845T_B apresentaram 13 regiões em comum (seleção). **B** – As amostras 683T_A e 683T_B não apresentaram regiões em comum (*Nexus v.5.0 -Biodiscovery*).

4.2.3 História familiar de Leiomiomas Uterinos

Neste estudo, seis pacientes relataram história de LU em parentes de primeiro e segundo grau (Anexo 2), totalizando dez amostras (046T_A, 046T_B, 084T, 629T_A, 629T_B, 633T, 683T_A, 683T_B, 842T e 857T). Para investigar a existência de um possível gene em comum associado ao risco familiar destes tumores, as amostras foram avaliadas individualmente quanto à presença de alterações frequentes entre os tumores, sem a utilização de critérios estridentes aplicados à análise global. Regiões de perdas em comum foram observadas em 1p36.33 nas amostras 046T_A, 084T, 629T_B e 633T e em 14q24.3 entre as amostras 046T_A, 629T_B e 842T (Figura 8A e B, respectivamente). Cada sonda gênica foi investigada para a descrição exata das perdas consecutivas em cada caso. Na região 1p36.33, apenas a sonda A_14_P131252 que mapeia o gene *PRKCZ* apresentou deleções entre os casos e na região 14q24.3, a sonda A_14_P200645 apresentou deleções entre os casos, mapeando o gene *VSX2* (Figura 9A e B, respectivamente). A sonda A_14_P200645, que mapeia o gene *VSX2*, está localizada no intron 3 deste gene (Figura 10B).

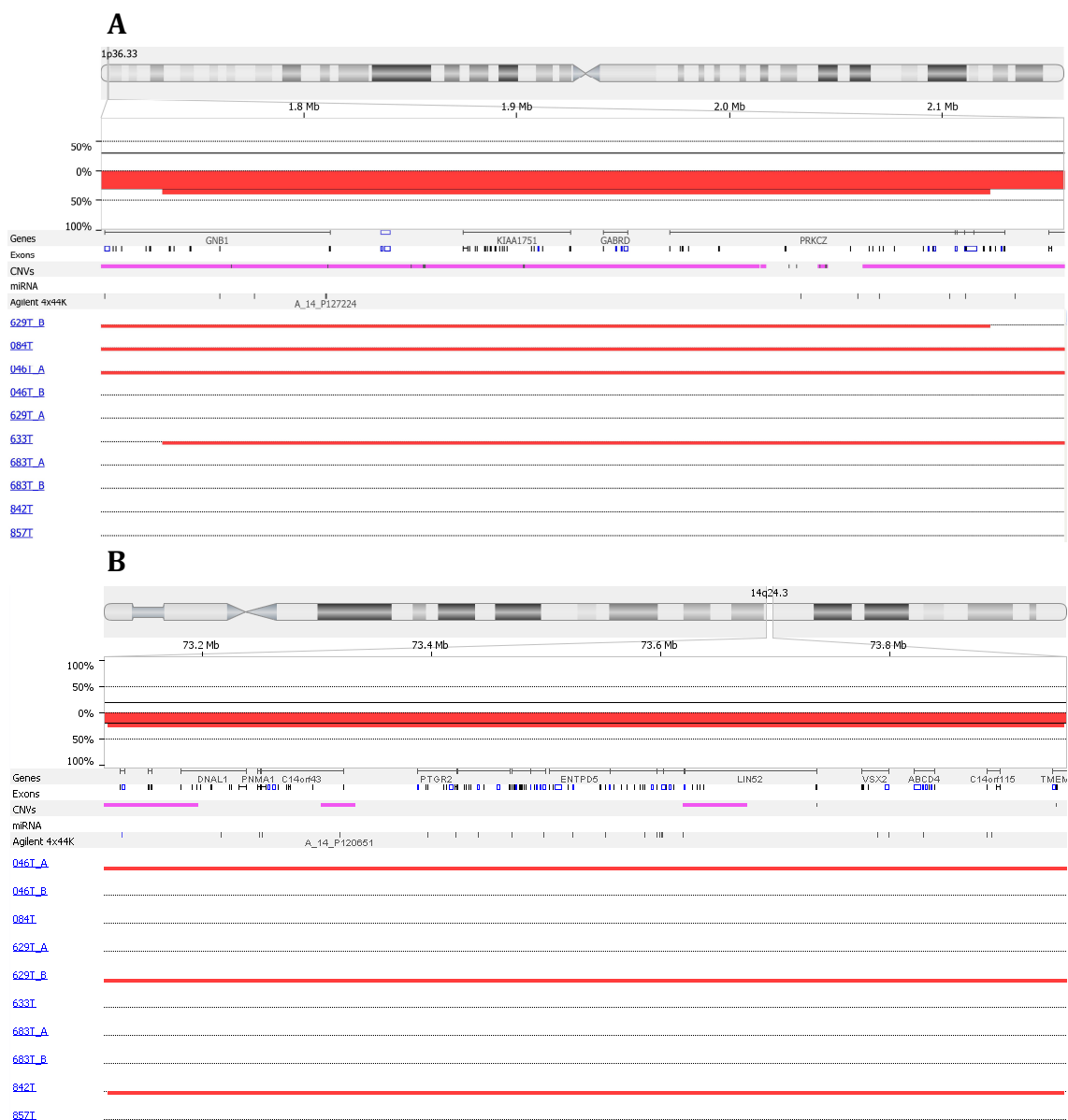


Figura 8 – Perdas genômicas identificadas nas amostras de pacientes com história de LU familiar. As linhas vermelhas ao centro da imagem representam as perdas entre os casos (em azul). **A** – Perda em 1p36.33 nas amostras 046T_A, 084T, 629T_B e 633T, onde estão mapeados genes como *GNB1*, *CALML6*, *TMEM52*, *KIAA1751*, *GABRD* e *PRKCZ*. **B** – Perda em 14q24.3 nas amostras 046T_A, 629T_B e 842T envolvendo vários genes, entre eles *DNAL1*, *PTGR2*, *V5X2* e *ABCD4* (Nexus v.5.0 -Biodiscovery).

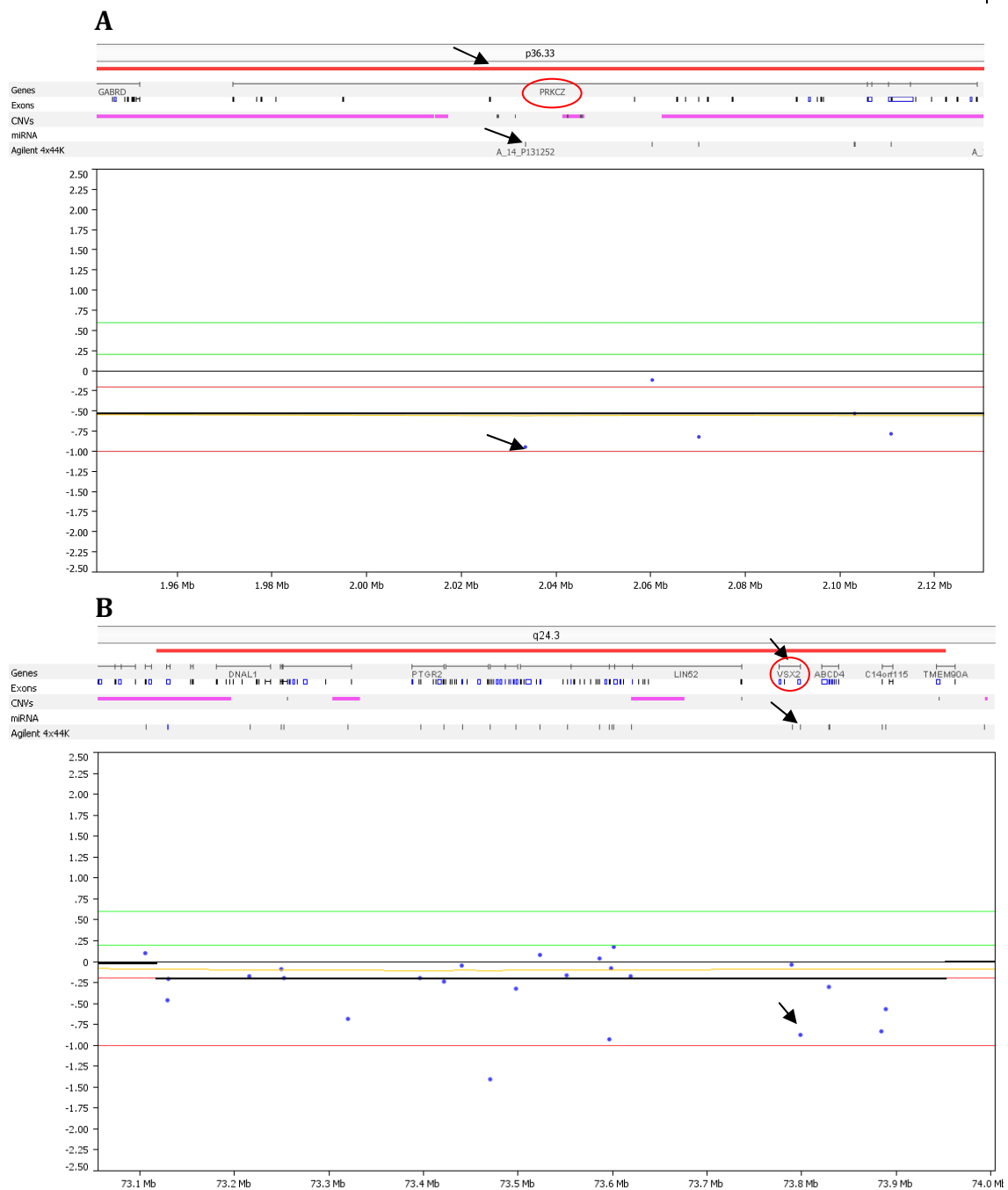
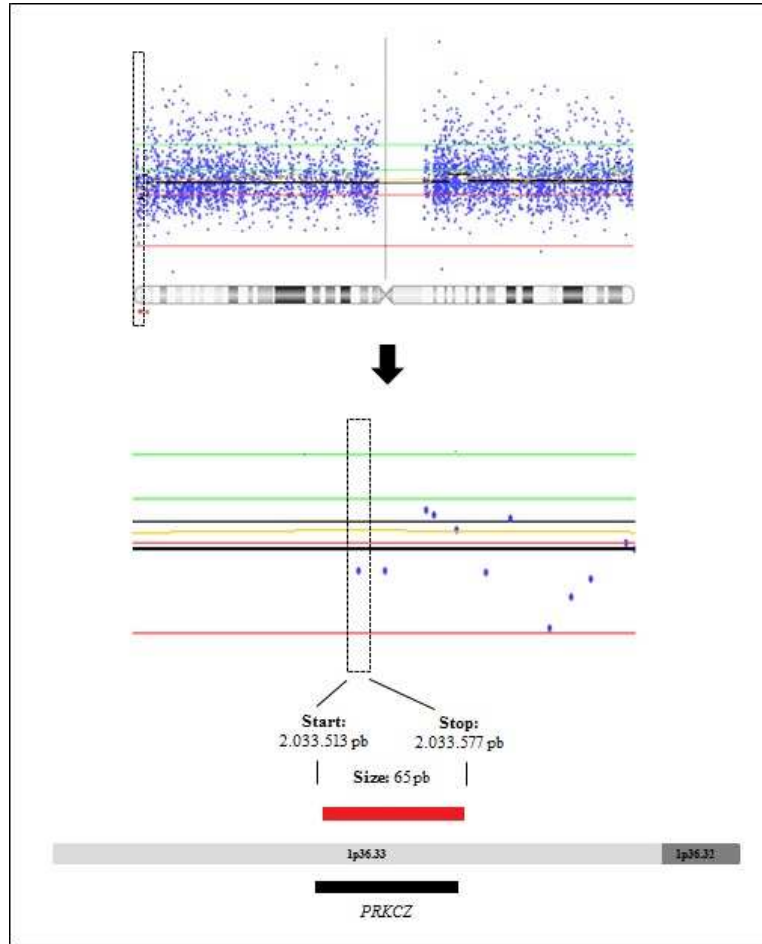


Figura 9. Representação esquemática da investigação das sondas gênicas nos casos que apresentaram as deleções mais frequentemente observadas. As flechas indicam os genes (círculo vermelho) e as sondas envolvidas. **A** - perda em 1p36.33 no caso 084T, mostrando a deleção da sonda A_14_P131252 que mapeia o gene *PRKCZ* envolvida em perda genômica nos casos 046T_A, 629T_B e 633T. **B**- perda em 14q24.3 no caso 842T, mostrando a sonda A_14_P200645 que mapeia o gene *VSX2*, envolvida em perda genômica comum nos casos 046T_A e 629T_B (*Nexus v.5.0 -Biodiscovery*).

A



B

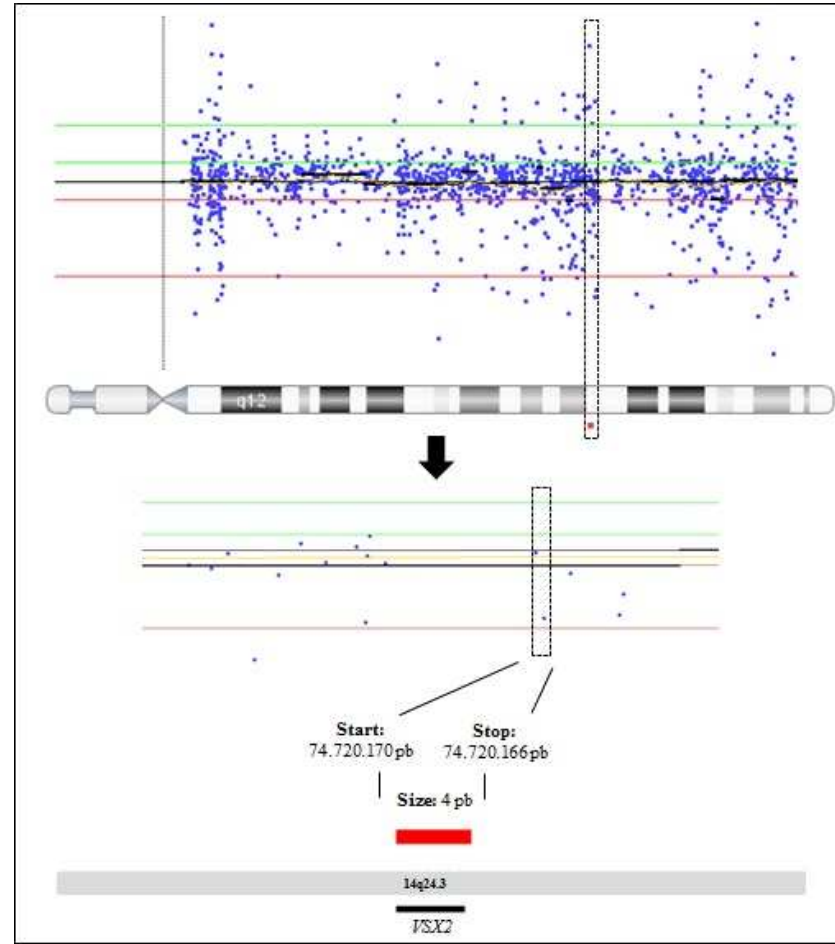


Figura 10. Representação esquemática da cobertura das sondas correspondentes aos genes candidatos envolvidos em perdas genômicas entre as amostras de LU familiar. **A** - perda em 1p36.33, mostrando a deleção da sonda A_14_P131252 que mapeia o gene *PRKCZ* com deleção de 65pb do íntron 4. **B**- perda em 14q24.3, mostrando a sonda A_14_P200645 que mapeia o gene *VSX2*, com deleção de 4pb do íntron 3 (*Nexus v.5.0 -Biodiscovery*).

4.2.4 Agrupamento hierárquico

Após a obtenção dos resultados das alterações genômicas gerais entre os 80 LU, foi aplicada a análise de agrupamento hierárquico. Foram observados dois grandes grupos de amostras, divididas entre o Grupo 1 ($n = 42$) e o Grupo 2 ($n = 38$) (Figura 11A). As CNVs que diferenciavam significativamente estes dois grupos foram os ganhos em 2q35 (10/42 casos), 7q22.1 (12/42 casos), 19p13.12-p13.11 e 19q13.32-q13.33 (10/42 casos), observados exclusivamente no Grupo 1 ($P < 0,01$) e apenas a perda em 7q22.1-q22.2 (10/38 casos), que foi observada exclusivamente no Grupo 2 ($P < 0,01$). A Figura 11B representa o perfil genômico das amostras agrupadas no Grupo 1 e Grupo 2. A Figura 12 mostra as alterações características do Grupo 1 e do Grupo 2, respectivamente. Entre as amostras do Grupo 1, foi possível observar uma maior frequência de tumores múltiplos em relação ao Grupo 2. No Grupo 1, pelo menos duas amostras dos casos 046T, 147T, 300T, 301T, 307T, 317T, 631T, 643T e 722T se agruparam, totalizando 20 amostras. No Grupo 2, foi observada uma maior frequência de amostras únicas, onde apenas as duas amostras do caso 844T se agruparam. Neste grupo, observamos também 11 amostras (317T_A, 629T_B, 630T_A, 631T_C, 642T_A, 643T_B, 683T_A, 722T_B, 729T_C, 733T_A, 845T_A) de tumores múltiplos isoladas de suas outras amostras, presentes no Grupo 1. As alterações dos Grupos 1 e 2 foram comparadas aos dados clínicos das pacientes, no qual foram estatisticamente associados com o número de amostras avaliadas: o Grupo 1 foi significativamente associado com tumores múltiplos (2 ou 3 amostras avaliadas - 28/42 amostras, $P < 0,01$) e o Grupo 2, com tumores não múltiplos (uma amostra avaliada - 25/39 amostras, $P < 0,01$) (Figura 13).

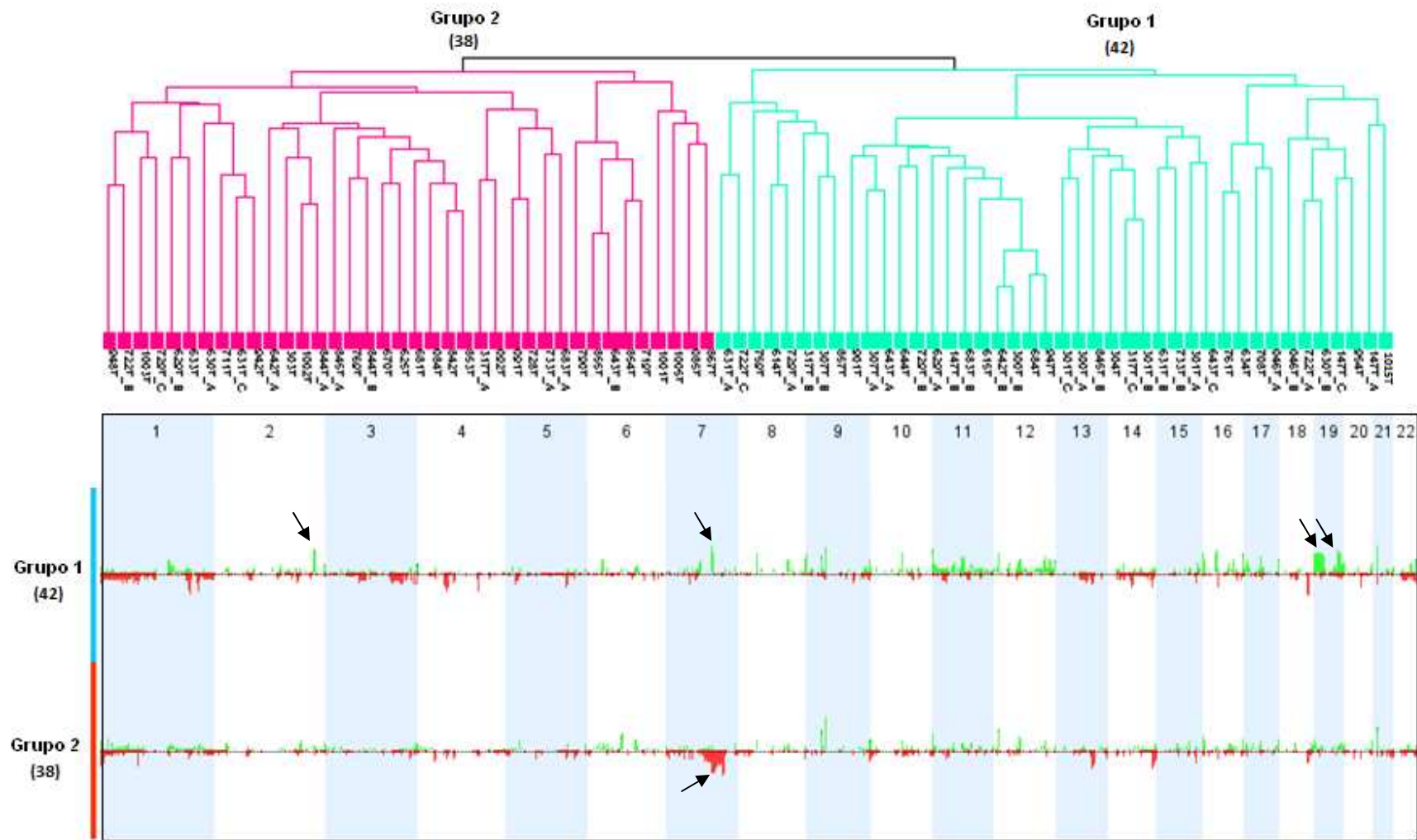
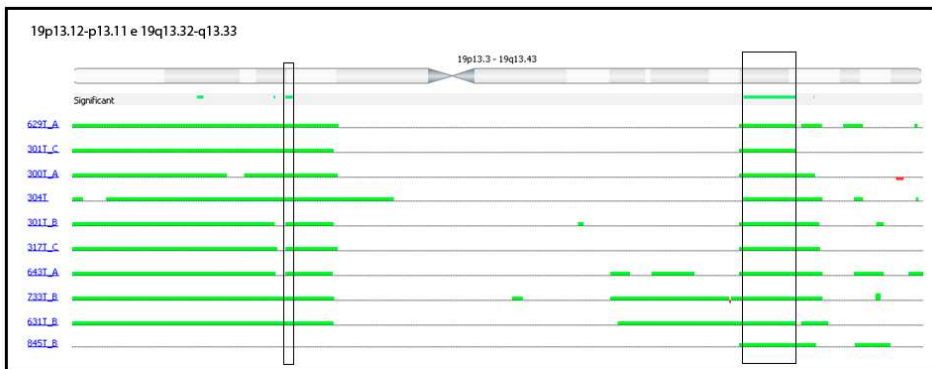
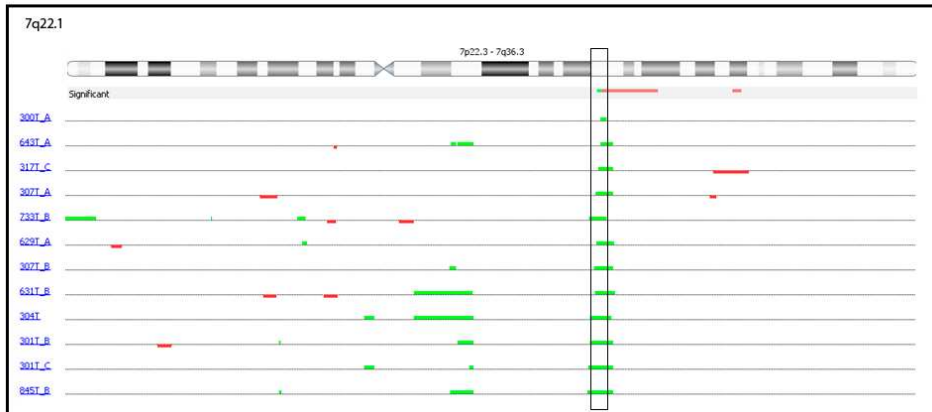
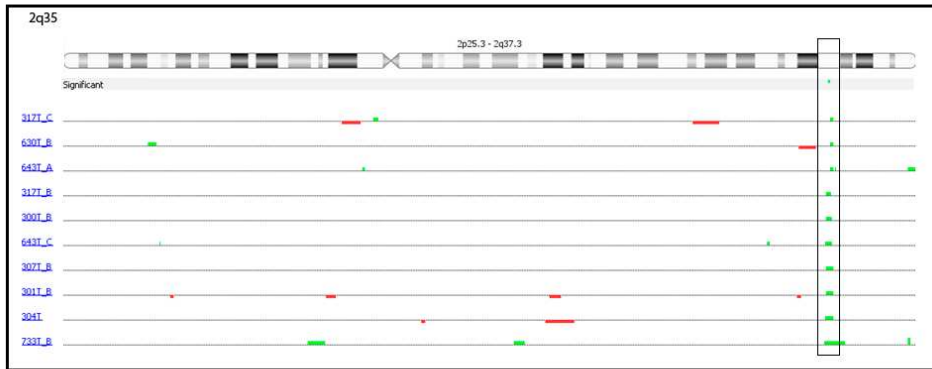


Figura 11. A - Análise de agrupamento hierárquico das CNVs identificadas gerou dois grandes grupos de amostras: Grupo 1 (n=42) e Grupo 2 (n=38). **B** - Ganhos em 2q35, 7q22.1, 19p13.12- 13.11 e 19q13.32-q13.33 foram observados exclusivamente no Grupo 1 ($P < 0,01$). Perdas em 7q22.1-q22.2 foram presentes exclusivamente no Grupo 2 ($P < 0,01$) (setas) (Nexus v.5.0 -Biodiscovery).

Grupo 1



Grupo 2

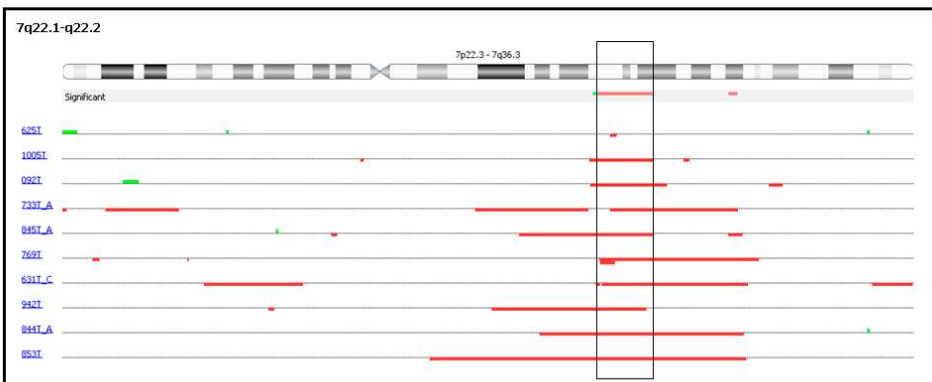


Figura 12. Regiões de ganhos e perdas exclusivas aos grupos identificados. **Grupo 1:** amostras que apresentaram ganhos em 2q35, 7q22.1, 2q35, 7q22.1, 19p13.12-p13.11 e 19q13.32-q13.33. **Grupo 2:** amostras com perda em 7q22.1-q22.2 (Nexus v.5.0 -Biodiscovery).

2 ou 3 amostras avaliadas
 1 amostra avaliada

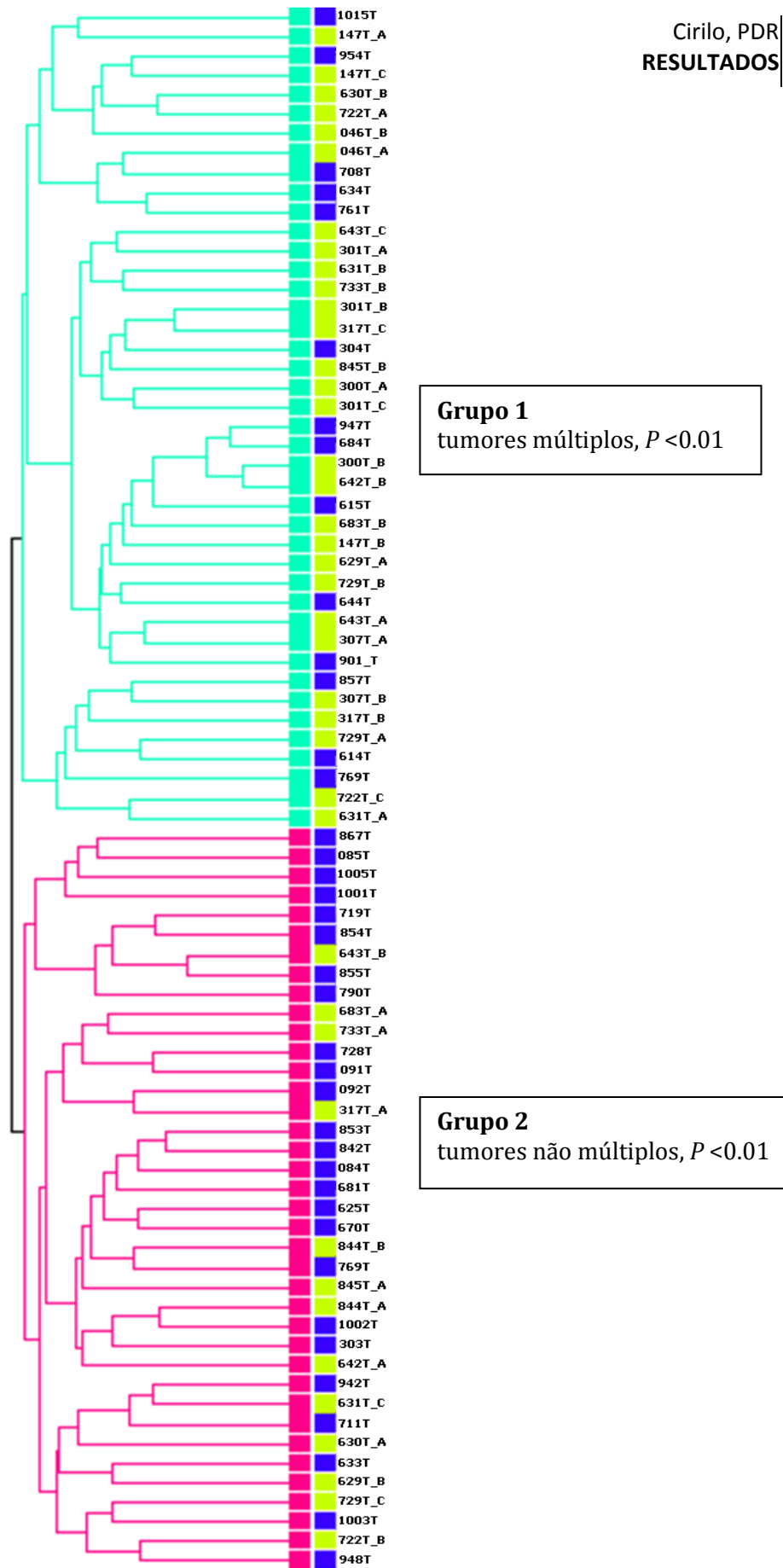


Figura 13. Grupos 1 e 2 estatisticamente associados com tumores múltiplos e tumores não múltiplos, respectivamente (*Nexus v.5.0 - Biodiscovery*).

4.2.5 Investigação por FISH

Como a deleção em 7q22.1-q22.2 foi exclusivamente observada em um subgrupo de amostras, esta região foi selecionada para um estudo mais detalhado. A sonda mapeada em 7q22.3-q31.1 (RP11-46J20) foi utilizada para FISH em um estudo prévio do grupo (Perez, 2004) em um subgrupo de LU os quais também foram avaliados por CGH array neste estudo. Preparações cromossômicas normais foram usadas como controle do experimento sendo detectado um padrão normal em 94,3% das células (Figura 14). Onze dos 27 casos (41%) de LU avaliados pela FISH apresentaram a perda em 7q22.3-q31.1 (711T, 729T_A, 759T, 769T, 801T, 842T, 844T_A, 845T_A, 853T, 854T e 947T). Para comparação com os achados de FISH, aos dados de CGH array foram consideradas apenas as regiões descritas com os critérios usados na identificação de CNVs e com exclusão de CNVs em >5% da população controle referentes a perdas em 7q. A Tabela 6 mostra todas as perdas identificadas em 7q por CGH array e os casos de FISH que foram concordantes com este achado. Foram identificadas 29 regiões envolvidas em perdas em 7q, distribuídas entre 49 amostras. A deleção em 7q22.3-q31.1 (região mínima de cobertura da sonda de FISH) evidenciada por CGH array, foi confirmada em quatro casos (769T, 844T_A, 845T_A e 853T). Outras regiões de perdas foram observadas em maior frequência, como 7q22.1 observada em 5 casos (625T, 631T_C, 728T, 769T e 842T) e a perda em 7q31.33 identificada também em 5 casos (722T_B, 729T_C, 845T_A, 948T e 1003T), porém sem concordância com a cobertura da sonda de FISH.

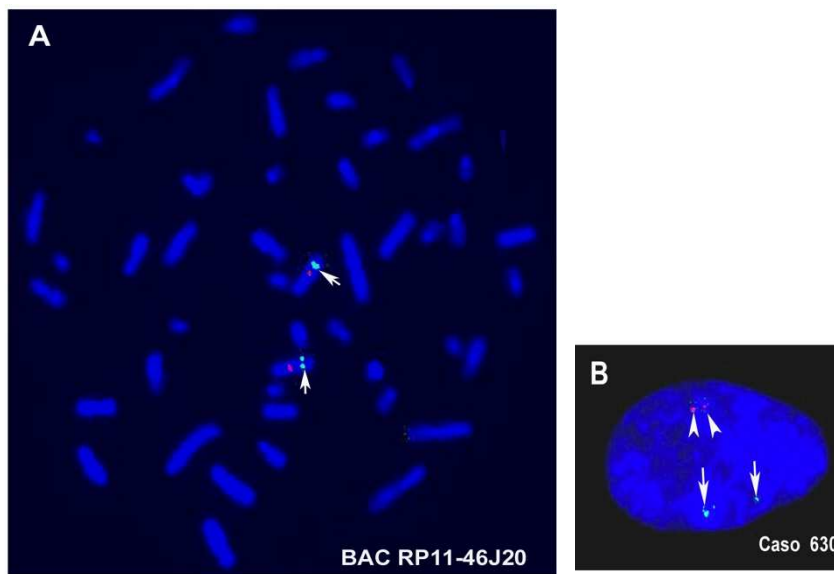


Figura 14. Análise citogenética para a sonda RP11-46J20 da região 7q22.3-q31.1. **A**- Metáfase obtida após cultura temporária de linfócitos de sangue periférico, confirmando o mapeamento da sonda em 7q (setas). **B** - Exemplo de núcleo interfásico de LU mostrando dois sinais da sonda teste (verde) e dois sinais do controle centrômero do cromossomo 7 (Dados obtidos de Perez, 2004).

Tabela 6. Deleções no braço longo do cromossomo 7 identificadas por CGH array e “validação” por FISH entre as amostras de Leiomiomas Uterinos.

Região	Amostras	Start	End	Símbolo gene	miRNAs	FISH
q11.21	317T_B	63,375,270	65,122,221	>10		
	733T_B	62,153,588	64,985,140	>10		
q11.21-q11.22	1015T	62,153,588	70,745,172	>10		
q11.21-q11.23	759T	64,178,574	76,030,684	>10	hsa-mir-590	
q11.22-q11.23	630T_A	71,399,429	73,199,550	>10		
q11.22-q32.1	853T	68,428,492	127,561,721	>10	hsa-mir-591, hsa-mir-1285-1, hsa-mir-548o, hsa-mir-590, hsa-mir-489, hsa-mir-106b, hsa-mir-25, hsa-mir-653, hsa-mir-592, hsa-mir-593, hsa-mir-93	DEL
q11.23-q21.3	733T_A	76,908,841	97,392,304	>10	hsa-mir-591, hsa-mir-1285-1, hsa-mir-489, hsa-mir-653	
q21.11-q21.12	854T	84,830,349	86,551,075	<i>GRM3, KIAA1324L</i>		
q21.11-q31.1	845T_A	85,102,183	110,503,042	>10	hsa-mir-591, hsa-mir-1285-1, hsa-mir-548o, hsa-mir-489, hsa-mir-106b, hsa-mir-25, hsa-mir-653, hsa-mir-93	DEL
	942T	80,135,755	108,866,979	>10	hsa-mir-591, hsa-mir-1285-1, hsa-mir-548o, hsa-mir-489, hsa-mir-106b, hsa-mir-25, hsa-mir-653, hsa-mir-93	
q21.13-q32.1	844T_A	88,958,429	127,101,490	>10	hsa-mir-591, hsa-mir-1285-1, hsa-mir-548o, hsa-mir-489, hsa-mir-106b, hsa-mir-25, hsa-mir-653, hsa-mir-592, hsa-mir-93	DEL
q22.1	625T	102,068,500	103,385,539	>10		

	631T_C	99,551,867	100,318,043	>10	
	728T	99,953,849	102,068,500	>10	hsa-mir-548o
	769T	100,114,185	103,058,551	>10	hsa-mir-548o
	842T	100,235,958	100,590,565	<i>ACHE, AP1S1, EPHB4, MUC12, MUC17, SERPINE1, SLC12A9, SRRT, TRIM56, TRIP6, UFSP1</i>	
q22.1-q22.2	759T	104,284,386	105,139,642	<i>ATXN7L1, EFCAB10, LHFPL3, LOC100216545, LOC723809, MLL5, PUS7, RINT1, SRPK2</i>	
	855T	103,958,095	105,139,642	<i>ATXN7L1, EFCAB10, LHFPL3, LOC100216545, LOC723809, MLL5, PUS7, RINT1, SRPK2</i>	
q22.1-q31.1	092T	98,368,589	112,721,449	>10	hsa-mir-548o, hsa-mir-106b, hsa-mir-25, hsa-mir-93
	1005T	98,191,548	110,352,000	>10	hsa-mir-548o, hsa-mir-106b, hsa-mir-25, hsa-mir-93
q22.1-q31.33	733T_A	102,068,500	125,951,548	>10	
q22.1-q32.1	631T_C	100,633,129	127,907,619	>10	hsa-mir-548o, hsa-mir-129-1, hsa-mir-592, hsa-mir-593
q22.1-q32.2	769T	103,058,551	129,940,385	>10	hsa-mir-335, hsa-mir-96, hsa-mir-183, hsa-mir-182, hsa-mir-129-1, hsa-mir-592, hsa-mir-593
q31.1	1002T	110,503,042	113,342,324	<i>C7orf53, C7orf60, DOCK4, GPR85, IFRD1, IMMP2L, LOC401397, LRRN3, PPP1R3A, TMEM168, ZNF277</i>	
q31.1-q31.2	1015T	112,571,792	117,037,157	>10	
q31.2	317T_A	115,662,896	116,947,755	<i>ASZ1, CAPZA2, CAV1, CAV2, CFTR, MET, ST7, ST7OT1, ST7OT2, ST7OT3, ST7OT4, TES, WNT2</i>	
	1005T	115,854,745	116,991,644	<i>ASZ1, CAPZA2, CAV1, CAV2, CFTR, MET, ST7, ST7OT1, ST7OT2, ST7OT3, ST7OT4, WNT2</i>	

q31.31-q31.32	307T_A	120,102,164	121,375,364	<i>C7orf58, FAM3C, ING3, KCND2, PTPRZ1, TSPAN12, WNT16</i>	
q31.31-q32.1	317T_C	120,721,032	127,440,289	>10	hsa-mir-592
q31.32	844T_B	121,303,093	122,078,243	<i>AASS, CADPS2, FEZF1, PTPRZ1</i>	
q31.32-q31.33	1002T	122,129,122	124,352,498	>10	
q31.32-q32.1	1015T	122,005,948	127,035,525	>10	hsa-mir-592
q31.33	722T_B	124,268,700	125,951,548	<i>GRM8, POT1</i>	
	729T_C	124,268,700	126,559,231	<i>GRM8, MIR592, POT1</i>	hsa-mir-592
	1003T	124,268,700	126,559,231	<i>GRM8, MIR592, POT1</i>	hsa-mir-592
	845T_A	124,234,571	126,759,965	<i>GRM8, MIR592, POT1</i>	hsa-mir-592
	948T	124,268,700	126,322,883	<i>GRM8, POT1</i>	
q32.3-q33	092T	131,709,708	134,399,994	<i>AGBL3, AKR1B1, AKR1B10, AKR1B15, BPGM, CALD1, CHCHD3, EXOC4, LRGUK, PLXNA4, SLC35B4</i>	
q34	759T	137,414,239	140,802,424	>10	
q36.1	085T	150,095,956	150,684,706	>10	hsa-mir-671
	633T	150,306,435	151,042,646	>10	hsa-mir-671
	759T	151,435,276	152,068,395	<i>FABP5L3, GALNT11, LOC100128822, MLL3, XRCC2</i>	
	855T	151,435,276	152,088,370	<i>ACTR3B, FABP5L3, GALNT11, LOC100128822, MLL3, XRCC2</i>	
q36.1-q36.3	631T_C	151,042,646	158,475,824	>10	hsa-mir-153-2, hsa-mir-595
q36.3	722T_B	155,187,858	156,434,124	<i>C7orf13, LMBR1, NCRNA00244, RBM33, RNF32, SHH</i>	
	854T	156,682,247	157,503,156	<i>DNAJB6, MIR153-2, PTPRN2, UBE3C</i>	hsa-mir-153-2

DEL=deleção da região mínima comum de cobertura da sonda de FISH (7q22.3-q31.1).

4.3 Análise funcional dos genes mapeados em regiões significativas

Os processos biológicos dos 1323 genes mapeados nas regiões de CNVs significativas foram obtidos utilizando o programa *Nexus (Biodiscovery)*. Foram identificadas 85 categorias moleculares, entre as quais, cinco principais processos biológicos que apresentaram genes associados com a regulação da transcrição (15,38%), *splicing* de RNA (5,53%), processos metabólicos de lipídeos (4,81%), regulação da transcrição da RNA polimerase II (4,33%) e processos metabólicos de lipídeos (2,88%) (Figura 15).

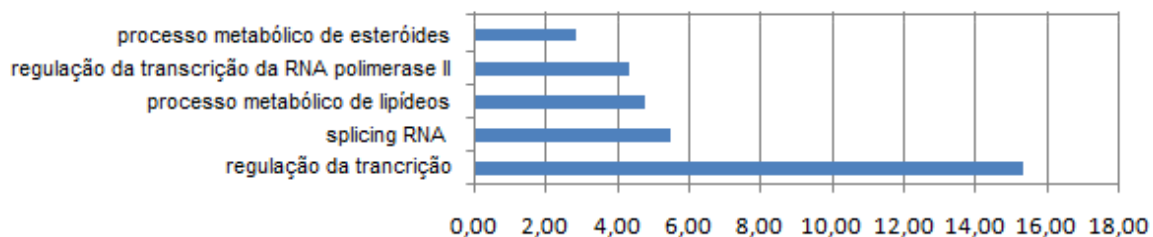


Figura 15. Os 1323 genes alterados quanto ao número de cópias nos LU estavam associados a várias funções, incluindo processo metabólico de esteróides, regulação da transcrição da RNA polimerase II, processo metabólico de lipídeos, mecanismos de *splicing* alternativo e principalmente genes associados à regulação da transcrição. Os valores indicados estão em porcentagem (*Nexus v.5.0 -Biodiscovery*).

Para confirmação das funções biológicas destes genes, análises funcionais foram geradas através do sistema *Ingenuity Pathway Analysis* (IPA). Os resultados obtidos mostraram 25 redes gênicas com valores de *scores* variando de 14 a 1. Valores de *scores* mais altos indicam maior associação entre os genes importados com as moléculas do banco de dados IPA. A Tabela 7 mostra as funções biológicas, doenças e funções moleculares, associadas com as moléculas em cada rede. A rede com *score* mais alto (14) associou 18 moléculas elegíveis á partir dos dados obtidos com 16 outras moléculas do IPA. As funções significativas associadas a estas moléculas foram expressão gênica, doença inflamatória e sinalização celular. As doenças e funções moleculares associadas a estas moléculas estão representadas na Figura 16A e 16B, respectivamente. As doenças associadas com as moléculas que apresentaram os cinco maiores valores de $-\log(P \text{ valor})$ foram câncer (2,8), doenças do sistema reprodutivo (2,8), doenças genéticas (2,2), doenças musculares e esqueléticas (2,2) e desenvolvimento e função do tecido conectivo (1,4). Da mesma forma, as cinco funções moleculares com os mais altos valores de $-\log(P \text{ valor})$ foram metabolismo de drogas (3,9), transporte molecular (3,9), bioquímica de pequenas moléculas (3,9), modificação pós-transdução (2,8) e proliferação e crescimento celular (2,7).

Tabela 7. Redes gênicas com os 10 *scores* mais altos, geradas a partir dos genes mapeados nas regiões de CNVs significativas geradas pelo IPA - *Ingenuity Pathway Analysis*.

Rede	Moléculas da rede*	Score	Moléculas em foco	Principais funções
1	<i>Actin, BACE1, ↑BRF2, ↑CD19, ↑CDC42BPG, DMAP1, ↑ETV4, ↑FANCA, ↑IFITM1, Lamin b, MSN, ↑PLEC, POU2F1, ↑PRMT1, ↑PRPF6, PRPF4B, RN7SK, ↑RRAS, ↑RTN3, ↑RUVBL2, ↑SCT, SMARCA4, ↑SMARCC2 (includes EG:6601), SNAPC1, SNAPC2, ↓SNAPC4, SOCS1, SP1, SP3, ↑SRCAP, STAT1, ↑STAT6, SWI-SNF, TNFRSF1A, ↑TRADD</i>	14	18	Expressão Gênica, Doença Inflamatória, Sinalização Celular
2	<i>14-3-3, A2M, ↑AKT1S1, ↑APOE, APP, ↑ATF5, ↑BAD, ↑BCAT2, ↑BRCA1, CCND1, CDC34 (includes EG:997), CDKN1B, ↑GLI1, ↑GSTP1, Hsp90, HSPA8, ↑IFI35, JUN, KRT19, ↑LCAT, LOC10431, ↑LRP1, MAP3K5 (includes EG:4217), ↑MAPK11, MBD1, MBD2, MBD3 (includes EG:53615), MECP2, NMI, ↑PPP1CA, ↑PPP5C, ↑SH2B3, SHC1, ↑TOMM40, ↑TUFM</i>	13	17	Expressão Gênica, Morte celular, Replicação de DNA, Recombinação e Reparo
3	<i>Alpha tubulin, ↑AP2A1, ↑ASH2L, ↑C16ORF53, ↑C5AR1, ↑CARD8, CHD8, DPY30, ↑GPA1, ↑GTPBP4, HCFC1, ↑HIF3A, IFN Beta, ↓IKZF1, ↑IRF3, ↑IRF7, MEN1, ↑MEOX1, MLL2, MLL3, MLL4, NFκB (complex), POLR2B, PRKCB, ↑PXN, SATB1, ↑SETD1A, ↑TALDO1, ↑TCEB2, TES, ↑VASP, VCL, VHL, WDR5, ZNF143</i>	13	17	Desenvolvimento Embrionário, Desenvolvimento Tecidual, Expressão Gênica
4	<i>B2M, CANX, CCL19, ↑CXCR2, ERBB2, ↑ERBB3, ESR1, ↑FCGRT, ↓GRB10, HDAC1, ↑HSF1, Hsp90, IL8, ↑IL27, IL12A, IL1B, IL6ST, IRF1, IRF4, JAK1, ↑MAPK3, NFKB1, NFKB2, ↑PLA2G16, ↑PTGES3 (includes EG:10728), RAF1, ↑RARRES3, REL, RELA, ↑RELB, ↑SIRT1, ↑SPIB, ↑STIP1, ↑SYMPK, TP53</i>	11	15	Desenvolvimento Celular, Doença Imunológica, Expressão Gênica
5	<i>↓AJAP1 (includes EG:55966), CCNH, ↑CD151, CDK7, CEBPB, ↑CTCF, CTNNB1, ↑DDIT3, DDX5, ERBB2, ↑ERCC2, ERCC3, ESR1, GTF2H1, GTF2H2, HNRNPA1, ↑HRAS, ITGA6, ITGB1, LGALS3, MAX, MDM2, MNAT1, MYC, NBN, NCL, ↑PA2G4, ↑PUF60, ↑RPL8, RPL11, ↑RPL18, ↑RPL27, ↑RPLP2, ↑SLC3A2, ↑SPN</i>	9	14	Expressão Gênica, Replicação de DNA, Recombinação e Reparo, Ciclo Celular
6	<i>Ap2 alpha, ↑BAX, ↑BBC3, BCL2, BHLHE40, CCNA2, CCNB1, CDC25A, CDK1, CDKN1A, ↓CENPF, ↑CLN3, CLU, ↑DYNLL1, EP300, ↑ERCC1, FEN1, FOXO1, ↑GDF11, HDAC1, HDAC3, ↑MMP25, MYC, PML, ↑PPP1R13L, ↑PRIM1, PRKCD, ↑RAD9A, RELA, SFN, SIN3A, ↓SKI, TOPBP1, TP53, ↓TP73</i>	8	13	Ciclo Celular, Morte Celular, Replicação de DNA, Recombinação e Reparo, Ciclo Celular
7	<i>↑BAX, BCL2, ↑BRCA1, CCNB1, CCND1, CDK1, ↑CDK2, CDK8, CDKN1B, ↑CPT1B, Cyclin A, CYP1B1, ↑E2F4, ↑ESRRA, GTF2B, ↑LIG1, MED1, MED21, PCNA, PIN1, ↑PNKP, POLB, ↑POLD1, POLR2D, ↑POLR2G, POLR2I, ↑POLR2L (includes EG:5441), PPARGC1A, RBL1, RBL2, ↑RDBP, RNA polymerase II, SKP2, TFF1, TFIIF</i>	7	12	Câncer, Ciclo Celular, Expressão Gênica

8	<p>↑ACD, ↑BCL2L12, ↑BRCA1, CEBPA, ↑CPSF1, CREBBP, EP300, ESR1, HDAC1, HDAC2, HEXIM1, Histone h4, Hsp70, ↑ITGAL, ↑MAPK12, MDM2, NBN, ↑NXF1, ↑ODF3B, ↑PGLYRP1, ↓POT1, RAD50, ↓RINT1, RUNX1, ↑SHANK3, TAF5, TAF7, TAF11, TAF12 (includes EG:6883), TERF1, TERF2, TERF2IP, TINF2, TNPO2, WRN (includes EG:7486)</p>	7	12	Organização e Montagem Celular, Função e Manutenção Celular, Replicação de DNA, Recombinação e Reparo
9	<p><i>AR, ARHGDI1A, ↑ATP2A2, ↑CEND1, CRSP5, ↑DCTPP1, ERBB2, ESR1, ↑FOXH1, GRB2, IL1R1, INPP5D, ↑LAT, MED1, MED14, MED24, ↑MED25, MED27, MED7 (includes EG:9443), ↑MYL2, MYOCD, ↓NTRK3, PI3K (complex), ↑PITPNM1, PLC gamma, ↓PRKCZ, PTK2, PTK2B, RHOA, RNA polymerase II, SHC1, SRC, ↑STAT2, TFF1, ↑TGFB111</i></p>	7	12	Expressão Gênica, Sinalização Celular, Morfologia Celular
10	<p><i>CBX1, CBX5, DDX20, EHMT1, ↑EHMT2, ESR1, FBL, GEMIN4, GEMIN5, GEMIN6, ↑GEMIN7, ↑GLTSCR2, Histone h3, ↑MVP, NCAPD3, NCAPG2, ↑NCAPH2, ↑NFATC2IP, PTEN, ↑RNF41, RPS6KB1, ↑RPS6KB2, SIP1, SMC2, SMC4, SMN1, SNRPD1, ↑SNRPD2, SNRPD3, SNRPE, SNRPF, ↑STAR, TERT, TP53, ↓VRK1</i></p>	6	11	Modificação pós-transcricional de RNA, Organização e Montagem Celular, Ciclo Celular

*Setas para cima (vermelhas) indicam genes mapeados em regiões de ganhos genômicos e setas para baixo (verdes) indicam genes mapeados em regiões de perdas genômicas. Em negrito, moléculas identificadas no estudo atual.

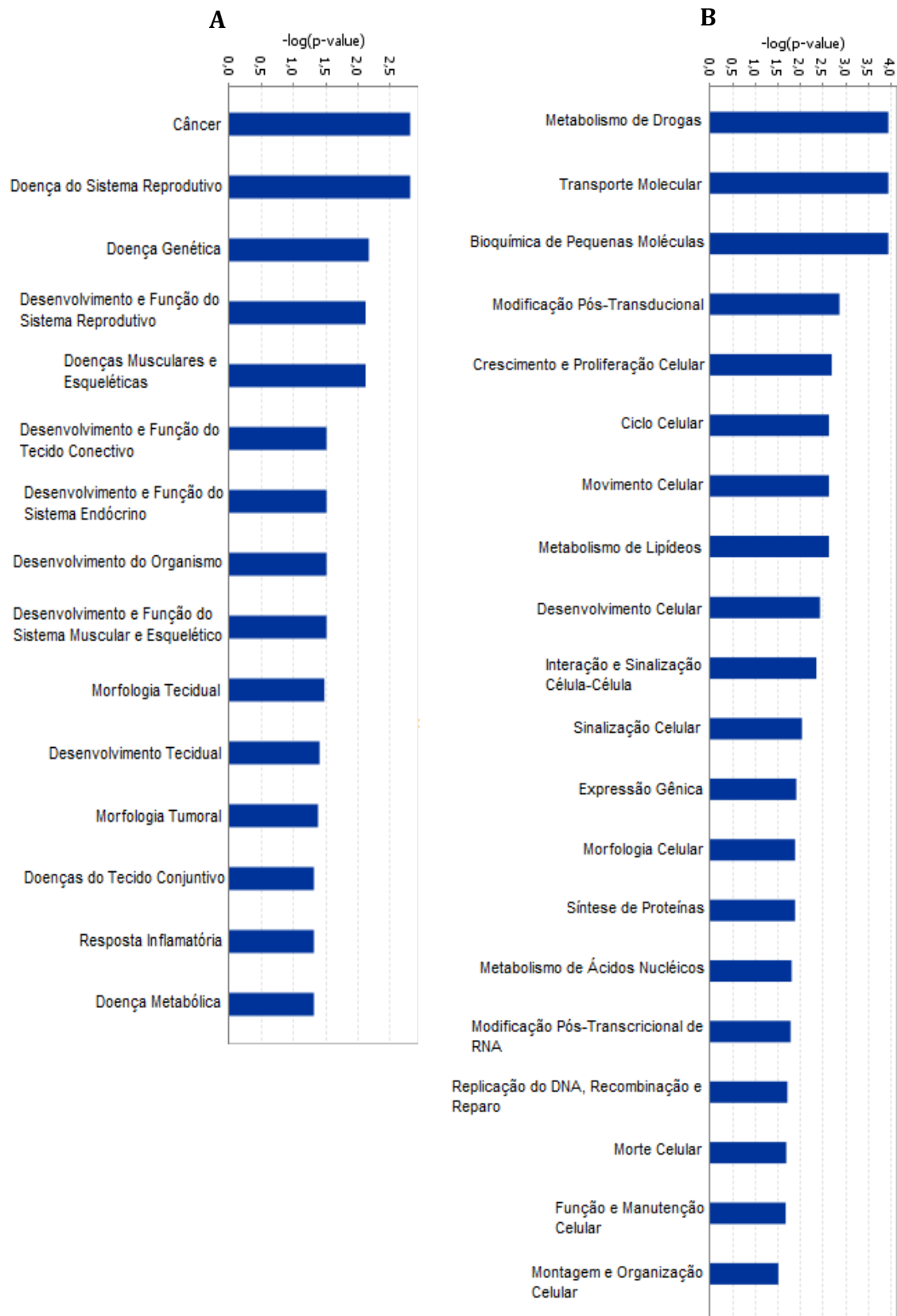


Figura 16. Funções biológicas associadas com as moléculas identificadas nas redes a partir dos genes mapeados em regiões de CNVs significativas. Os valores são denotados em $-\log(P\text{ value})$. **A.** Doenças; **B.** Funções Moleculares (IPA).

As vias canônicas geradas a partir da razão entre os genes deste estudo e as moléculas depositadas no banco de dados IPA revelaram as cinco principais vias associadas com, metabolismo dos hormônios esteróides (72%), sinalização de quimiocinas (69%), reparo a quebras de dupla-fita de DNA por recombinação homóloga (65%), via de reparo por excisão de nucleotídeos (55%) e sinalização ATM (53%) (Figura 17).

Outras vias de interesse entre os estudos dos LU também foram identificadas como Sinalização de IGF-1 (50%), TGF-beta (35%) e Sinalização do Receptor de Estrógeno (29%) (Figuras 18, 19 e 20, respectivamente).

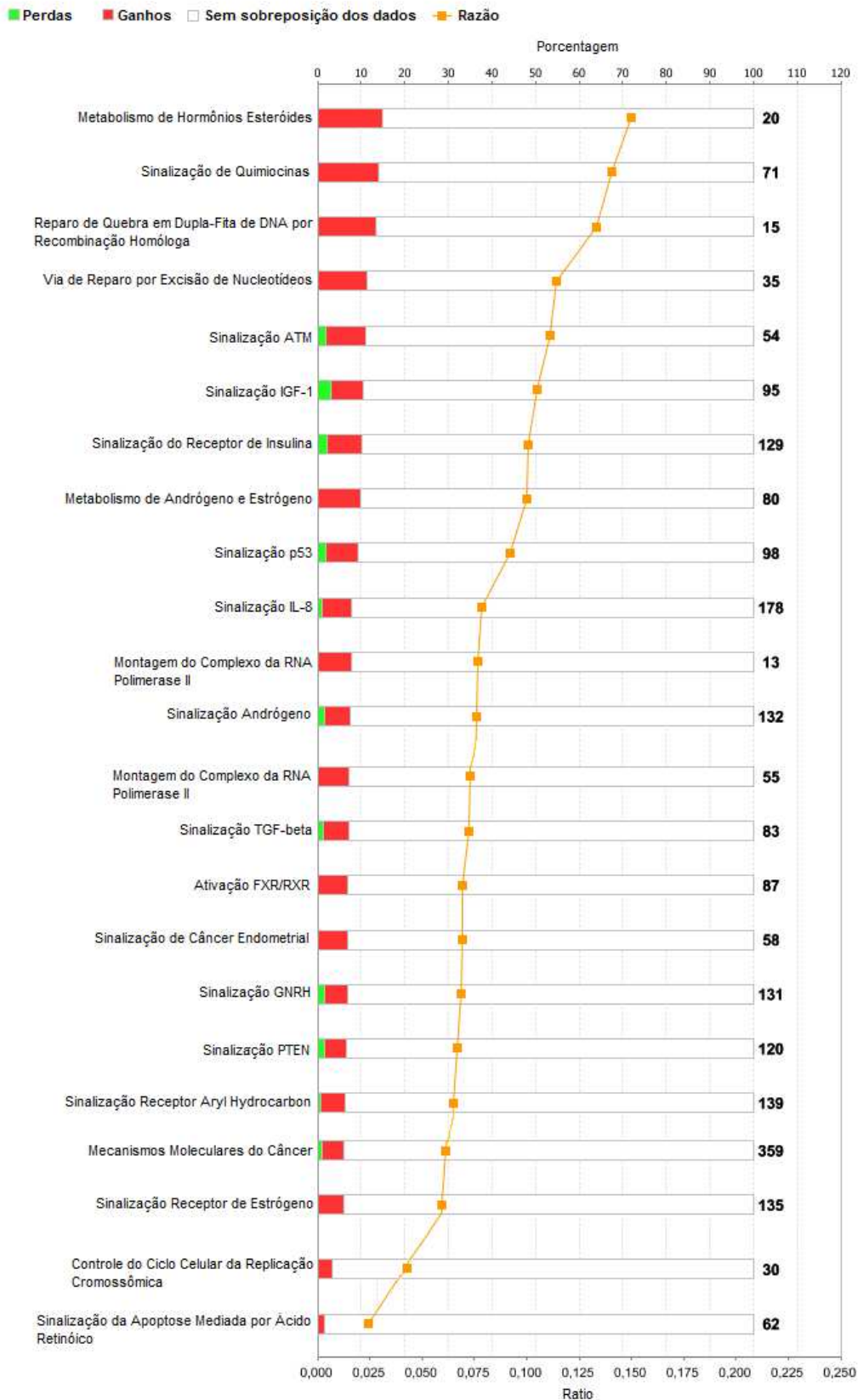


Figura 17. Vias canônicas geradas a partir da razão entre os genes mapeados em regiões de CNVs significativas e as vias depositadas no banco de dados IPA (IPA).

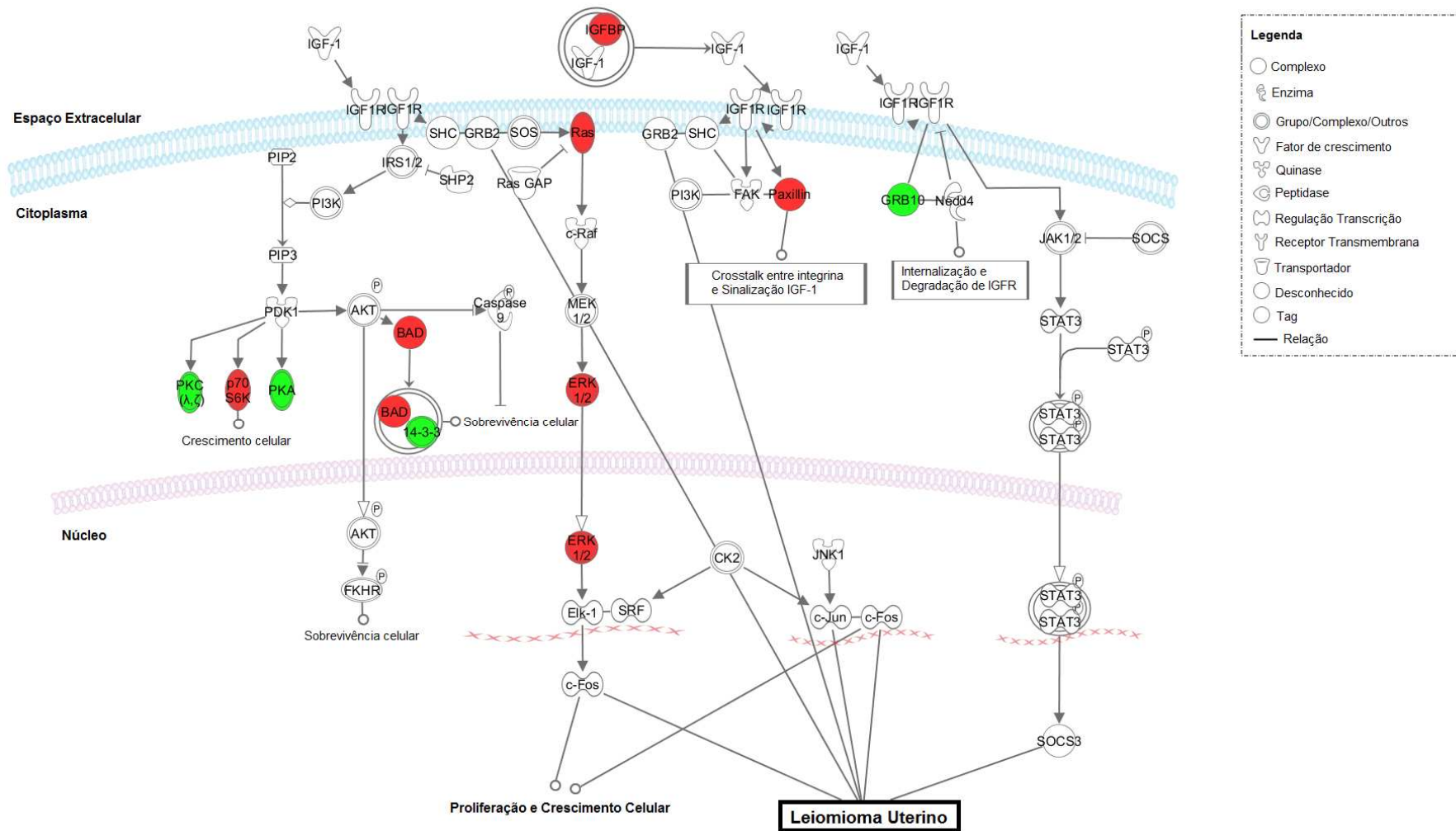


Figura 18. Representação gráfica da via de Sinalização IGF-1. Esta via têm sido associada aos mecanismos de proliferação celular em Leiomomas Uterinos. Em verde, moléculas envolvidas em perdas genômicas e em vermelho, moléculas envolvidas em ganhos (IPA).

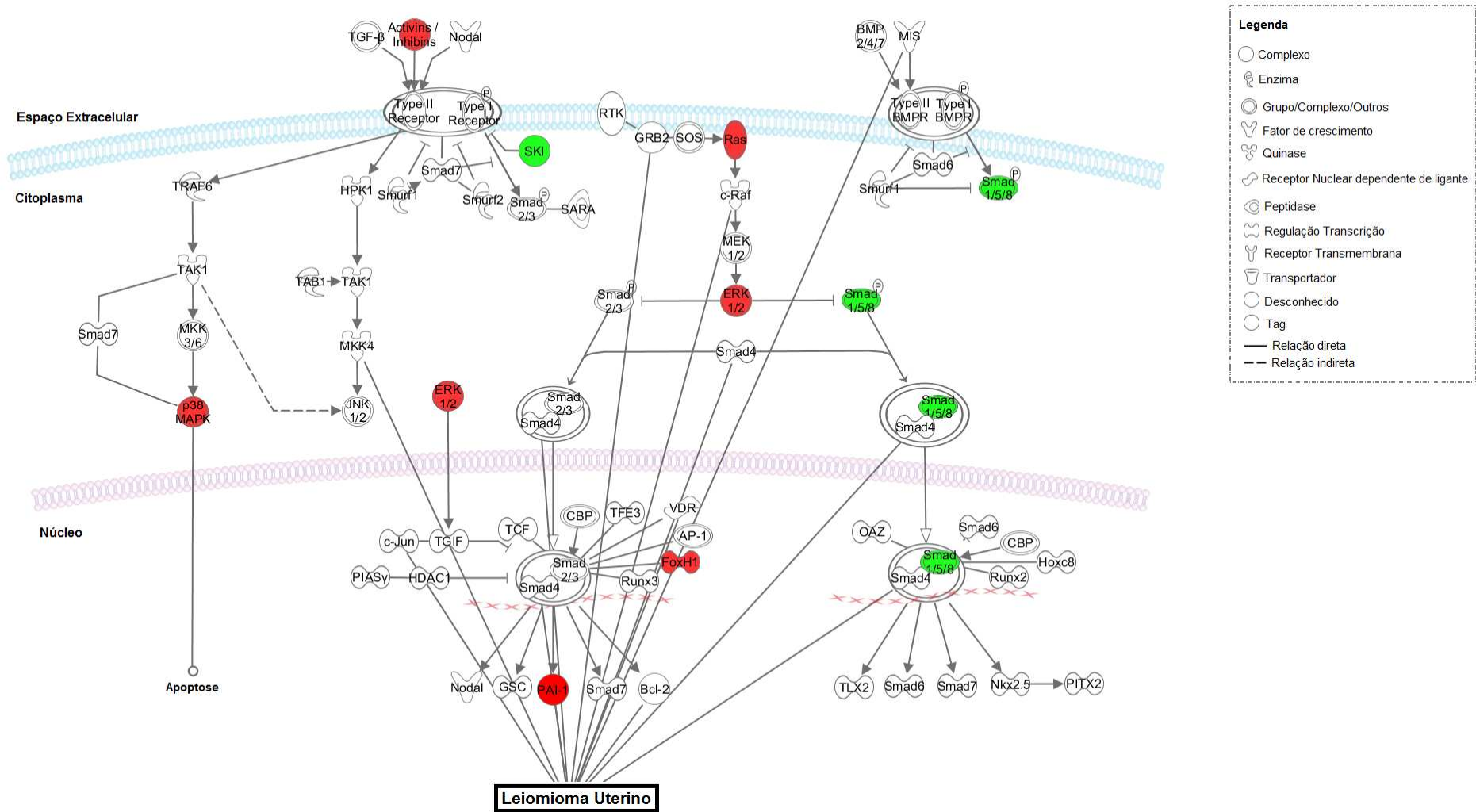


Figura 19. Representação gráfica da via de Sinalização TGF-beta. Estudos prévios associaram genes desta via com a característica fibróide dos Leiomiomas Uterinos. Em verde, moléculas envolvidas em perdas genômicas e em vermelho, moléculas envolvidas em ganhos (IPA).

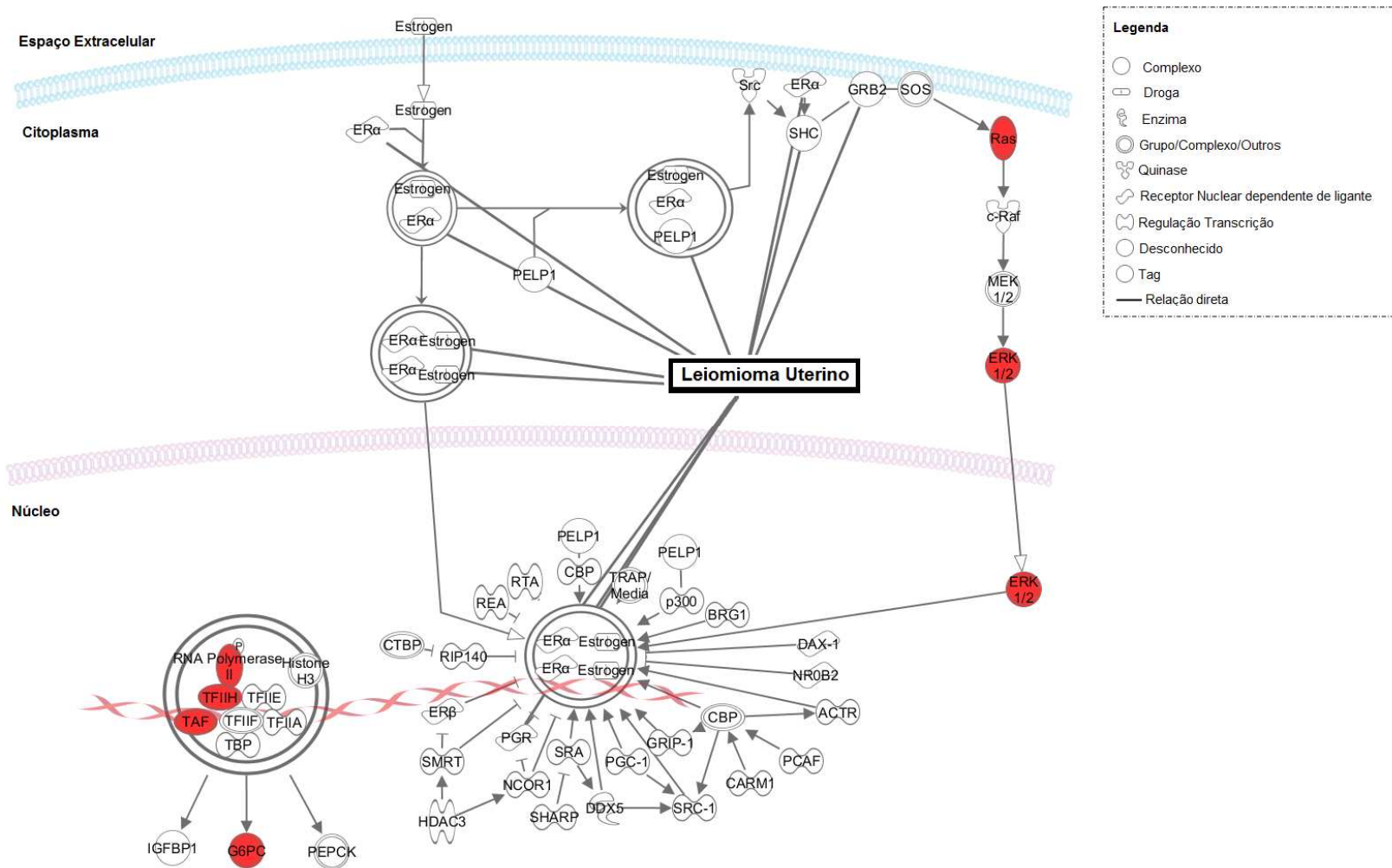


Figura 20. Representação gráfica da via de Sinalização do Receptor de Estrógeno. Os estrógenos podem desencadear estímulos de proliferação celular pela via Ras e induzir a transcrição de genes alvos. Em vermelho, moléculas envolvidas em ganhos genômicos (IPA).

4.3.1 Análise funcional dos genes mapeados em 2q, 7q, 19p e 19q

O agrupamento hierárquico não supervisionado dos dados de CGH array revelou dois grupos segundo o padrão de alterações genômicas. Com o objetivo de identificar as funções biológicas associadas a cada grupo foi realizada uma análise funcional dos genes mapeados nas regiões de ganhos (2q35, 7q22.1, 19p13.12-13.11 e 19q13.32-q13.33) observadas no Grupo 1 e a região de perda (7q22.1) observado no Grupo 2. A análise comparativa dos dados revelou funções biológicas distintas das principais redes (altos *scores*) entre os grupos. O Grupo 1 foi associado com 25 redes e o Grupo 2 com seis redes. A Rede 1 do Grupo 1, apresentou moléculas com funções envolvidas com metabolismo de drogas, desenvolvimento e função do sistema imune e metabolismo de lipídeos, enquanto a Rede 1 do Grupo 2, foi associada com moléculas relacionadas ao desenvolvimento celular. As Figuras 21A e 21B mostram as doenças e funções moleculares associadas com cada grupo, respectivamente. Doenças gastrointestinais, genéticas e inflamatórias foram similares entre os dois grupos, contudo, doenças como câncer e do tecido conjuntivo foram exclusivas apenas aos genes do Grupo 1.

As análises das redes foram sobrepostas para verificar a similaridade das funções exercidas pelos genes entre os grupos. As Redes 3 a 6 do Grupo 2 foram inter-relacionadas com as Redes 1 a 17 do Grupo 1 (Figura 22). No entanto, as Redes 18 a 25 do Grupo 1 e as Redes 1 e 2 do Grupo 2 comportaram-se como redes órfãs, não apresentando conectividade entre elas.

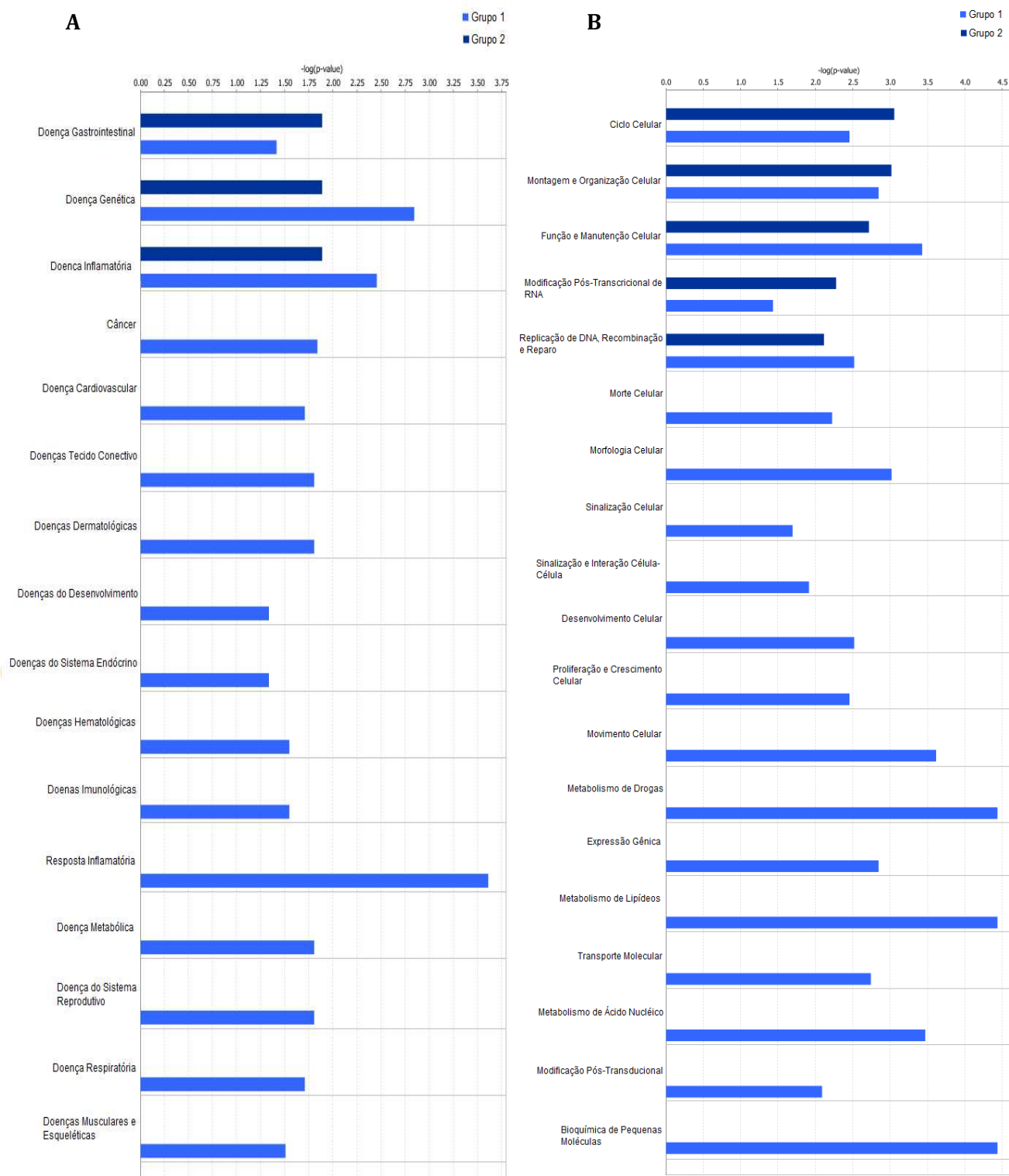


Figura 21. Análise comparativa entre as moléculas do Grupo 1 e Grupo 2. **A e B** – Doenças e funções moleculares similares e exclusivas associadas aos genes de cada grupo, respectivamente (IPA).

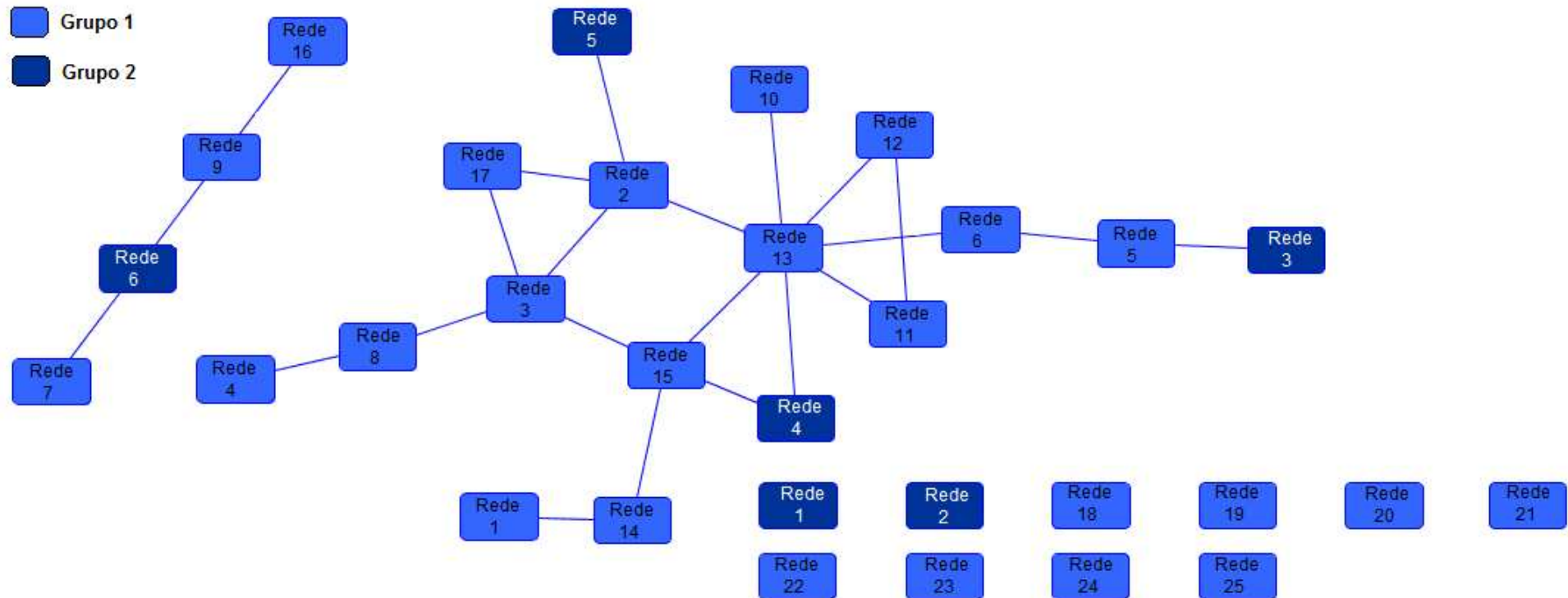


Figura 22. Análise de sobreposição de redes gênicas. As Redes 3 a 6 do Grupo 2 apresentaram sobreposição de funções entre as Redes 1 a 17 do Grupo 1. Também foram observadas geração de redes órfãs (Redes 18 a 25 - Grupo 1 e Redes 1 e 2 - Grupo 2) (IPA).

4.4 Perfil de expressão gênica global

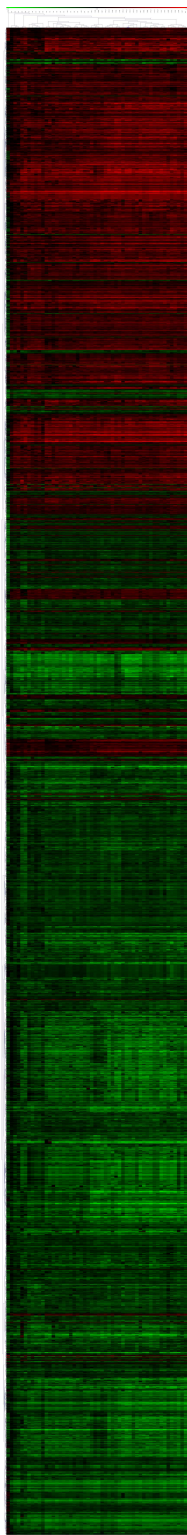
A análise de expressão de transcritos foi realizada em 51 amostras de LU obtidas de 34 pacientes, utilizando a plataforma 4x44K (Agilent Technologies). O teste estatístico SAM (Tusher et al., 2001) com valor de FDR<0,05, foi utilizado para analisar e identificar os genes diferencialmente expressos nos tumores em relação ao miométrio adjacente. O teste de agrupamento hierárquico não-supervisionado, baseado na Correlação de *Pearson*, foi aplicado para verificar se as amostras se agrupavam de acordo com a fase do ciclo menstrual e se todos os tumores se agrupavam diferentemente dos miométrios normais.

Este teste revelou que os miométrios se agruparam diferentemente dos tumores, porém não foi observado um agrupamento das amostras quanto à fase do ciclo menstrual, proliferativa (20 amostras de 13 pacientes) e secretora (31 amostras de 21 pacientes) (Figura 23). Foi observado um agrupamento entre nove amostras na fase secretora do ciclo (642T_B, 729T_C, 722T_C, 317T_C, 317T_A, 317T_B, 300T_B, 300T_A e 684T, nesta ordem), porém este agrupamento não foi significativo.

Os LU múltiplos avaliados de uma única paciente apresentaram uma tendência de agrupamento entre 23 das 29 amostras de 12 pacientes (Figura 23, a partir da amostra 629T_B até a 729T_B). Dezenove casos que tiveram apenas uma amostra avaliada (22 amostras de 22 pacientes) mostraram-se em um outro agrupamento (Figura 23, a partir da amostra 304T até a 615T). As amostras 300T_A e B; 301T_A, B e C; 307T_A e B e 317T_A, B e C agruparam-se, entretanto, um caso 722T teve suas amostras em diferentes grupos (casos 722T_A e 722T_B agruparam-se enquanto o 722T_C estava em outro grupo). Quando foram consideradas as amostras que tiveram diagnóstico de tumores múltiplos (42 amostras/25 pacientes) ou tumores únicos (9 amostras/9 pacientes), não foi observado um agrupamento característico entre as amostras, embora foram observadas oito amostras de tumores únicos em maior proximidade dentro do grupo (855T, 670T, 761T, 708T, 719T, 634T, 644T e 614T, nesta ordem).

Entre os casos avaliados neste estudo, três pacientes relataram história de LU na família, totalizando cinco amostras (629T_A, 629T_B, 683T_A, 683T_B e 857T). Foi investigado se estas amostras se agruparam diferentemente, considerando esta característica. Os resultados indicaram que, pelo menos em relação história de LU na família, as amostras não apresentaram uma tendência de agrupamento (Figura 23).

A



Fase do ciclo

■ Proliferativa

■ Secretora

N° amostras avaliadas

■ 2 ou 3 amostras avaliadas

■ 1 amostra avaliada

Diagnóstico n° tumores

■ Tumor múltiplo

■ Tumor único

B

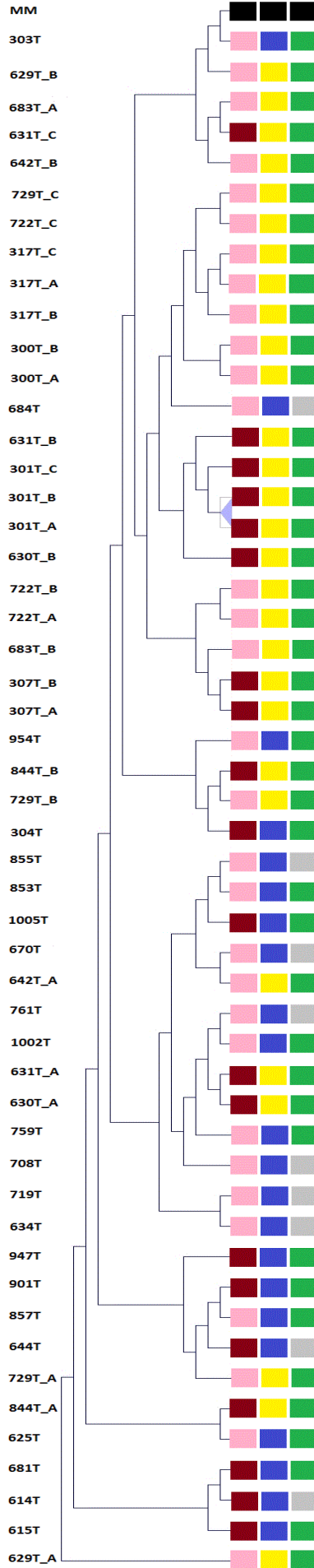


Figura 23. Análise de agrupamento hierárquico não supervisionado entre as 51 amostras de LU e miométrio adjacente. **A** – Perfil de expressão gênica resultante da análise do SAM (Tusher et al., 2001). **B** – Dois subgrupos foram detectados: os miométrios adjacentes (MM) separaram-se dos Leiomiomas Uterinos (LU). Porém, não foi observado um agrupamento significativo das amostras em relação ao período menstrual, presença de tumores múltiplos ou únicos e história de LU na família (*TMeV v.4.5*).

4.4.1 Genes diferencialmente expressos

As análises prévias indicaram que os LU poderiam ser avaliados como amostras independentes e pertencentes a um único grupo. Foram detectados 3325 genes com alteração significativa quanto ao nível de transcritos, os quais foram utilizados na análise integrada. Nesta seção, serão apresentados os dados referentes aos genes diferencialmente expressos nos LU quando comparados ao miométrio adjacente.

Os 3325 genes que apresentaram nível de transcritos considerados significativos na comparação entre os LU e os miométrios, foram avaliados baseados na razão entre leiomioma *versus* miométrio adjacente segundo Roth et al. (2007); as razões foram convertidas em *fold-change* com valor de *cut-off* de 1,5 (Hoffman et al., 2004). Esta comparação permitiu a identificação de 206 genes diferencialmente expressos na comparação entre os LU e os miométrios adjacentes. Foram detectados 136 genes com expressão diminuída (*fold-change* ≤ -1.5) e 70 genes com expressão aumentada (*fold-change* ≥ 1.5), com valores de *fold-change* variando de -9.46 a 50.82 (Tabela 8).

A função relacionada á atividade enzimática foi a mais frequentemente observada entre os genes que apresentaram expressão diminuída (28/136 genes). Entre eles, os genes *EXO1*, *POLQ*, *PPIH* que codificam proteínas envolvidas nos processos de recombinação e reparo a danos no DNA, o gene *UBA7*, alvo do ácido retinóico que degrada o complexo PML/RAR alpha e os genes *GBP2*, *RAB25* que atuam como supressores de tumor.

A segunda categoria mais frequente exercida pelos genes descritos com expressão diminuída foi a regulação da transcrição (16 genes) tais como, *JUN*, *FOS*, *UHRF1*, *CDKN2A*, *SIX1*, *HOXC13* e *TFDP1*.

Entre os genes com expressão aumentada, encontram-se aqueles que exercem funções associadas com a proliferação celular como *IGF1*, *FGF7*, *INSM1*, *STC2*, *SCGB1D2*, *HTATIP2*. Os genes *COL3A1* e *COL4A1* expressos em tecidos conectivos, como o útero, também foram identificados com expressão aumentada.

Tabela 8. Genes diferencialmente expressos nos Leiomiomas Uterinos em relação ao miométrio adjacente.

<i>Gene ID</i>	<i>Símbolo</i>	<i>Nome do Gene</i>	<i>Localização cromossômica</i>	<i>Localização celular</i>	<i>Atividade</i>	<i>Fold-Change</i>
<i>KIF11</i>	<i>KIF11</i>	<i>kinesin family member 11</i>	10q23.33	Núcleo	Outros	-1,500
<i>RFX2</i>	<i>RFX2</i>	<i>regulatory factor X, 2 (influences HLA class II expression)</i>	19p13.3	Núcleo	Regulador transcrição	-1,502
<i>GBP2</i>	<i>GBP2</i>	<i>guanylate binding protein 2, interferon-inducible</i>	1p22.2	Citoplasma	Enzima	-1,504
<i>NEK2</i>	<i>NEK2</i>	<i>NIMA (never in mitosis gene a)-related kinase 2</i>	1q32.3	Citoplasma	Quinase	-1,505
<i>EXO1</i>	<i>EXO1</i>	<i>exonuclease 1</i>	1q43	Núcleo	Enzima	-1,506
<i>CDC20</i>	<i>CDC20</i>	<i>cell division cycle 20 homolog (S. cerevisiae)</i>	1p34.2	Núcleo	Outros	-1,506
<i>MAGEA12</i>	<i>MAGEA12</i>	<i>melanoma antigen family A, 12</i>	Xq26	ND	Outros	-1,507
<i>GPR160</i>	<i>GPR160</i>	<i>G protein-coupled receptor 160</i>	3q26.2	Membrana Celular	Receptor acoplado a proteína-G	-1,507
<i>ATP8B2</i>	<i>ATP8B2</i>	<i>ATPase, class I, type 8B, member 2</i>	1q21.3	Membrana Celular	Transportador	-1,507
<i>UHRF1</i>	<i>UHRF1</i>	<i>ubiquitin-like with PHD and ring finger domains 1</i>	19p13.3	Núcleo	Regulador transcrição	-1,508
<i>GPR126</i>	<i>GPR126</i>	<i>G protein-coupled receptor 126</i>	6q24.1	Membrana Celular	Receptor acoplado a proteína-G	-1,510
<i>NETO2</i>	<i>NETO2</i>	<i>neuropilin (NRP) and tolloid (TLL)-like 2</i>	16q12.1	ND	Outros	-1,512
<i>PCNA</i>	<i>PCNA</i>	<i>proliferating cell nuclear antigen</i>	20p12.3	Núcleo	Outros	-1,512
<i>CBS</i>	<i>CBS</i>	<i>cystathionine-beta-synthase</i>	21q22.3	Citoplasma	Enzima	-1,513
<i>PRLR</i>	<i>PRLR</i>	<i>prolactin receptor</i>	5p13.2	Membrana Celular	Receptor transmembrana	-1,517
<i>TMEM97</i>	<i>TMEM97</i>	<i>transmembrane protein 97</i>	17q11.2	Espaço Extracelular	Outros	-1,520
<i>ZFP36L1</i>	<i>ZFP36L1</i>	<i>zinc finger protein 36, C3H type-like 1</i>	14q24.1	Núcleo	Regulador transcrição	-1,520
<i>RNF128</i>	<i>RNF128</i>	<i>ring finger protein 128</i>	Xq22.3	Citoplasma	Enzima	-1,521

<i>C1orf106</i>	<i>C1orf106</i>	<i>chromosome 1 open reading frame 106</i>	1q32.1	ND	Outros	-1,522
<i>GALNT14</i>	<i>GALNT14</i>	<i>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 14 (GalNAc-T14)</i>	2p23.1	Citoplasma	Enzima	-1,523
<i>CCNF</i>	<i>CCNF</i>	<i>cyclin F</i>	16p13.3	Núcleo	Outros	-1,525
<i>SDC1</i>	<i>SDC1</i>	<i>syndecan 1</i>	2p24.1	Membrana Celular	Outros	-1,526
<i>ESPL1</i>	<i>ESPL1</i>	<i>extra spindle pole bodies homolog 1 (S. cerevisiae)</i>	12q13.13	Núcleo	Peptidase	-1,526
<i>BNC2</i>	<i>BNC2</i>	<i>basonuclin 2</i>	9p22.3-p22.2	Núcleo	Outros	-1,527
<i>SERPINB8</i>	<i>SERPINB8</i>	<i>serpin peptidase inhibitor, clade B (ovalbumin), member 8</i>	18q21.33-q22.1	Citoplasma	Outros	-1,529
<i>SGOL1</i>	<i>SGOL1</i>	<i>shugoshin-like 1 (S. pombe)</i>	3p24.3	Núcleo	Outros	-1,532
<i>CHKA</i>	<i>CHKA</i>	<i>choline kinase alpha</i>	11q13.2	Citoplasma	Quinase	-1,532
<i>ZBTB16</i>	<i>ZBTB16</i>	<i>zinc finger and BTB domain containing 16 phosphoribosylaminoimidazole carboxylase,</i>	11q23.2	Núcleo	Regulador transcrição	-1,532
<i>PAICS</i>	<i>PAICS</i>	<i>phosphoribosylaminoimidazole succinocarboxamide synthetase</i>	4q12	ND	Enzima	-1,533
<i>UBAP2L</i>	<i>UBAP2L</i>	<i>ubiquitin associated protein 2-like</i>	1q21.3	ND	Outros	-1,535
<i>PRSS2</i>	<i>PRSS2</i>	<i>protease, serine, 2 (trypsin 2)</i>	7q34	Espaço Extracelular	Peptidase	-1,535
<i>HOMER2</i>	<i>HOMER2</i>	<i>homer homolog 2 (Drosophila)</i>	15q25.2	Membrana Celular	Outros	-1,536
<i>CBX7</i>	<i>CBX7</i>	<i>chromobox homolog 7</i>	22q13.1	Núcleo	Outros	-1,537
<i>WDR62</i>	<i>WDR62</i>	<i>WD repeat domain 62</i>	19q13.12	Núcleo	Outros	-1,537
<i>GTSE1</i>	<i>GTSE1</i>	<i>G-2 and S-phase expressed 1</i>	22q13.31	Citoplasma	Outros	-1,538
<i>ABCC9</i>	<i>ABCC9</i>	<i>ATP-binding cassette, sub-family C (CFTR/MRP), member 9</i>	12p12.1	Membrana Celular	Canal iônico	-1,545

<i>HAPLN1</i>	<i>HAPLN1</i>	<i>hyaluronan and proteoglycan link protein 1</i>	5q14.3	Espaço Extracelular	Outros	-1,545
<i>STK4</i>	<i>STK4</i>	<i>serine/threonine kinase 4</i>	20q13.12	Citoplasma	Quinase	-1,545
<i>PKMYT1</i>	<i>PKMYT1</i>	<i>protein kinase, membrane associated tyrosine/threonine 1</i>	16p13.3	Citoplasma	Quinase	-1,546
<i>ONECUT2</i>	<i>ONECUT2</i>	<i>one cut homeobox 2</i>	18q21.31	Núcleo	Regulador transcrição	-1,549
<i>BQ233242</i>	<i>RCC2</i>	<i>regulator of chromosome condensation 2</i>	1p36.13	Núcleo	Outros	-1,551
<i>PRAME</i>	<i>PRAME</i>	<i>preferentially expressed antigen in melanoma</i>	22q11.22	Núcleo	Outros	-1,554
<i>FXVD3</i>	<i>FXVD3</i>	<i>FXVD domain containing ion transport regulator 3</i>	19q13.11-q13.12	Membrana Celular	Canal iônico	-1,555
<i>ENST00000369158</i>	<i>HIST2H3C</i>	<i>histone cluster 2, H3c</i>	1q21.2	Núcleo	Outros	-1,555
<i>SPIB</i>	<i>SPIB</i>	<i>Spi-B transcription factor (Spi-1/PU.1 related)</i>	19q13.33	Núcleo	Regulador transcrição	-1,558
<i>CTTNBP2</i>	<i>CTTNBP2</i>	<i>cortactin binding protein 2</i>	7q31.2-q31.31	ND	Outros	-1,562
<i>BMP7</i>	<i>BMP7</i>	<i>bone morphogenetic protein 7</i>	20q13.31	Espaço Extracelular	Fator de crescimento	-1,564
<i>YWHAE</i>	<i>YWHAE</i>	<i>tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, epsilon polypeptide</i>	17p13.3	Citoplasma	Outros	-1,565
<i>NT5DC2</i>	<i>NT5DC2</i>	<i>5'-nucleotidase domain containing 2</i>	3p21.1	ND	Outros	-1,568
<i>FAM63A</i>	<i>FAM63A</i>	<i>family with sequence similarity 63, member A</i>	1q21.2	ND	Outros	-1,572
<i>AK026697</i>	<i>CDS1</i>	<i>CDP-diacylglycerol synthase (phosphatidate cytidyltransferase) 1</i>	4q21.23	Citoplasma	Enzima	-1,576
<i>KIFC1</i>	<i>KIFC1</i>	<i>kinesin family member C1</i>	6p21.32	Núcleo	Outros	-1,577
<i>AICDA</i>	<i>AICDA</i>	<i>activation-induced cytidine deaminase</i>	12p13	Citoplasma	Enzima	-1,579
<i>PTPRJ</i>	<i>PTPRJ</i>	<i>protein tyrosine phosphatase, receptor type, J</i>	11p11.2	Membrana Celular	Fosfatase	-1,580
<i>POLQ</i>	<i>POLQ</i>	<i>polymerase (DNA directed), theta</i>	3q13.33	Núcleo	Enzima	-1,582
<i>C21orf113</i>	<i>NCRNA00162</i>	<i>non-protein coding RNA 162</i>	21q22.3	ND	Outros	-1,582
<i>TNFRSF10B</i>	<i>TNFRSF10B</i>	<i>tumor necrosis factor receptor superfamily,</i>	8p21.3	Membrana Celular	Receptor	-1,584

		member 10b			transmembrana	
CCDC21	CCDC21	<i>coiled-coil domain containing 21</i>	1p36.11	ND	Outros	-1,585
UBE1L	UBA7	<i>ubiquitin-like modifier activating enzyme 7</i>	3p21.31	Citoplasma	Enzima	-1,586
CDKN2A	CDKN2A	<i>cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)</i>	9p21	Núcleo	Regulador transcrição	-1,586
CNOT6	CNOT6	<i>CCR4-NOT transcription complex, subunit 6</i>	5q35.3	Núcleo	Outros	-1,588
SVEP1	SVEP1	<i>sushi, von Willebrand factor type A, EGF and pentraxin domain containing 1</i>	9q31.3	Citoplasma	Outros	-1,588
CD1A	CD1A	<i>CD1a molecule</i>	1q23.1	Membrana Celular	Outros	-1,594
CDC25A	CDC25A	<i>cell division cycle 25 homolog A (S. pombe)</i>	3p21.31	Núcleo	Fosfatase	-1,595
SLC19A1	SLC19A1	<i>solute carrier family 19 (folate transporter), member 1</i>	21q22.3	Membrana Celular	Transportador	-1,595
AP1S3	AP1S3	<i>adaptor-related protein complex 1, sigma 3 subunit</i>	2q36.1	Citoplasma	Transportador	-1,597
ALDH4A1	ALDH4A1	<i>aldehyde dehydrogenase 4 family, member A1</i>	1p36.13	Citoplasma	Enzima	-1,601
SCD	SCD	<i>stearoyl-CoA desaturase (delta-9-desaturase)</i>	10q24.3	Citoplasma	Enzima	-1,603
XAF1	XAF1	<i>XIAP associated factor 1</i>	17p13.2	Núcleo	Outros	-1,603
SIX1	SIX1	<i>SIX homeobox 1</i>	14q23.1	Núcleo	Regulador transcrição	-1,603
PPIH	PPIH	<i>peptidylprolyl isomerase H (cyclophilin H)</i>	1p34.2	Núcleo	Enzima	-1,605
ABI3	ABI3	<i>ABI family, member 3</i>	17q21.32	Membrana Celular	Outros	-1,605
KIF4A	KIF4A	<i>kinesin family member 4A</i>	Xq13.1	Núcleo	Outros	-1,607
PPIA	PPIA	<i>peptidylprolyl isomerase A (cyclophilin A)</i>	7p13	Citoplasma	Enzima	-1,608
MAGEB2	MAGEB2	<i>melanoma antigen family B, 2</i>	Xp21.2	Citoplasma	Outros	-1,609
FASN	FASN	<i>fatty acid synthase</i>	17q25.3	Citoplasma	Enzima	-1,611
BG680979	FAM110A	<i>family with sequence similarity 110, member A</i>	20p13	ND	Outros	-1,612
HLA-DRB4	HLA-DRB4	<i>major histocompatibility complex, class II, DR</i>	6p21.3	Membrana Celular	Receptor transmembrana	-1,614

		<i>beta 4</i>				
MYO10	MYO10	<i>myosin X</i>	5p15.1	Citoplasma	Outros	-1,615
ENPP2	ENPP2	<i>ectonucleotide pyrophosphatase/phosphodiesterase 2</i>	8q24.12	Membrana Celular	Enzima	-1,616
HOXC13	HOXC13	<i>homeobox C13</i>	12q13.3	Núcleo	Regulador transcrição	-1,618
CDCA3	CDCA3	<i>cell division cycle associated 3</i>	12p13.31	Citoplasma	Outros	-1,625
TFDP1	TFDP1	<i>transcription factor Dp-1</i>	13q34	Núcleo	Regulador transcrição	-1,625
PMAIP1	PMAIP1	<i>phorbol-12-myristate-13-acetate-induced protein 1</i>	18q21.32	Citoplasma	Outros	-1,627
FANCI	FANCI	<i>Fanconi anemia, complementation group I</i>	15q26.1	Núcleo	Outros	-1,628
PAOX	PAOX	<i>polyamine oxidase (exo-N4-amino)</i>	10q26.3	Citoplasma	Enzima	-1,630
GPR19	GPR19	<i>G protein-coupled receptor 19</i>	12p13.1	Membrana Celular	Receptor acoplado a proteína-G	-1,630
RNASE1	RNASE1	<i>ribonuclease, RNase A family, 1 (pancreatic)</i>	14q11.2	Espaço Extracelular	Enzima	-1,637
DLL3	DLL3	<i>delta-like 3 (Drosophila)</i>	19q13.2	Espaço Extracelular	Outros	-1,640
GPT2	GPT2	<i>glutamic pyruvate transaminase (alanine aminotransferase) 2</i>	16q11.2-q12.1	ND	Enzima	-1,641
FGL2	FGL2	<i>fibrinogen-like 2</i>	7q11.23	Espaço Extracelular	Peptidase	-1,644
MCM10	MCM10	<i>minichromosome maintenance complex component 10</i>	10p13	Núcleo	Outros	-1,649
SLC25A10	SLC25A10	<i>solute carrier family 25 (mitochondrial carrier; dicarboxylate transporter), member 10</i>	17q25.3	Citoplasma	Transportador	-1,649
PGAM5	PGAM5	<i>phosphoglycerate mutase family member 5</i>	12q24.33	Citoplasma	Enzima	-1,652
CDT1	CDT1	<i>chromatin licensing and DNA replication factor 1</i>	16q24.3	Núcleo	Outros	-1,656
PLK1	PLK1	<i>polo-like kinase 1</i>	16p12.1	Núcleo	Quinase	-1,658
PRSS3	PRSS1/PRSS3	<i>protease, serine, 1 (trypsin 1)</i>	9p13.3	Espaço Extracelular	Peptidase	-1,661

<i>GPX2</i>	<i>GPX2</i>	<i>glutathione peroxidase 2 (gastrointestinal)</i>	14q23.3	Citoplasma	Enzima	-1,665
<i>C10orf54</i>	<i>C10orf54</i>	<i>chromosome 10 open reading frame 54</i>	10q22.1	ND	Outros	-1,672
<i>ENST00000284767</i>	<i>PDLIM3</i>	<i>PDZ and LIM domain 3</i>	4q35.1	Citoplasma	Outros	-1,674
<i>TNNC1</i>	<i>TNNC1</i>	<i>troponin C type 1 (slow)</i>	3p21.1	Citoplasma	Outros	-1,687
<i>DGAT2</i>	<i>DGAT2</i>	<i>diacylglycerol O-acyltransferase 2</i>	11q13.5	Citoplasma	Enzima	-1,699
<i>PACSIN1</i>	<i>PACSIN1</i>	<i>protein kinase C and casein kinase substrate in neurons 1</i>	6p21.31	Citoplasma	Quinase	-1,701
<i>CACNG6</i>	<i>CACNG6</i>	<i>calcium channel, voltage-dependent, gamma subunit 6</i>	19q13.42	Membrana Celular	Canal iônico	-1,702
<i>ARID5A</i>	<i>ARID5A</i>	<i>AT rich interactive domain 5A (MRF1-like)</i>	2q11.2	Núcleo	Regulador transcrição	-1,702
<i>C4BPB</i>	<i>C4BPB</i>	<i>complement component 4 binding protein, beta</i>	1q32.2	Espaço Extracelular	Outros	-1,717
<i>AZGP1</i>	<i>AZGP1</i>	<i>alpha-2-glycoprotein 1, zinc-binding</i>	7q22.1	Espaço Extracelular	Transportador	-1,719
<i>AURKB</i>	<i>AURKB</i>	<i>aurora kinase B</i>	17p13.1	Núcleo	Quinase	-1,733
<i>C16orf57</i>	<i>C16orf57</i>	<i>chromosome 16 open reading frame 57</i>	16q13	ND	Outros	-1,740
<i>MTHFD1L</i>	<i>MTHFD1L</i>	<i>methylenetetrahydrofolate dehydrogenase (NADP+ dependent) 1-like</i>	6q25.1	Citoplasma	Enzima	-1,746
<i>RAB25</i>	<i>RAB25</i>	<i>RAB25, member RAS oncogene family</i>	1q22	Citoplasma	Enzima	-1,749
<i>C5orf34</i>	<i>C5orf34</i>	<i>chromosome 5 open reading frame 34</i>	5p12	ND	Outros	-1,752
<i>HBE1</i>	<i>HBE1</i>	<i>hemoglobin, epsilon 1</i>	11p15.4	Citoplasma	Transportador	-1,772
<i>KIAA1305</i>	<i>NYNRIN</i>	<i>NYN domain and retroviral integrase containing eukaryotic translation elongation factor 1 alpha</i>	14q12	ND	Outros	-1,775
<i>EEF1A2</i>	<i>EEF1A2</i>	<i>2</i>	20q13.33	Citoplasma	Regulador tradução	-1,791
<i>JUN</i>	<i>JUN</i>	<i>jun proto-oncogene</i>	1p32-p31	Núcleo	Regulador transcrição	-1,814
<i>GPR116</i>	<i>GPR116</i>	<i>G protein-coupled receptor 116</i>	6p12.3	Membrana Celular	Receptor acoplado a proteína-G	-1,829
<i>FAM107A</i>	<i>FAM107A</i>	<i>family with sequence similarity 107, member A</i>	3p14.2	Núcleo	Outros	-1,831

<i>DIO1</i>	<i>DIO1</i>	<i>deiodinase, iodothyronine, type I</i>	1p32.3	Citoplasma	Enzima	-1,867
<i>NRG1</i>	<i>NRG1</i>	<i>neuregulin 1</i>	8p12	Espaço Extracelular	Fator de crescimento	-1,867
<i>ACACA</i>	<i>ACACA</i>	<i>acetyl-CoA carboxylase alpha</i>	17q12	Citoplasma	Enzima	-1,870
<i>HBD</i>	<i>HBD</i>	<i>hemoglobin, delta</i>	11p15.5	Citoplasma	Transportador	-1,877
<i>SIX6</i>	<i>SIX6</i>	<i>SIX homeobox 6</i>	14q23.1	Núcleo	Regulador transcrição	-1,880
<i>VGLL1</i>	<i>VGLL1</i>	<i>vestigial like 1 (Drosophila)</i>	Xq26.3	Núcleo	Regulador transcrição	-1,912
<i>NME1</i>	<i>NME1</i>	<i>non-metastatic cells 1, protein (NM23A) expressed in</i>	17q21.3	Núcleo	Quinase	-1,920
<i>RAB17</i>	<i>RAB17</i>	<i>RAB17, member RAS oncogene family</i>	2q37.3	Citoplasma	Enzima	-1,936
<i>GPRIN1</i>	<i>GPRIN1</i>	<i>G protein regulated inducer of neurite outgrowth 1</i>	5q35.2	Membrana Celular	Outros	-1,946
<i>MAPK10</i>	<i>MAPK10</i>	<i>mitogen-activated protein kinase 10</i>	4q21.3	Citoplasma	Quinase	-1,949
<i>CDCA2</i>	<i>CDCA2</i>	<i>cell division cycle associated 2</i>	8p21.2	Núcleo	Outros	-1,956
<i>C11orf41</i>	<i>C11orf41</i>	<i>chromosome 11 open reading frame 41</i>	11p13	Citoplasma	Outros	-1,988
<i>C1orf67</i>	<i>DNAH14</i>	<i>dynein, axonemal, heavy chain 14</i>	1q42.12	ND	Outros	-1,999
<i>FOS</i>	<i>FOS</i>	<i>FBJ murine osteosarcoma viral oncogene homolog</i>	6q31	Núcleo	Regulador transcrição	-2,088
<i>SFN</i>	<i>SFN</i>	<i>stratifin</i>	1p36.11	Citoplasma	Outros	-2,156
<i>CXCL5</i>	<i>CXCL5</i>	<i>chemokine (C-X-C motif) ligand 5</i>	4q13.3	Espaço Extracelular	Citocina	-2,205
<i>EPB41L4B</i>	<i>EPB41L4B</i>	<i>erythrocyte membrane protein band 4.1 like 4B</i>	9q31.3	ND	Transportador	-2,555
<i>PTBP1</i>	<i>PTBP1</i>	<i>polypyrimidine tract binding protein 1</i>	19p13.3	Núcleo	Enzima	-9,456
<i>EIF2C2</i>	<i>EIF2C2</i>	<i>eukaryotic translation initiation factor 2C, 2</i>	8q24.3	Citoplasma	Regulador tradução	1,506
<i>RGS17</i>	<i>RGS17</i>	<i>regulator of G-protein signaling 17</i>	6q25.2	Citoplasma	Outros	1,518
<i>E2F5</i>	<i>E2F5</i>	<i>E2F transcription factor 5, p130-binding</i>	5p13.2	Núcleo	Regulador transcrição	1,524
<i>TMEM22</i>	<i>TMEM22</i>	<i>transmembrane protein 22</i>	3q22.3	Citoplasma	Outros	1,528
<i>RTN3</i>	<i>RTN3</i>	<i>reticulon 3</i>	11q13.1	Citoplasma	Outros	1,542

VAPB	VAPB	VAMP (vesicle-associated membrane protein)-associated protein B and C	20q13.32	Membrana Celular	Outros	1,567
RPL19	RPL19	ribosomal protein L19	17q12	Citoplasma	Outros	1,581
SOX4	SOX4	SRY (sex determining region Y)-box 4	6p22.3	Núcleo	Regulador transcrição	1,584
DIAPH1	DIAPH1	diaphanous homolog 1 (Drosophila)	5q31.3	Citoplasma	Outros	1,585
IARS2	IARS2	isoleucyl-tRNA synthetase 2, mitochondrial	1q41	Citoplasma	Enzima	1,588
ATP8B4	ATP8B4	ATPase, class I, type 8B, member 4	15q21.2	ND	Transportador	1,589
TEAD1	TEAD1	TEA domain family member 1 (SV40 transcriptional enhancer factor)	11p15.2	Núcleo	Regulador transcrição	1,593
HIST1H4C	HIST4H4	histone cluster 1, H4c	6p22.1	Núcleo	Outros	1,597
MPZL1	MPZL1	myelin protein zero-like 1	1q24.2	Membrana Celular	Outros	1,597
TUBA1C	TUBA1C	tubulin, alpha 1c	12q13.12	Citoplasma	Outros	1,601
RGR	RGR	retinal G protein coupled receptor	10q23.1	Membrana Celular	Receptor acoplado a proteína-G	1,605
ST7	ST7	suppression of tumorigenicity 7	7q31.2	ND	Outros	1,613
TUBA1B	TUBA1B	tubulin, alpha 1b	12q13.12	Citoplasma	Outros	1,623
LYPD1	LYPD1	LY6/PLAUR domain containing 1	2q21.2	Membrana Celular	Receptor acoplado a proteína-G	1,639
KCNMA1	KCNMA1	potassium large conductance calcium-activated channel, subfamily M, alpha member 1	10q22.3	Membrana Celular	Canal iônico	1,657
MT1E	MT1E	metallothionein 1E	16q13	Citoplasma	Outros	1,671
TOE1	TOE1	target of EGR1, member 1 (nuclear)	1p34.1	Núcleo	Outros	1,673
PIGH	PIGH	phosphatidylinositol glycan anchor biosynthesis, class H	14q24.1	Citoplasma	Enzima	1,685
MT2A	MT2A	metallothionein 2A	16q13	Citoplasma	Outros	1,694
CDKN2D	CDKN2D	cyclin-dependent kinase inhibitor 2D (p19,	19p13.2	Núcleo	Regulador transcrição	1,705

		<i>inhibits CDK4)</i>				
COLEC12	COLEC12	<i>collectin sub-family member 12</i>	18p11.32	Membrana Celular	Receptor transmembrana	1,710
TUBA4A	TUBA4A	<i>tubulin, alpha 4a</i>	2q35	Citoplasma	Outros	1,714
FGD3	FGD3	<i>FYVE, RhoGEF and PH domain containing 3</i>	9q22.31	Citoplasma	Outros	1,731
CHCHD2	CHCHD2	<i>coiled-coil-helix-coiled-coil-helix domain containing 2</i>	7p11.2	Citoplasma	Outros	1,741
PYGL	PYGL	<i>phosphorylase, glycogen, liver</i>	14q22.1	ND	Enzima	1,761
FAM40B	FAM40B	<i>family with sequence similarity 40, member B</i>	7q32.1	ND	Outros	1,796
MAL	MAL	<i>mal, T-cell differentiation protein</i>	2q11.1	Membrana Celular	Transportador	1,847
CCDC109B	CCDC109B	<i>coiled-coil domain containing 109B</i>	4q25	ND	Outros	1,852
ANXA3	ANXA3	<i>annexin A3</i>	4q21.21	Citoplasma	Enzima	1,895
GALNT3	GALNT3	<i>UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 3 (GalNAc-T3)</i>	2q24.3	Citoplasma	Enzima	1,898
AK124698	GRIN2A	<i>glutamate receptor, ionotropic, N-methyl D-aspartate 2A</i>	16p13.2	Membrana Celular	Canal iônico	1,902
ASRGL1	ASRGL1	<i>asparaginase like 1</i>	11q12.3	Citoplasma	Enzima	1,909
PHLDA3	PHLDA3	<i>pleckstrin homology-like domain, family A, member 3</i>	1q32.1	Membrana Celular	Outros	1,917
RPS10	RPS10	<i>ribosomal protein S10</i>	6p21.31	Citoplasma	Outros	1,923
SKAP2	SKAP2	<i>src kinase associated phosphoprotein 2</i>	7p15.2	Citoplasma	Outros	1,926
DLG4	DLG4	<i>discs, large homolog 4 (Drosophila)</i>	17p13.1	Membrana Celular	Quinase	1,929
CD247	CD247	<i>CD247 molecule</i>	1q24.2	Membrana Celular	Receptor transmembrana	1,938
ABCC2	ABCC2	<i>ATP-binding cassette, sub-family C (CFTR/MRP), member 2</i>	10q24.2	Membrana Celular	Transportador	1,965
P2RY2	P2RY2	<i>purinergic receptor P2Y, G-protein coupled, 2</i>	11q13.4	Membrana Celular	Receptor acoplado a	1,980

proteína-G

HTATIP2	HTATIP2	HIV-1 Tat interactive protein 2, 30kDa	11p15.1	Núcleo	Regulador transcrição	2,062
INSM1	INSM1	insulinoma-associated 1	20p11.23	Núcleo	Outros	2,108
IGF1	IGF1	insulin-like growth factor 1 (somatomedin C)	12q23.2	Espaço Extracelular	Fator de crescimento	2,125
STC2	STC2	stanniocalcin 2	5q35.2	Espaço Extracelular	Outros	2,126
BG001037	TXNRD1	thioredoxin reductase 1	12q23.3	Citoplasma	Enzima	2,193
FGF7	FGF7	fibroblast growth factor 7	15q21.1	Espaço Extracelular	Fator de crescimento	2,282
WWOX	WWOX	WW domain containing oxidoreductase	16q23.1	Citoplasma	Enzima	2,286
ATP6V1C2	ATP6V1C2	ATPase, H+ transporting, lysosomal 42kDa, V1 subunit C2	2p25.1	ND	Transportador	2,349
PBX1	PBX1	pre-B-cell leukemia homeobox 1	1q23.3	Núcleo	Regulador transcrição	2,352
GCH1	GCH1	GTP cyclohydrolase 1	14q22.2-q22.3	Citoplasma	Enzima	2,356
IGJ	IGJ	immunoglobulin J polypeptide, linker protein for immunoglobulin alpha and mu polypeptides	4q13.3	Espaço Extracelular	Outros	2,367
COL17A1	COL17A1	collagen, type XVII, alpha 1	10q25.1	Membrana Celular	Outros	2,569
ODC1	ODC1	ornithine decarboxylase 1	2p25.1	Citoplasma	Enzima	2,588
AK021531	COL3A1	collagen, type III, alpha 1	2q32.2	Espaço Extracelular	Outros	2,647
SEMA4D	SEMA4D	sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4D	9q22.2	Membrana Celular	Outros	2,943
AL050204	EXPH5	exophilin 5	11q22.3	ND	Outros	3,178
STRBP	STRBP	spermatid perinuclear RNA binding protein	9q33.2	Citoplasma	Outros	3,228
COL4A1	COL4A1	collagen, type IV, alpha 1	13q34	Espaço Extracelular	Outros	3,528
ZDHHC2	ZDHHC2	zinc finger, DHHC-type containing 2	8p22	Núcleo	Enzima	3,596
SCGB1D2	SCGB1D2	secretoglobin, family 1D, member 2	11q12.3	Espaço Extracelular	Outros	3,614
NAV2	NAV2	neuron navigator 2	11p15.1	Núcleo	Outros	3,716

<i>CENPK</i>	<i>CENPK</i>	<i>centromere protein K</i>	5q12.3	Núcleo	Outros	3,771
<i>DHFRL1</i>	<i>DHFRL1</i>	<i>dihydrofolate reductase-like 1</i>	3q11.2	ND	Outros	4,328
<i>UNC5D</i>	<i>UNC5D</i>	<i>unc-5 homolog D (C. elegans)</i>	8p12	ND	Outros	4,776
<i>TOMM34</i>	<i>TOMM34</i>	<i>translocase of outer mitochondrial membrane 34</i>	20q13.12	Citoplasma	Outros	10,860
<i>ZNF516</i>	<i>ZNF516</i>	<i>zinc finger protein 516</i>	18q23	Núcleo	Outros	50,818

Em negrito, genes identificados em estudos prévios em Leiomiomas Uterinos. *Gene ID* = identificação do gene; *Fold-Change* = valores negativos, genes com expressão diminuída e valores positivos, genes com expressão aumentada; ND = não determinado.

4.4.2 Análise funcional dos genes diferencialmente expressos

Os genes diferencialmente expressos foram submetidos a análises de funções, redes e vias canônicas no sistema IPA. Foram geradas 13 redes com valores de *score* variando de 12 a 52, no qual os valores mais altos indicam maior associação entre as moléculas. As principais funções identificadas foram associadas com mecanismos de regulação do ciclo celular (Redes 1, 2, 3, 4, 6 e 10 - Tabela 9). A Rede 1 (*score* 52) apresentou conexões entre os genes associados com ciclo celular, câncer, doença do sistema reprodutivo. A Rede 2 contém genes com funções associadas ao ciclo celular, doenças hematológicas e metabolismo de carboidratos (*score* 45). A Rede 3 mostra genes em associação com a morfologia do organismo, desenvolvimento celular, organização e montagem celular. Contudo, outras redes foram geradas a partir de genes que desempenham funções associadas com LU, como a Rede 8 que contém genes relacionados a função de regulação da expressão gênica, desenvolvimento embrionário e do organismo (*score* 18). A Rede 9 (*score* 18) apresenta genes envolvidos com o câncer, doenças e condições dermatológicas e doenças do sistema reprodutivo. A Figura 24A e B é uma representação gráfica da interação entre os genes das Redes 1 e 2, respectivamente. As funções e doenças foram posicionadas em valores de $-\log(P \text{ valor})$. Esta classificação revelou as principais funções dos genes associadas com morte celular (6.0), ciclo celular (5.5), proliferação e crescimento celular (3.9), organização e montagem celular (3.6) e manutenção e função celular (3.6) (Figura 25). Entre as principais doenças que aparecem com maiores valores de $-\log(P \text{ valor})$ encontram-se o câncer (9.9), doenças gastrointestinais (9.8), doenças genéticas (9.8), doenças do sistema reprodutivo (4.5) e doenças musculares e esqueléticas (4.4) (Figura 26).

As vias canônicas identificadas foram descritas como a razão dos valores de frequência dos genes diferencialmente expressos em relação às moléculas do sistema IPA (Figura 27). As principais vias descritas foram associadas com o reparo a erros de

pareamento (*Mismatch Repair*) em eucariotos (69%), regulação do *checkpoint* G2/M do ciclo celular em resposta a danos no DNA (55%), regulação do *checkpoint* G1/S do ciclo celular (45%), sinalização da apoptose mediada por MYC (44%) e regulação do ciclo celular e ciclinas (30%). A via controle do ciclo celular relacionada a duplicação cromossômica teve o maior número de genes identificados, porém, com baixa associação entre as moléculas com base no banco de dados do sistema IPA (aproximadamente 12% de associação).

Outras vias já descritas em LU foram geradas com menores valores de associação, como sinalização IGF-1 e TGF-beta (25% e 19%, respectivamente). Entre estas vias, foram selecionadas as vias de regulação do ciclo celular G1/S e a via de sinalização IGF-1 para representar graficamente os dois principais achados entre os genes diferencialmente expressos incluindo os genes envolvidos nas vias de mecanismos do ciclo celular e proliferação e crescimento celular (Figuras 28 e 29, respectivamente).

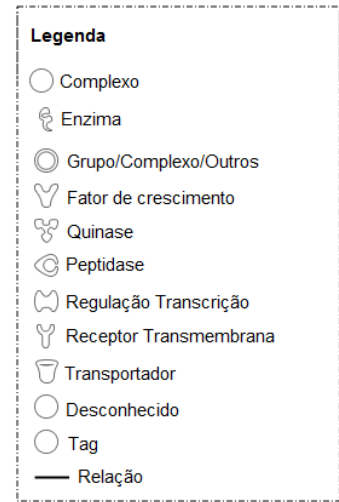
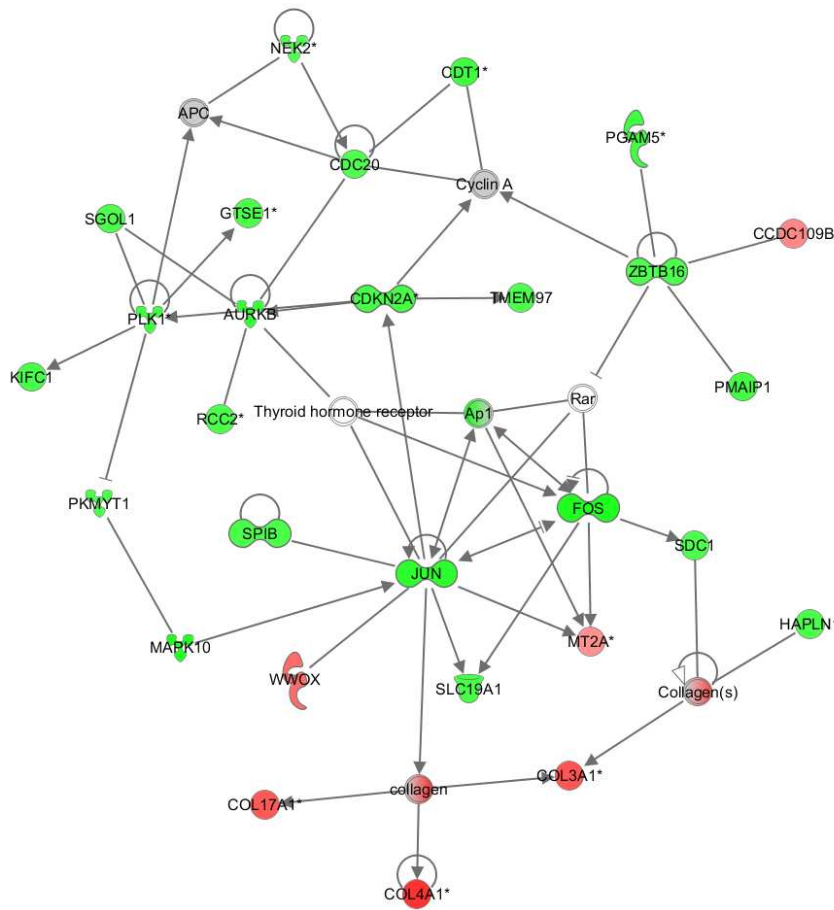
Tabela 9. Redes e principais funções geradas a partir da associação entre os genes diferencialmente expressos analisados pelo sistema IPA (*Ingenuity Pathway Analysis*).

Redes	Moléculas na rede	Score	Moléculas em foco	Principais funções
1	<i>Ap1, APC, ↓AURKB, ↑CCDC109B, ↓CDC20, ↓CDKN2A, ↓CDT1, ↑COL17A1, ↑COL3A1, ↑COL4A1, collagen, Collagen(s), Cyclin A, ↓FOS, ↓GTSE1, ↓HAPLN1, ↓JUN, ↓KIFC1, ↓MAPK10, ↑MT2A, ↓NEK2, ↓PGAM5, ↓PKMYT1, ↓PLK1, ↓PMAIP1, Rar, ↓RCC2, ↓SDC1, ↓SGOL1, ↓SLC19A1, ↓SPIB, Thyroid hormone receptor, ↓TMEM97, ↑WWOX, ↓ZBTB16</i>	52	28	Ciclo Celular, Câncer, Doença do Sistema Reprodutivo
2	<i>14-3-3, ↓C1orf106, ↓C5orf34, ↑CD247, ↓CDC25A, ↓DIO1, E2f, ↑E2F5, E2f-Tfdp1, ↓EXO1, ↓FASN, FKBP14, ↓HBD, ↓HBE1, ↓HIST2H3C (includes others), ↑HIST4H4 (includes others), Histone h3, Histone h4, ↑HTATIP2, INS, ↓MCM10, MIR124 (human), PAPD7, ↓PCNA, ↓POLQ, ↓PPIA, ↓PRLR, ↓PTBP1, RNA polymerase II, ↓SFN, ↑TEAD1, ↓TFDP1, ↓UHRF1, ↓VGLL1, ↓YWHAE</i>	45	25	Ciclo Celular, Doença Hematológica, Metabolismo de Carboidratos
3	<i>↓ACACA, ATP11C, ↓ATP8B2, ↑ATP8B4, ↓BRCA1, cacn, ↓CACNA1H, CACNG2, CACNG5, ↓CACNG6, CACNG7, CARD16, ↓CDS1, CTBP1, ↓GPT2, ↑IARS2, IFI6, IRF7, ↑KCNMA1, KDM5B, KIF1B, ↑LYPD1, Mg2+-ATPase, ↑MT1E, ↓ONECUT2, ↓RAB17, S100A8, SATB1, ↓SIX6, ↑TOMM34, ↓UBA7, VCP, ↓XAF1, ZNF217</i>	24	16	Morfologia do Organismo, Desenvolvimento celular, Organização e Montagem Celular
4	<i>↓AB13, Adaptor protein 1, ALDOC, ↓AP1S3, ↓CHKA, choline, CTBP1, ↓ENPP2, ↓ESR1, F2, F2RL3, ↓FGL2, HTT, ↑IGJ, ↓MAGEA12, ↓NCRNA00162, ↓NME1, NOS3, PAK1, PIGR, ↓PRSS1/PRSS3, ↓RAB25, ↓RNASE1, ↑RPS10, ↓SERPINB8, SNAI1, STAT5A, TAF12, TAF1A, TAF4B, TBP, UBTF, ↓WDR62, ↑ZDHHC2, ↑ZNF516</i>	24	17	Desenvolvimento Embrionário, Desenvolvimento Tecidual, Movimento Celular
5	<i>↓AZGP1, ↓BACE1, ↓C4BPB, ↓CDCA3, ↓CTTNBP2, CTTNBP2NL, ↑DIAPH1, ↑DLG4, ↓DNAH14, DYNLL1, DYNLT1, Erbb4 dimer, ↓FAM107A, FAM40A, ↑FAM40B, Grin-Grin-Grin-Dlg4, ↑GRIN2A, GRIN2D, GRIN3A, ↓KIF11, ↓LRP1, ↓NRG1, ↓NT5DC2, ↑PPP2R1B, ↓PRSS2, Ras homolog, ↑RTN3, SIKE1, ↓SLC25A10, SLMAP, STK24, SYNGAP1, TRAF2, TRAF3IP3, USP53</i>	24	16	Doença Genética, Doença Oftalmológica, Desenvolvimento do Sistema Nervoso
6	<i>AHSA1, ↑ANXA3, ↓CBX7, ↓CNOT6, CSTF1, CTBP1, DNAJB6, DR4/5, E2F4, ↓EPB41L4B, ↑FGF7, ↑GCH1, HDAC2, HMG20B, HOXB7, HSP90AA1, ↓KIF4A, LBR, MIR1, ↓MTHFD1L, ↓MYO10, ↓NEK2, ↓NETO2, NOC2L, ↓PACSIN1, ↑PBX1, ↑PHLDA3, ↑SERPINB5, ↓SMC4, TMSB10/TMSB4X, ↓TNFRSF10B, ↑TOE1, TP53, TPM4, TRIO</i>	23	16	Morte Celular, Organização e Montagem Celular, Função e Manutenção Celular
7	<i>ARRB2, BAI1, CHEMOKINE, CHRM4, ↓CXCL5, CXCR1, ↓EEF1A2, F2RL1, F2RL3, ↓FANCI, FZD1, FZD2, ↓FZD4, ↑FZD7, GNRHR, Gpcr, ↓GPR19, GPR50, ↓GPR116, ↓GPR126, ↓GPR160, ↓GPRC5A, GRM1, ↓HOMER2, ↑HTR2A, HTR2C, LPHN3, ↑P2RY2, ↓PPIH, ↑RGR, ↑STC2, ↓TNNC1, UBC, USP15</i>	18	13	Metabolismo Carboidratos, Metabolismo Lipídeos, Bioquímica de Pequenas Moléculas
8	<i>↑ABCC2, AGTPBP1, APOC3, ↓C16orf57, ↓CBS, ↓CCDC21, ↑CHCHD2, ↓DGAT2, DPM1, DPM2, ↑E2F5, HNF4A,</i>	18	13	Expressão Gênica, Desenvolvimento

	<i>JUB</i> , ↓ MAGEB2 , <i>NR2F2</i> , ↓ NYNRIN , <i>PARD3B</i> , <i>PDZK1</i> , <i>PIGA</i> , <i>PIGC</i> , ↑ PIGH , <i>PIGP</i> , <i>PIGQ</i> , <i>PIGY</i> , ↓ PRAME , <i>RARA</i> , ↑ RPL28 , <i>RPL31</i> , <i>RPL18A</i> , ↓ SCD , <i>SKI</i> , <i>SMAD4</i> , <i>STK19</i> , ↓ TCF7 , ↑ TXNRD1			Embrionário, Desenvolvimento do Organismo
9	<i>2-methoxyestradiol</i> , <i>ACTN2</i> , <i>Alpha tubulin</i> , ↑ ASCL2 , ↓ CCNF , <i>CD48</i> , <i>CD55</i> , ↓ CDCA2 , ↓ CDK1 , ↓ ESPL1 , <i>FYN</i> , ↓ GBP2 , <i>GTP</i> , ↑ INCENP , ↑ INSM1 , <i>ITPRIP</i> , <i>LIMK1</i> , <i>MAG</i> , ↑ MAL , <i>MAP1S</i> , ↓ PDLIM3 , <i>PITPNM1</i> , <i>RASSF5</i> , ↑ ST7 , ↓ STK4 , <i>TBCD</i> , <i>TMEM109</i> , ↑ TUBA1B , ↑ TUBA1C , ↑ TUBA4A , <i>TUBB3</i> , <i>TUBB2B</i> , <i>Tubulin</i> , ↑ VAPB , <i>VDAC1</i>	18	13	Câncer, Doenças e Condições Dermatológicas, Doença do Sistema Reprodutivo
10	↓ AICDA , ↓ BNC2 , <i>CCNA1</i> , <i>CCNG1</i> , ↓ CDKN2A , <i>Cdkn2a-Mdm2</i> , ↑ CENPK , <i>CTGF</i> , <i>Cyclin D</i> , ↓ DLL3 , <i>E4F1</i> , ↑ EIF2C2 , <i>FANCL</i> , ↓ FXRD3 , <i>GNL3</i> , ↑ IGF1 , ↓ MCM10 , <i>MDM2</i> , <i>MDM4</i> , <i>MIR24 (human)</i> , ↑ NAV2 , ↑ NEFM , <i>NEUROD1</i> , <i>NOTCH1</i> , <i>NOV</i> , <i>NR4A1</i> , ↓ PAICS , <i>PIAS2</i> , <i>PRKRA</i> , <i>RB1</i> , <i>RPL11</i> , <i>RRM2B</i> , ↓ SIX1 , ↑ TBX2 , ↓ TFDP1	17	13	Ciclo Celular, Doença Auditiva, Doença Neurológica
11	<i>ACVR2A</i> , ↓ BMP7 , ↓ C10orf54 , ↑ CDKN2D , <i>CSH1/CSH2</i> , ↑ E2F5 , <i>ENG</i> , ↑ EXPH5 , <i>FCRL6</i> , <i>FSTL3</i> , <i>GAD2</i> , ↑ GALNT3 , ↓ GALNT14 , <i>HLA-DR</i> , <i>HLA-DRA</i> , <i>Hla-Drb</i> , <i>HLA-DRB3</i> , ↓ HLA-DRB4 , <i>HLA-DRB5</i> , <i>MHC II-β</i> , <i>MUC2</i> , <i>MUC7</i> , <i>MUC5AC/MUC5B</i> , <i>MYC</i> , ↑ ODC1 , <i>polypeptide N-acetylgalactosaminyltransferase</i> , ↑ PYGL , <i>RFX1</i> , ↓ RFX2 , <i>RFX4</i> , <i>SMAD3</i> , <i>SMAD9</i> , ↓ SVEP1 , <i>UNC45A</i>	17	12	Doença Genética, Doença Neurológica, Morfologia Tumoral
12	↓ ARID5A , ↓ ATXN1 , <i>ATXN2</i> , ↓ C11orf41 , <i>DTX2</i> , <i>DVL2</i> , ↓ EFEMP2 , ↓ ENO1 , ↑ GAPDH , ↓ GPRIN1 , ↓ HOXC13 , <i>ITGA4</i> , ↑ KIF23 , <i>LMNA</i> , ↓ NONO , <i>PRKD1</i> , <i>PRRC2A</i> , ↓ PTBP1 , <i>PXN</i> , ↓ RBPMS , <i>REPS2</i> , <i>RHOXF2/RHOXF2B</i> , ↓ RNF128 , ↑ RPL19 , <i>SSBP1</i> , ↑ STRBP , <i>TOLLIP</i> , <i>TUBA1A</i> , ↓ UBAP2L , <i>UPF1</i> , <i>USP8</i> , <i>WNK1</i> , <i>YWHAZ</i> , <i>ZBTB32</i> , ↓ ZFP36L1	13	10	Morte Celular, Doença Inflamatória, Morfologia Celular
13	↓ ABCC9 , <i>Antigenic lipid-CD1A-β2M</i> , <i>ARHGAP17</i> , <i>B2M</i> , <i>Cd1</i> , <i>CD8</i> , <i>CD46</i> , <i>CD55</i> , <i>CD72</i> , ↓ CD1A , <i>CD3G</i> , <i>CEACAM1 (includes others)</i> , ↑ COLEC12 , <i>DLGAP3</i> , <i>DLGAP4</i> , ↓ FAM110A , <i>FCGR2C</i> , ↓ GPX2 , <i>GRB2</i> , <i>HNFB1A</i> , <i>KIR2DL3 (includes others)</i> , <i>KIR3DL1</i> , ↓ LCK , ↑ MPZL1 , <i>PLXNB1</i> , <i>PTPN1</i> , ↓ PTPN22 , ↓ PTPRJ , <i>Self lipid-CD1A-β2M</i> , ↑ SEMA4D , ↑ SKAP2 , ↑ SOX4 , <i>SRC</i> , ↓ ST5 , <i>TRAT1</i>	12	10	Sinalização e Interação Célula-Célula, Função e Desenvolvimento do Sistema Imunológico, Tráfego de Células Imune

Setas ao lado esquerdo do gene em vermelho indicam genes com expressão aumentada; Setas ao lado esquerdo em verde indicam os genes com expressão diminuída. **Em negrito**, genes identificadas em nosso estudo e com *fold-change* ≤ -1.5 e ≥ 1.5.

A



B

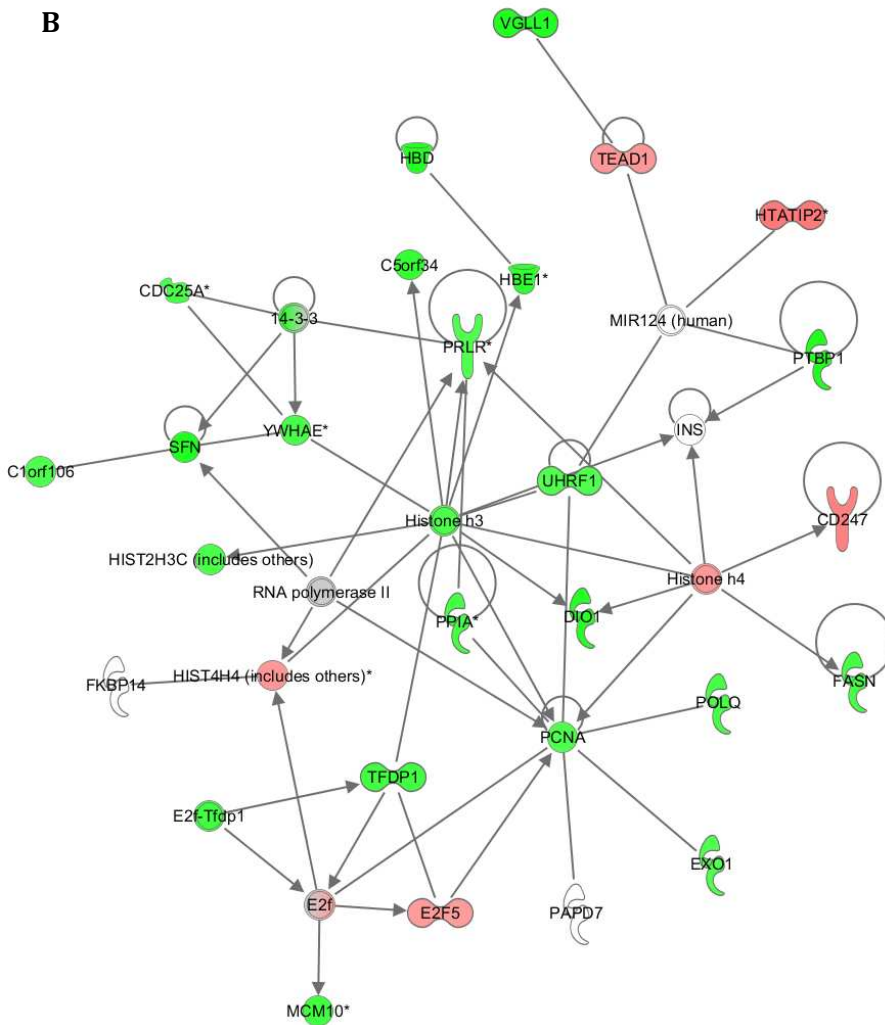


Figura 24. Redes geradas a partir das interações entre os genes diferencialmente expressos e o banco de dados IPA. **A - Rede 1:** genes associados com ciclo celular, câncer e doença do sistema reprodutivo; **B - Rede 2:** genes envolvidos com ciclo celular, doença hematológica e metabolismo de carboidratos. Em verde, genes com expressão diminuída e em vermelho, genes com expressão aumentada (IPA).

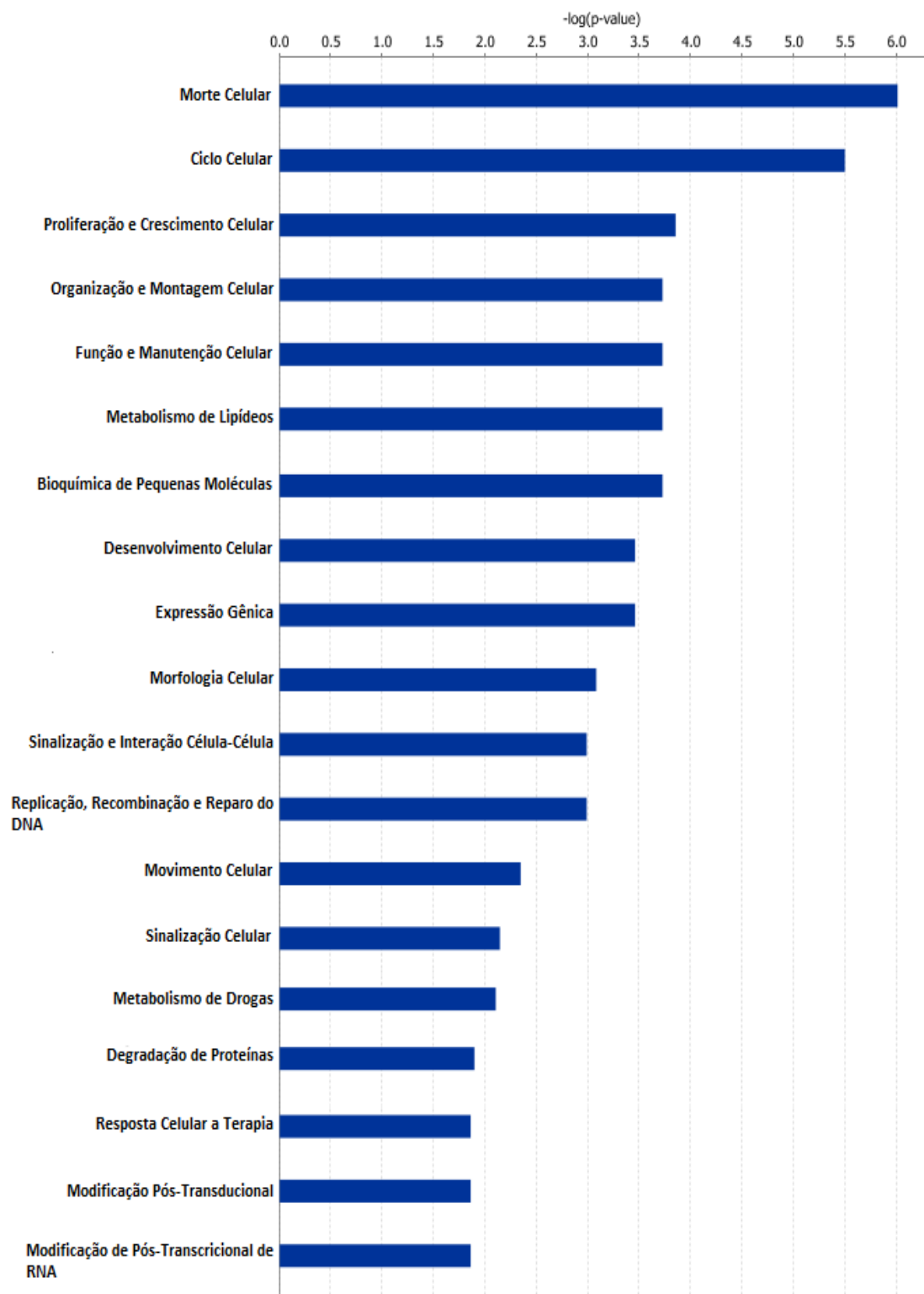


Figura 25. As funções moleculares exercidas pelos genes diferencialmente expressos, em valores de $-\log(P$ valor). Observam-se as funções associadas com ciclo celular e proliferação celular com maiores valores (IPA).

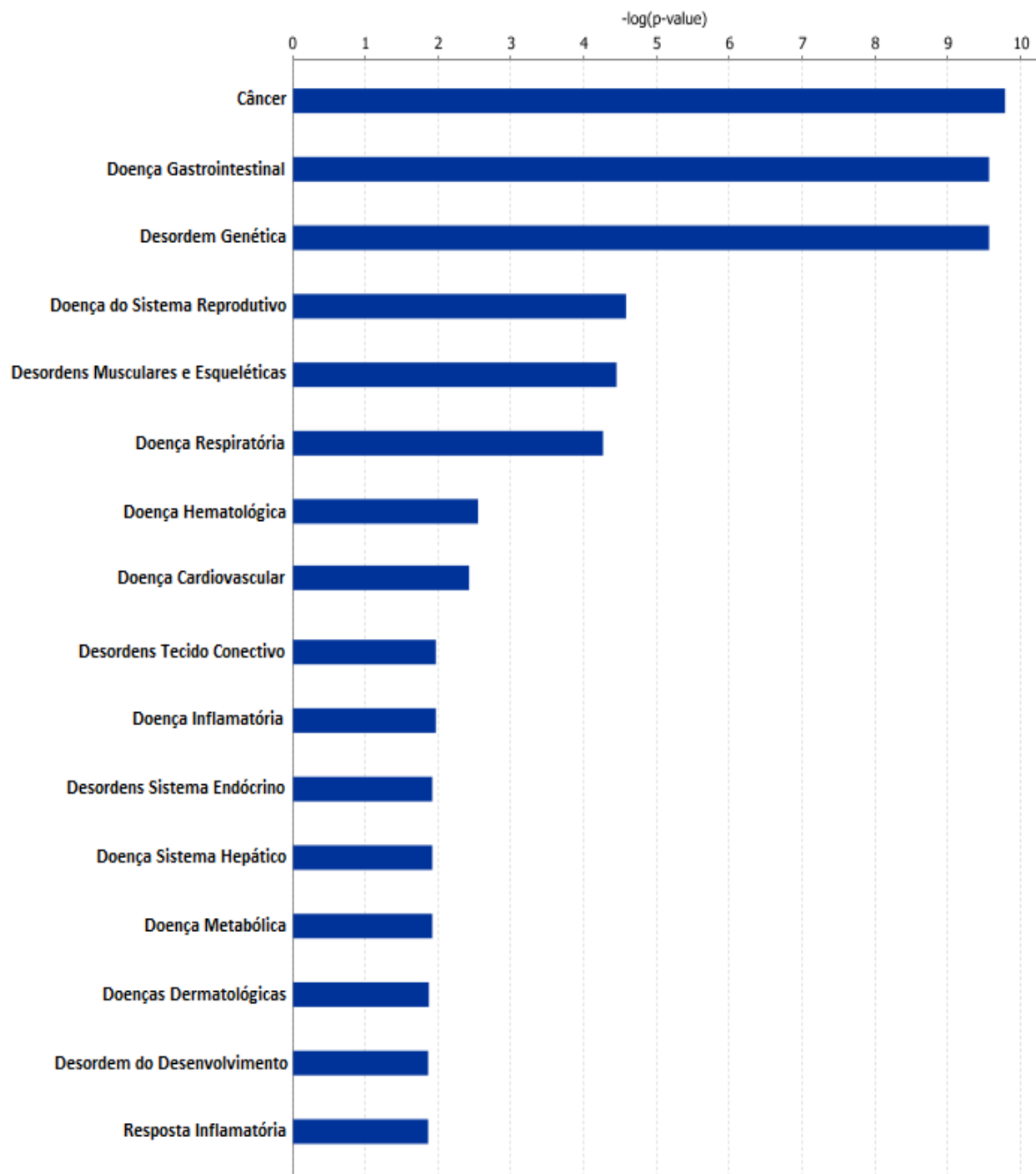


Figura 26. Doenças associadas aos genes diferencialmente expressos, em valores de $-\log(P$ valor). Câncer, doença gastrointestinal e doença genética aparecem com valores próximos ao P valor máximo observado, seguidas por doença do sistema reprodutivo e doenças musculares e esqueléticas (IPA).

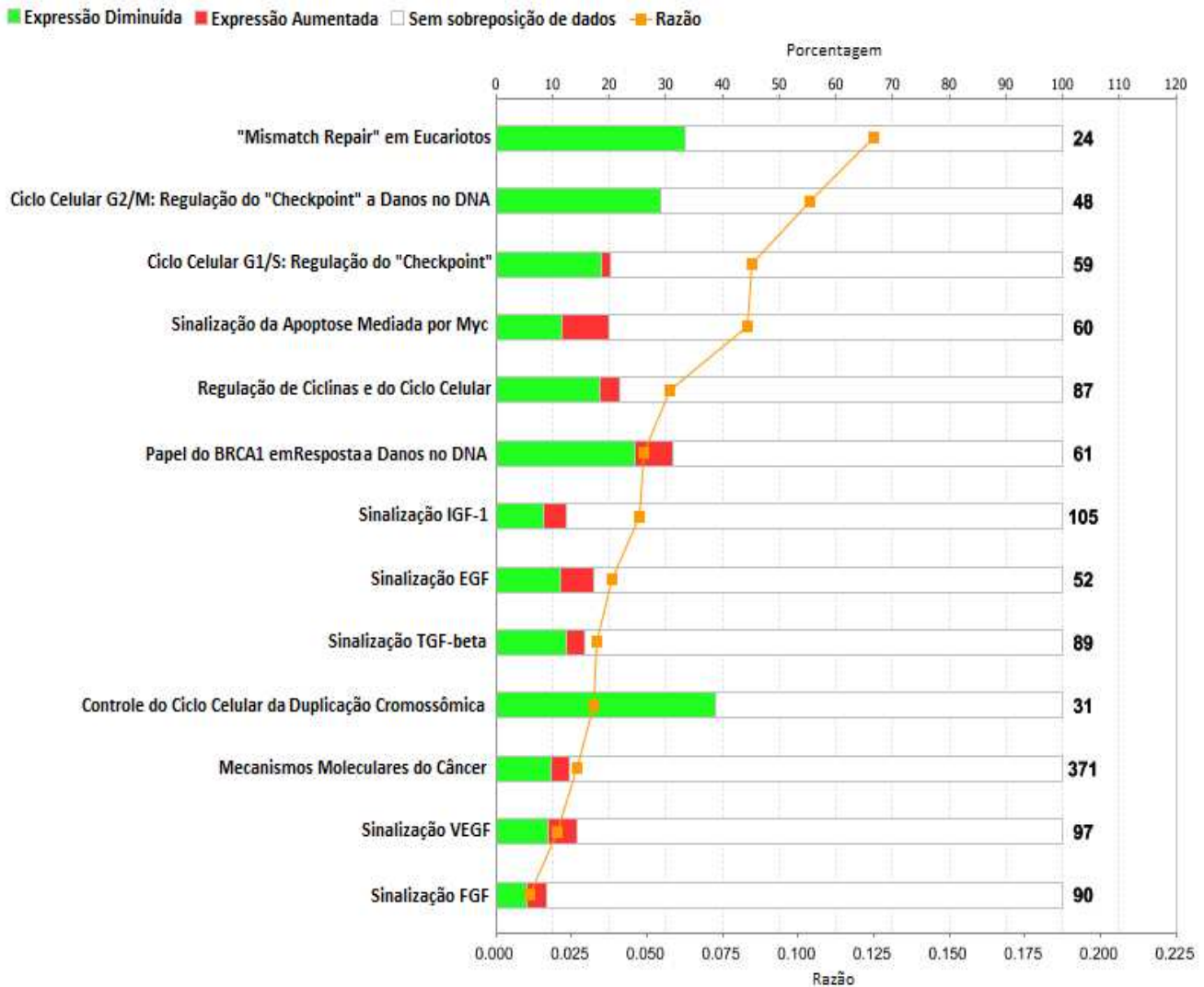


Figura 27. As vias canônicas associadas aos genes diferencialmente expressos estão descritas como valores de razão em relação à frequência das moléculas presentes no banco de dados. As principais vias desempenham funções associadas com reparo a erros de pareamento (*Mismatch Repair*) em eucariotos, regulação do *checkpoint* G2/M do ciclo celular em resposta a danos no DNA, regulação do *checkpoint* G1/S do ciclo celular, sinalização da apoptose mediada por MYC e regulação do ciclo celular e ciclinas (IPA).

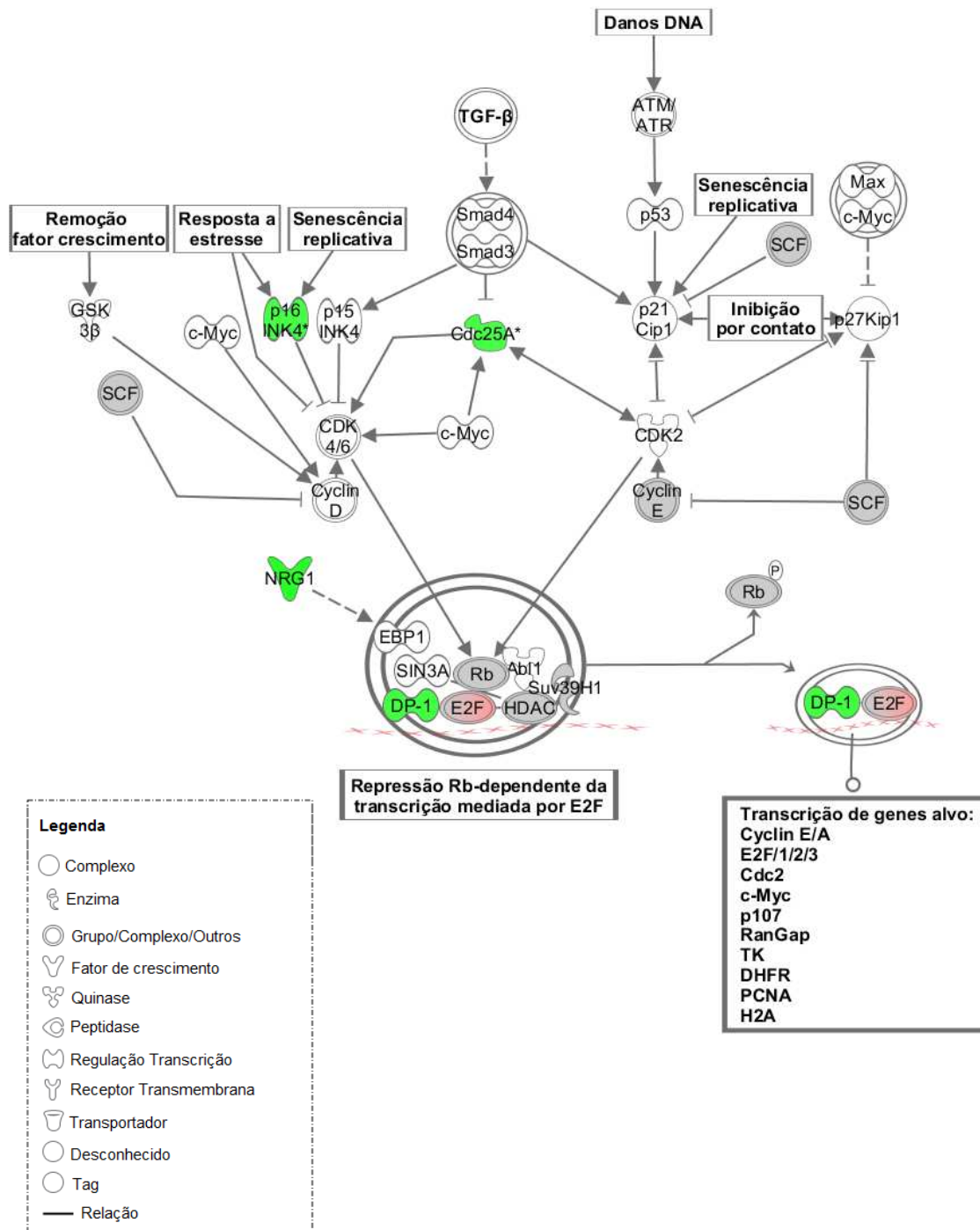


Figura 28. Via da regulação do *checkpoint* do ciclo celular em G1/S. A estimulação pelo TGF-beta, em resposta a danos no DNA, ativa a cascata de sinalização de moléculas envolvidas no *checkpoint* do ciclo celular e na transcrição de genes alvos. Em verde, moléculas codificadas por genes com expressão diminuída e em vermelho, moléculas codificadas por genes com expressão aumentada (IPA).

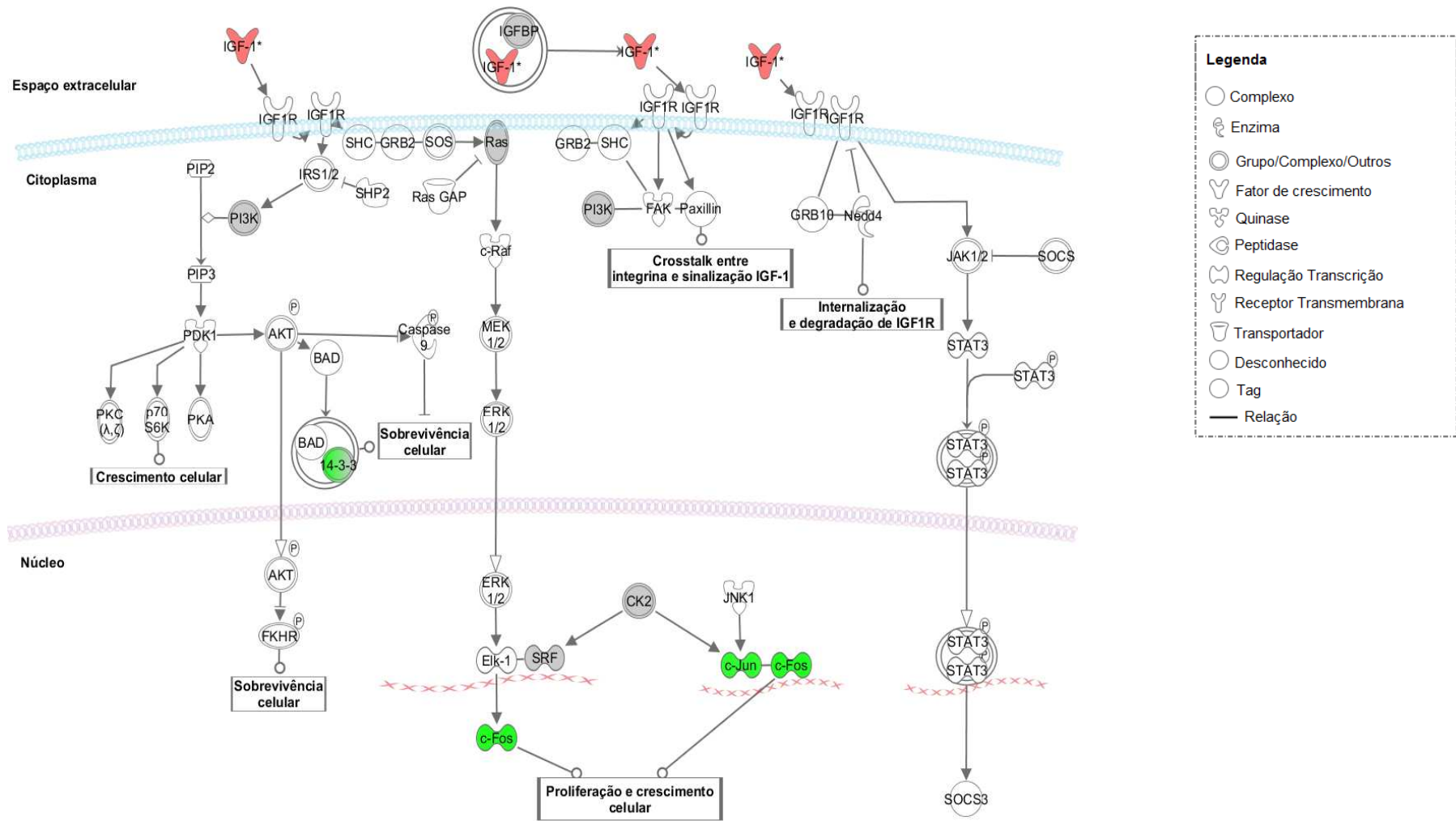


Figura 29. Via Sinalização IGF-1. O IGF-1, que atua como fator de crescimento, ativa a cascata de sinalização Ras, PI3K e JAK/STAT, que induz a transcrição de genes envolvidos com proliferação e crescimento celular. Em verde, moléculas codificadas por genes com expressão diminuída e em vermelho, moléculas codificadas por genes com expressão aumentada (IPA).

4.5 Análise integrada dos dados genômicos e transcriptômicos

Os resultados da análise integrada foram obtidos a partir da comparação entre os dados genômicos e transcriptômicos em 51 amostras de Leiomiomas Uterinos. Como apresentado na seção Materiais e Métodos, devido à abordagem utilizada para a integração dos dados, foi realizada uma nova análise dos dados de CGH array nos 51 casos pareados para as duas análises. As etapas da obtenção dos resultados de CGH array e microarray de expressão estão representadas na Figura 30.

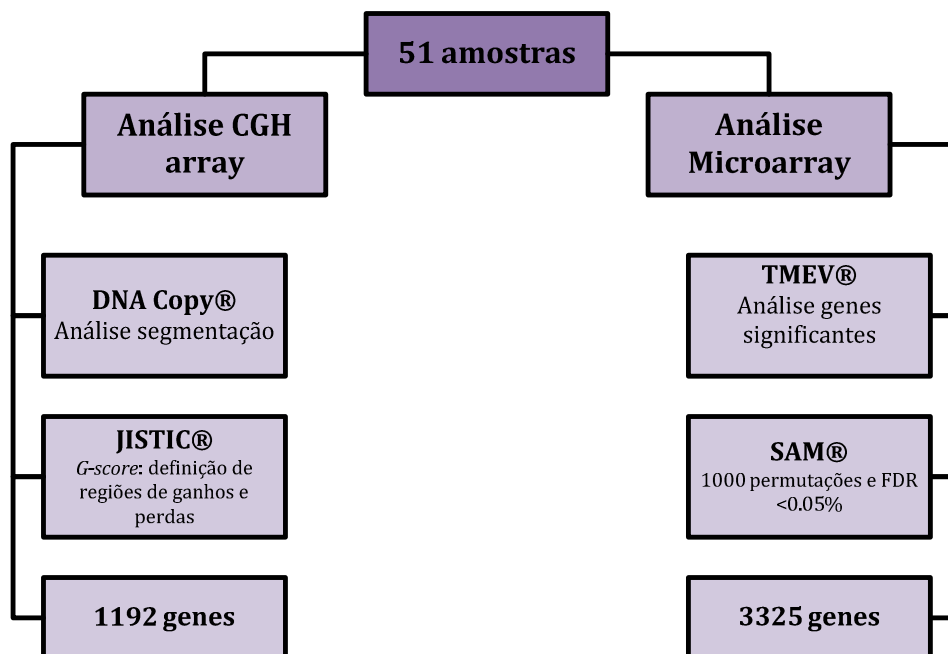


Figura 30. Etapas da análise dos dados de CGH array (genômicos) e de microarray de expressão (transcriptômicos) para obtenção dos genes significantes.

4.5.1 CGH array

A análise dos dados genômicos foi dividida em duas etapas: a segmentação e a obtenção das regiões com alterações significativas. A primeira envolveu a identificação de alterações no número de cópias ou de pontos de quebra. Para isso foi utilizado o método não-paramétrico *DNA Copy* (Olshen et al., 2004). Os perfis genômicos de cada caso foi gerado. A Figura 31 é um exemplo que ilustra o perfil da amostra 759T. A segunda etapa foi a utilização do algoritmo JISTIC (Beroukhim et al., 2007) para as 51 amostras segmentadas. Nesta análise, foi utilizado um limiar de sensibilidade de 0,3 (*q*-valor). Os genes mapeados em regiões significativamente alteradas foram escolhidos como genes candidatos a moduladores. Com esse algoritmo, foi possível identificar 1192 genes mapeados em regiões com *G-score* significativo (picos de alterações) para ganhos e perdas genômicas.

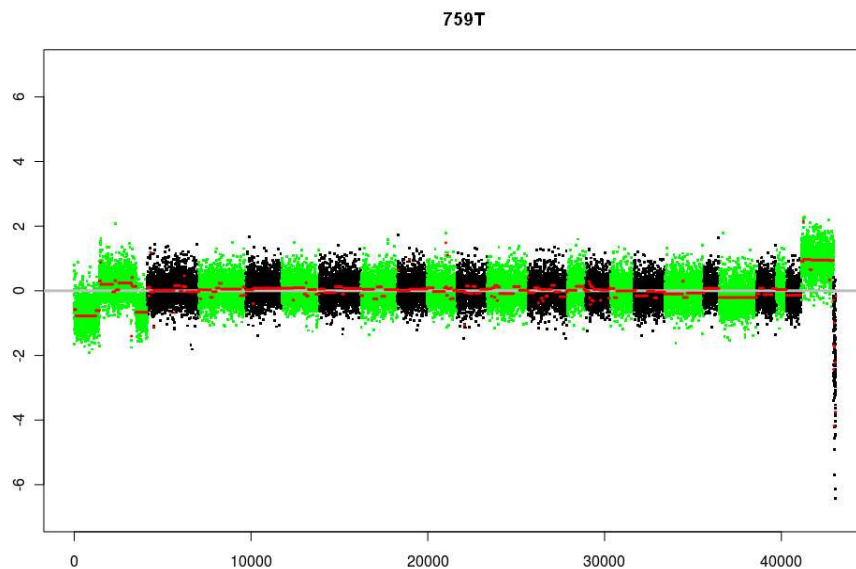


Figura 31. Análise de segmentação realizada pelo *DNA Copy*. No eixo x, os valores de 0 a 40.000 representam o número de marcadores do algoritmo que representam todos os cromossomos. As cores alternativas representam os cromossomos (1 a 22, X e Y), no qual o primeiro grupo de pontos verdes indica o cromossomo 1. No eixo y, os valores de \log_2 mostram as variações abaixo de zero (linha vermelha no eixo central), indicando regiões envolvidas em perdas e as variações acima de zero, indicam as regiões envolvidas em ganhos genômicos. Regiões sem oscilação do eixo central não apresentam alterações no número de cópias de DNA (*DNA Copy*).

4.5.2 Perfil de expressão gênica global

Nesta análise, foram utilizados os 3325 genes significantes, com seus respectivos valores de alteração ($\log_2\text{ratio}$), obtidos na análise prévia de dados transcriptômicos (SAM, 1000 permutações aleatórias e taxa de falsos positivos representado por um FDR <0.05%). Transcritos com valores de $\log_2\text{ratio} \leq -1.0$ e ≥ 1.0 (*fold-change*) foram classificados com diminuição ou aumento de expressão, respectivamente. Aproximadamente 45% dos genes com expressão aumentada e 38% dos genes com expressão diminuída possuem funções associadas a processos celulares. Entre os genes com expressão diminuída, 74,4% deles estão associados com funções relacionadas ao ciclo celular. As funções biológicas dos genes identificados nesta análise foram geradas pelo programa FUNNET (<http://www.funnet.info/>) e estão representadas na Figura 32.

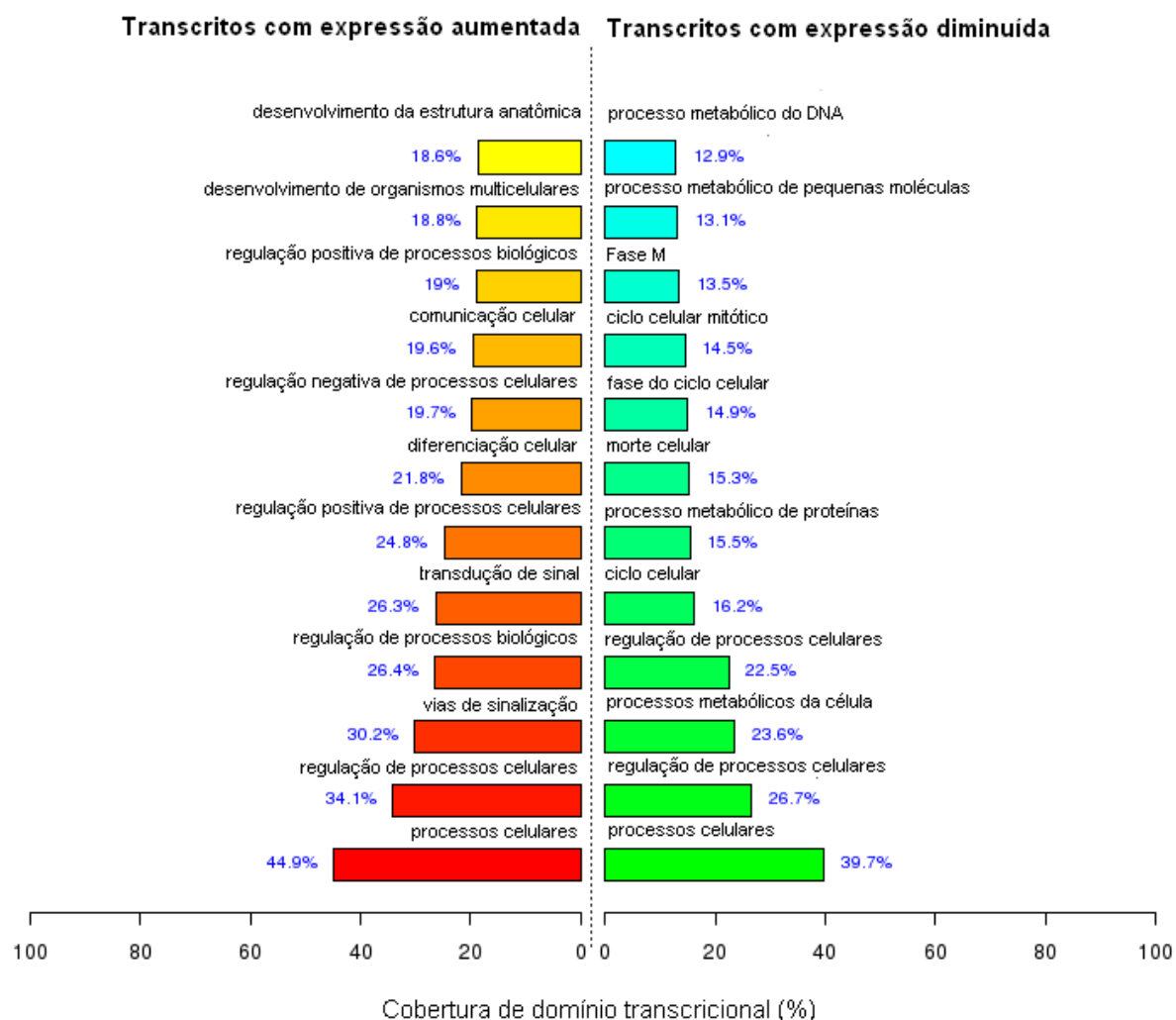


Figura 32. Funções biológicas dos 3325 genes identificados pela análise de microarray. A categorização dos transcritos foi baseada em \log_2ratio com o $fold-change \leq -1.0$ e ≥ 1.0 , indicando diminuição ou aumento de expressão, respectivamente. Quase metade dos genes com expressão aumentada (44,9%) possuem funções associadas com processos celulares e aproximadamente 38% dos genes com expressão diminuída, apresentam esta mesma categorial funcional. 74,4% dos genes com expressão diminuída estão relacionados a funções associadas com ciclo celular (*FUNNET*).

4.5.3 Integração dos dados – CONEXIC

Após a obtenção das regiões genômicas e genes significantes, o algoritmo CONEXIC (Akavia et al., 2010) foi utilizado na integração dos dados. Este algoritmo é baseado na premissa de que genes moduladores (ou *drivers*) são diferencialmente expressos devido a diferentes mecanismos biológicos, como ganhos e perdas genômicas, fornecendo uma vantagem adaptativa para o desenvolvimento do tumor. Para a obtenção destes genes, foram realizadas três etapas fundamentais: 1) seleção dos genes candidatos a moduladores (JISTIC); 2) associação entre os genes candidatos a moduladores e os módulos gênicos (análise de expressão gênica) e 3) utilização de redes de aprendizagem que aumentam o poder de seleção dos genes candidatos. Após estas etapas, foram identificados 75 genes moduladores que explicam o comportamento de expressão dos 3325 genes identificados na análise de expressão gênica. O Anexo 5 mostra a distribuição dos 75 genes identificados separados de acordo com o padrão de ganhos e perdas genômicas e expressão aumentada ou diminuída entre as amostras. As regiões genômicas dos 75 genes moduladores foram anotadas e estavam distribuídos nas seguintes regiões genômicas, 1p36.13, 1q41, 2q32.1, 2q32.2, 2q35, 4p14, 5q31.2, 5q35.3, 7q22.1, 8p12, 8q24.3, 10p15.3, 10q21.3, 11p15.5, 11q13.2, 12p11.21, 12p13.31, 14q13.2, 16p11.2, 16q24.3, 17q21.31, 19q13.32, 19q13.33, 20p11.2 e 20p13. A Tabela 10 mostra os valores do JISTIC e CONEXIC e informações adicionais de cada gene.

Tabela 10. Genes moduladores resultantes da análise integrada dos dados genômicos e transcriptômicos.

q-valor JISTIC	Pontuação CONEXIC	CGH array*	Microarray*	Região	Posição	Símbolo do gene	Nome do gene
<0.01	9832.22	+	+	1p36.13	16.213.110-16.217.872	<i>HSPB7</i>	<i>heat shock 27kDa protein family, member 7</i>
<0.01	2280.2	-	-	1q41	212,843,155-212,904,537	<i>CENPF</i>	<i>centromere protein F</i>
<0.01	3214.2	+	+	2q32.1	187,916,094-188,021,266	<i>CALCLL</i>	<i>calcitonin receptor-like</i>
<0.01	3211.1	+	+	2q32.2	189,547,344-189,585,717	<i>COL3A1</i>	<i>collagen, type III, alpha 1</i>
<0.01	2576.68	+	+	2q35	217,245,073-217,268,517	<i>IGFBP5</i>	<i>insulin-like growth factor binding protein 5</i>
<0.01	15242.9	+	+		218,372,757-218,517,041	<i>TNS1</i>	<i>tensin 1</i>
<0.01	1774	+	+		218,972,722-218,978,908	<i>CTDSP1</i>	<i>CTD (carboxy-terminal domain, RNA polymerase II, polypeptide A) small phosphatase 1</i>
<0.01	3222	+	-		218,996,724-219,022,487	<i>VIL1</i>	<i>villin 1</i>
<0.01	8213.98	+	-		219,141,922-219,167,243	<i>RQCD1</i>	<i>RCD1 required for cell differentiation1 homolog</i>
<0.01	3009.88	-	-	4p14	39,874,922-39,922,676	<i>RHOH</i>	<i>ras homolog gene family, member H</i>
<10 ⁻⁹	21227.32	+	-	5q31.2	137,543,248-137,551,261	<i>KIF20A</i>	<i>kinesin family member 20A</i>
<0.01	2487.12	+	-		137,648,858-137,695,415	<i>CDC25C</i>	<i>cell division cycle 25 homolog C</i>
<0.01	2290.81	+	-	5q35.3	176,663,441-176,666,556	<i>PRELID1</i>	<i>PRELI domain containing 1</i>
<0.01	2001.89	+	-		176,761,745-176,769,183	<i>F12</i>	<i>coagulation factor XII (Hageman factor)</i>
<0.01	3678.31	+	+		176,816,220-176,833,300	<i>DBN1</i>	<i>drebrin 1</i>
<0.01	3222.8	+	-	7q22.1	98,874,499-98,892,932	<i>CPSF4</i>	<i>cleavage and polyadenylation specific factor 4</i>
<0.01	4222.7	+	-		98,893,720-98,901,744	<i>ATP5J2</i>	<i>ATP synthase, H+ transporting, mitochondrial Fo complex, subunit F2</i>
<0.01	1908.18	+	-		98,994,384-99,012,013	<i>ZNF655</i>	<i>zinc finger protein 655</i>
<0.01	1232.98	+	-		99,402,286-99,411,623	<i>AZGP1</i>	<i>alpha-2-glycoprotein 1, zinc-binding</i>
<0.01	442.98	+	-		99,528,340-99,537,363	<i>MCM7</i>	<i>minichromosome maintenance complex component 7</i>

0.05	1132.32	+	-		99,613,474-99,649,946	<i>STAG3</i>	<i>stromal antigen 3</i>
<0.01	1883.3	+	-	8q24.3	144,520,168-144,522,180	<i>C8orf51</i>	<i>chromosome 8 open reading frame 51</i>
<0.01	998.3	+	-		144,757,532-144,762,887	<i>PYCR1</i>	<i>pyrroline-5-carboxylate reductase-like</i>
<0.01	774.1	+	-		145,221,948-145,224,416	<i>CYC1</i>	<i>cytochrome c-1</i>
<0.01	523.2	+	+		145,619,808-145,624,735	<i>VPS28</i>	<i>vacuolar protein sorting 28 homolog (S. cerevisiae)</i>
<0.01	1124.2	+	-		145,624,971-145,640,620	<i>TONSL</i>	<i>tonsoku-like, DNA repair protein</i>
<0.01	523.65	+	-		145,707,479-145,714,008	<i>RECQL4</i>	<i>RecQ protein-like 4</i>
<0.01	235.65	+	-		145,714,199-145,721,365	<i>LRRC14</i>	<i>leucine rich repeat containing 14</i>
<0.01	998.4	+	+	8p12	37,773,932-37,820,650	<i>GPR124</i>	<i>G protein-coupled receptor 124</i>
<0.01	1534.34	+	-		38,007,177-38,037,040	<i>EIF4EBP1</i>	<i>eukaryotic translation initiation factor 4E binding protein 1</i>
<0.01	1552.55	+	+		38,387,813-38,445,509	<i>FGFR1</i>	<i>fibroblast growth factor receptor 1</i>
<0.01	1523.52	+	-	10p15.3	1,075,964-1,085,061	<i>IDI1</i>	<i>isopentenyl-diphosphate delta isomerase 1</i>
<0.01	1774.98	+	+		311,432-725,606	<i>DIP2C</i>	<i>disco-interacting protein 2 homolog C</i>
<0.01	1883.3	+	-	10q21.3	69,539,196-69,641,779	<i>MYPN</i>	<i>myopalladin</i>
0.01	998.3	+	-		70,385,898-70,414,285	<i>DDX21</i>	<i>DEAD (Asp-Glu-Ala-Asp) box polypeptide 21</i>
<0.01	774.1	+	-		70,650,065-70,697,321	<i>HKDC1</i>	<i>hexokinase domain containing 1</i>
<0.01	523.2	+	+	11p15.5	303,991-305,272	<i>IFITM1</i>	<i>interferon induced transmembrane protein 1 (9-27)</i>
<0.01	1124.2	+	-		384,217-394,908	<i>PKP3</i>	<i>plakophilin 3</i>
0.02	523.65	+	-		522,242-525,550	<i>HRAS</i>	<i>v-Ha-ras Harvey rat sarcoma viral oncogene homolog</i>
<0.01	235.65	+	-		551,450-554,018	<i>RASSF7</i>	<i>Ras association (RalGDS/AF-6) domain family (N-terminal) member 7</i>
<0.01	998.4	+	-		737,432-755,024	<i>TALDO1</i>	<i>transaldolase 1</i>
<0.01	1534.34	+	-		780,475-786,221	<i>SLC25A22</i>	<i>solute carrier family 25 (mitochondrial carrier: glutamate), member 22</i>
<0.01	1552.55	+	-	11q13.2	67,576,902-67,645,434	<i>CHKA</i>	<i>choline kinase alpha</i>
<0.01	1523.52	-	-	12p11.21	32,151,452-32,422,408	<i>BICD1</i>	<i>bicaudal D homolog 1</i>

<0.01	1774.98	+	-	12p13.31	8,646,029-8,656,706	<i>AICDA</i>	<i>activation-induced cytidine deaminase</i>
<0.01	2001.32	+	+		8,689,807-8,706,700	<i>MFAP5</i>	<i>microfibrillar associated protein 5</i>
<0.01	1232	+	-	14q13.2	34,291,688-34,414,604	<i>BAZ1A</i>	<i>bromodomain adjacent to zinc finger domain, 1A</i>
<0.01	1822.11	+	+	16p11.2	28,456,163-28,457,996	<i>NUPR1</i>	<i>nuclear protein, transcriptional regulator, 1</i>
<0.01	909.81	+	-		28,850,761-28,858,164	<i>CD19</i>	<i>CD19 molecule</i>
<0.01	987.05	+	+		29,730,910-29,734,703	<i>PRRT2</i>	<i>proline-rich transmembrane protein 2</i>
<0.01	1034.55	+	+		29,739,288-29,766,842	<i>MVP</i>	<i>major vault protein</i>
<0.01	1993.31	+	-		30,023,632-30,032,379	<i>GDPD3</i>	<i>glycerophosphodiester phosphodiesterase domain containing 3</i>
<0.01	1738.2	+	-		30,102,427-30,107,898	<i>CORO1A</i>	<i>coronin, actin binding protein, 1A</i>
<0.01	783.5	+	+		30,488,524-30,491,229	<i>ZNF688</i>	<i>zinc finger protein 688</i>
<0.01	984.2	+	+	16q22.1	65,775,770-65,781,608	<i>EXOC3L1</i>	<i>exocyst complex component 3-like 1</i>
0.05	578.87	+	-		65,790,529-65,795,428	<i>ELMO3</i>	<i>engulfment and cell motility 3</i>
0.01	1378.55	+	+	16q24.3	88,301,042-88,314,895	<i>C16orf7</i>	<i>chromosome 16 open reading frame 7</i>
<0.01	566.98	+	-		88,331,460-88,410,566	<i>FANCA</i>	<i>Fanconi anemia, complementation group A</i>
<0.01	665.35	+	-		88,517,246-88,530,006	<i>TUBB3</i>	<i>tubulin, beta 3</i>
<0.01	533.8	+	-	17q21.31	38,449,840-38,530,994	<i>BRCA1</i>	<i>breast cancer 1, early onset</i>
<0.01	968.66	+	+		38256727-38263666	<i>AOC3</i>	<i>amine oxidase, copper containing 3 (vascular adhesion protein 1)</i>
<0.01	1372.76	+	+	19q13.32	50,702,528-50,722,080	<i>VASP</i>	<i>vasodilator-stimulated phosphoprotein</i>
<0.01	1799.91	+	+		50,784,865-50,787,557	<i>GPR4</i>	<i>G protein-coupled receptor 4</i>
0.01	879.54	+	-		50,882,558-50,887,282	<i>SNRPD2</i>	<i>small nuclear ribonucleoprotein D2 polypeptide</i>
<0.01	876.1	+	+		50,964,816-50,977,655	<i>DMPK</i>	<i>dystrophia myotonica-protein kinase</i>
<0.01	866.98	+	+		51,134,631-51,168,497	<i>NOVA2</i>	<i>neuro-oncological ventral antigen 2</i>
<0.01	655.5	+	+		51,605,929-51,608,681	<i>CCDC8</i>	<i>coiled-coil domain containing 8</i>
<0.01	675.01	+	-		51,796,352-51,805,879	<i>CALM3</i>	<i>calmodulin 3 (phosphorylase kinase, delta)</i>
<0.01	989.32	+	+		51,842,709-51,856,235	<i>DACT3</i>	<i>dapper, antagonist of beta-catenin, homolog 3 (Xenopus laevis)</i>
<0.01	377.3	+	-		51,969,980-51,983,653	<i>SLC1A5</i>	<i>solute carrier family 1 (neutral amino acid transporter), member</i>

							5
<0.01	656.33	+	-		52,033,263-52,046,043	<i>AP2S1</i>	<i>adaptor-related protein complex 2, sigma 1 subunit</i>
<0.01	886.44	+	-	19q13.33	55,614,028-55,624,060	<i>SPIB</i>	<i>Spi-B transcription factor (Spi-1/PU.1 related)</i>
<0.01	1252.32	-	+	20p11.2	23,007,993-23,014,977	<i>CD93</i>	<i>CD93 molecule</i>
0.17	0	+	-	20p13	309,308-326,203	<i>TRIB3</i>	<i>tribbles homolog 3</i>
0.23	0	+	-		364,124-391,187	<i>TBC1D20</i>	<i>TBC1 domain family, member 20</i>

*Os sinais positivos (+) e negativos (-) representam o *status* dos genes em relação ao padrão de ganhos (+) e perdas (-) e ao aumento (+) e diminuição (-) de expressão gênica. JISTIC = análise de segmentação dos dados de CGH array; CONEXIC = classificação do algoritmo para os genes moduladores.

4.5.3.1 Predição de regulação gênica por miRNA

Entre os 75 genes moduladores identificados, 26 apresentaram correlação positiva (ganho genômico/aumento expressão) e três apresentaram correlação negativa (perda genômica/diminuição expressão) e 46 deles apresentaram correlação inversa (45 apresentaram ganho genômico e expressão diminuída e um gene teve perda genômica e expressão aumentada). Sendo assim, foi realizada uma análise de predição de regulação gênica por miRNA com o objetivo de verificar se o comportamento de correlação inversa poderia ser explicado por este mecanismo. Foi realizada uma busca nos bancos de dados *TargetScan* (<http://www.targetscan.org/>) e *PicTar* (http://pictar.mdc-berlin.de/cgi-bin/PicTar_vertbrate.cgi) que são ferramentas de predição de regulação gênica por miRNAs, para os 46 genes que apresentaram correlação inversa (Tabela 11). Os genes *ATP5J2*, *AZGP1*, *CD19*, *CORO1A*, *CYC1*, *ELMO3*, *GDPD3*, *NFKBIL2*, *PKP3*, *PYCRL*, *RASSF7*, *RECQL4*, *SNRPD2*, *STAG3* e *TALDO1* não apresentaram predição por miRNA. Os genes *FANCA* e *TBC1D20*, que tiveram expressão gênica diminuída, são regulados pelos miRNAs *hsa-miR-26b* e *hsa-miR-150*, ambos identificados em regiões envolvidas em ganhos genômicos neste estudo (2q35 e 19q13.33, respectivamente).

Tabela 11. Genes identificados na análise integrada baseada nos achados de correlação inversa (ganho genômico/diminuição de expressão e perda genômica/ aumento de expressão gênica) e seus miRNAs preditos.

Símbolo gene	TargetScan	PicTar
<i>AICDA</i>	<i>hsa-miR-155</i>	<i>hsa-miR-155</i>
<i>AP2S1</i>	-	<i>hsa-miR-34b, hsa-miR-34c, hsa-miR-34a</i>
<i>ATP5J2</i>	-	-
<i>AZGP1</i>	-	-
<i>BAZ1A</i>	<i>hsa-miR-137</i>	-
<i>BRCA1</i>	<i>hsa-miR-218</i>	<i>hsa-miR-197, hsa-miR-143, hsa-miR-205, hsa-miR-132, hsa-miR-370, hsa-miR-30a-3p, hsa-miR-140, hsa-miR-185, hsa-miR-154</i>
<i>C8orf51</i>	ND	ND
<i>CALM3</i>	<i>hsa-miR-22</i>	<i>hsa-miR-122a, hsa-miR-22, hsa-miR-27b, hsa-miR-27a, hsa-miR-196b, hsa-miR-196a, hsa-miR-29c, hsa-miR-29b, hsa-miR-29a, hsa-miR-320,</i>
<i>CD19</i>	-	-
<i>CDC25C</i>	<i>hsa-miR-767-3p</i>	-
<i>CD93</i>	<i>hsa-miR-216a</i>	-
<i>CHKA</i>	<i>hsa-miR-30e, hsa-miR-30a, hsa-miR-30d, hsa-miR-30b, hsa-miR-30c</i>	-
<i>CORO1A</i>	-	-
<i>CPSF4</i>	<i>hsa-miR-214</i>	<i>hsa-miR-214, hsa-miR-23a, hsa-miR-23b, hsa-let-7c, hsa-let-7g, hsa-let-7b, hsa-let-7f, hsa-let-7i, hsa-let-7a, hsa-miR-98, hsa-let-7e, hsa-let-7d</i>
<i>CYC1</i>	-	-
<i>DDX21</i>	<i>hsa-miR-607</i>	-
<i>EIF4EBP1</i>	<i>hsa-miR-125b, hsa-miR-125a-5p</i>	<i>hsa-miR-125b, hsa-miR-125a</i>
<i>ELMO3</i>	-	-
<i>F12</i>	<i>hsa-miR-330-3p</i>	-
<i>FANCA</i>	-	<i>hsa-miR-26a, hsa-miR-26b</i>
<i>GDPD3</i>	-	-
<i>HKDC1</i>	<i>hsa-miR-876-5p, hsa-miR-1243</i>	-
<i>HRAS</i>	<i>hsa-miR-892a</i>	-
<i>IDI1</i>	<i>hsa-miR-570</i>	-
<i>KIF20A</i>	<i>hsa-miR-153</i>	<i>hsa-miR-369, hsa-miR-374</i>
<i>LRRC14</i>	-	<i>hsa-miR-370, hsa-miR-28, hsa-miR-122a</i>
<i>MCM7</i>	<i>hsa-miR-519b-3p, hsa-miR-519a, hsa-miR-519c-3p, hsa-miR-548p</i>	-
<i>MYPN</i>	<i>hsa-miR-214</i>	-
<i>NFKBIL2</i>	-	-

<i>PKP3</i>	-	-
<i>PRELID1</i>	<i>hsa-miR-22</i>	-
<i>PYCR1</i>	-	-
<i>RASSF7</i>	-	-
<i>RECQL4</i>	-	-
<i>RQCD1</i>	ND	ND
<i>SLC1A5</i>	<i>hsa-miR-125a-5p, hsa-miR-125b</i>	<i>hsa-miR-122a, hsa-miR-137</i>
<i>SLC25A22</i>	<i>hsa-miR-613, hsa-miR-1, hsa-miR-206</i>	<i>hsa-miR-296, hsa-miR-1, hsa-miR-206, hsa-miR-337</i>
<i>SNRPD2</i>	-	-
<i>SPIB</i>	<i>hsa-miR-1299, hsa-miR-520a-5p, hsa-miR-525-5p</i>	<i>hsa-miR-146, hsa-miR-218, hsa-miR-328</i>
<i>STAG3</i>	-	-
<i>TALDO1</i>	-	-
<i>TBC1D20</i>	<i>hsa-miR-150</i>	-
<i>TRIB3</i>	<i>hsa-miR-24</i>	-
<i>TUBB3</i>	<i>hsa-miR-429, hsa-miR-200c, hsa-miR-200b</i>	-
<i>VIL1</i>	-	-
<i>ZNF655</i>	<i>hsa-miR-181c, hsa-miR-181a, hsa-miR-181b, hsa-miR-181d</i>	-

Em negrito, miRNAs mapeados em regiões identificadas como alteradas por CGH array (*DNA Copy*); *TargetScan* (<http://www.targetscan.org/>); *PicTar* (http://pictar.mdc-berlin.de/cgi-bin/PicTar_vertbrate.cgi); ND=genes não identificados nos bancos de dados; (-) sem predição de miRNA.

4.5.3.2 miRNAs mapeados em regiões de ganhos e perdas genômicas

Os resultados anteriores revelaram pelo menos dois miRNAs mapeados em regiões de instabilidade genômica. A partir deste achado, os genes codificadores de miRNAs mapeados em regiões de ganhos e perdas genômicas deste estudo foram anotados assim como os genes por eles regulados. Os achados genômicos de alterações nos miRNAs foram comparados com o seu padrão de expressão dos genes que eles regulam. No total foram identificados 21 miRNAs em regiões genômicas alteradas elegíveis para regulação gênica. Estes miRNAs foram importados ao sistema IPA (que utiliza os bancos de dados de predição *TargetScan* e *PicTar*), onde foi gerada uma lista de 6635 genes alvos altamente preditos. Estes genes foram comparados aos 3325 genes identificados por análise de expressão gênica global neste estudo, resultando em 780 genes sobrepostos.

Os resultados foram organizados de acordo com o mapeamento dos miRNAs, número de amostras que apresentaram ganho ou perda dos miRNAs e os genes regulados por estes miRNAs que apresentaram aumento ou diminuição de expressão (*log₂ratio*, cutoff de 1.0) (Tabela 12).

Tabela 12. miRNAs mapeados em regiões de ganhos e perdas genômicas e o padrão de expressão de seus genes preditos.

miRNA	Região	Posição	CGH array (n° casos)	Gene	Microarray*
<i>hsa-mir-26b</i>	2q35	218,975,612-218,975,689	Ganho (6)	<i>FA2H, OVOL2, E2F7, DEPDC1, NUP210, DEPDC1, DEPDC1B, SLC16A6, ABCD3, CDC6, NFE2L3, ZIC5, CTNND2, AP1S3, PMAIP1, FANCA, NUA2, MAP2, DSCR6, MREG, SFXN1, FLVCR1, C11orf41, RAB3IP, GNPAT1, FAM49B, BID, RAB3IP, RFC3, TFAM, C9orf40, AGPAT5, GMDS, GALNT10, DNMT3B, PTER, PCNA, CELSR1, BID, ACSL3, C20orf24, STRBP, C20orf24, ROD1, ETNK1, LARP4, YWHAE, COX5A, NUP50, ALDH5A1, SLC7A6, BLOC1S2, YWHAE, SERBP1, NIP7, BCCIP, ERLIN1, RFK, ZNF655, GSR, CPSF2, RNF138, YWHAE, STRBP</i>	-
			Perda (0)	<i>SMAD1, CTDSP2, LINGO1, USP25, PPP3CB, MRAS, ULK2, ZNF516, RAB31, DLG4, ZNF516, LRRC2, CDH11, PECAM1, GREB1, DNAJB5, DMD, KCNE4, DMD, CD200, CACNB2, PRELP, ADAMTS19, HSD17B6, CILP, SOX17, PGR, MUM1L1, ASPN</i>	+
<i>hsa-mir-1269</i>	4q13.2	66,825,136-66,825,241	Ganho (0)	<i>JPH3, TTF2, TSR1, C16orf74, TSR1</i>	-
			Perda (6)	<i>PLB1</i>	+
<i>hsa-mir-1236</i>	6p21.32	32,032,594-32,032,696	Ganho (4)	<i>CDKN2A, C15orf48, GALNT6, RHOTB3, GFI1, DIO1, PVR, DPF3, SLA, CLPB, WDR3, VAPB</i>	-
			Perda (1)	<i>EPN2, MYCT1, SNED1, PPP1R1A, IGFBP5, ATP8B4</i>	+
<i>hsa-mir-25</i>	7q22.1	99,529,118-99,529,202	Ganho (9)	<i>DDX3Y, SLC7A11, PRDM13, PAX3, SOX11, JAKMIP1, GPR158, CXCL5, CLGN, PRDM15, MAGEC2, TRIP13, MCOLN2, MTHFD2, AURKA, HOXC8, ADM, CSMD3, GCLM, PTPN22, GCLM, ME1, TFAP2A, ZDHHC2, PLEKHB2, INSIG1, DOCK5, RGS17, TCFL5, PFKFB4, NF2, ASPH, TMED5, ASB7, GPR180, NLK, CHKA, ARFGF2, ASB7, QSER1, NCBP1, NPC1, CDCA7L, IDH1, C20orf111, C18orf25, CCT6A</i>	-
			Perda (3)	<i>LHFPL2, WDR81, ARHGEF17, PPP1R12C, FOXN3, VCL, PDCD6IP, ZFYVE21, KIAA1109, ARID5B, GPR124, PLEKHG3, RBPMS2, CACNA1H, NFIA, ACTC1, PCDH7, MITF, NFIX, CDKN1C, SPSB1, GATA6, FAM110B, ANGPTL2, ITPR1, RAB23, REV3L, HOXD10, C18orf1, FZD10, STARD13, JPH2, MSRB3, ADAMTS9, HAND2, ANGPTL2, COL1A2, GPR124, CACNA1C, NEFM, ACTC1, CACNA1C, PPAP2B, ATP2B4, LMO2, SCRG1, SLC24A3,</i>	+
<i>hsa-mir-106b</i>	7q22.1	99,529,551-99,529,633	Ganho (9)	<i>GDA, PLS1, RRM2, ESCO2, RRM2, ATAD2, ATAD2, ATAD2, LAMP3, RNF128, RAD51AP1, RAB11FIP4, C1orf135, GRHL2, RACGAP1, E2F2, FAM40B, MASTL, POLQ, MYO5B, RUNX3, HN1, KIF23, EYA1, CIT, CDCA7, HMGCS1, DPYSL5, CENPO, CKAP2, CKAP2, PRR15, RAB11FIP4, EPB41L5, SC4MOL, OLIG1, NIPA1, KPNA2, FNDC3B, PTPN3, E2F5, CKAP2, EIF2S1, SERPINB8, HAS2, SOLH, SH2D5, UBE3C, CENPO, SS18L1, E2F1, PANK3, DCBLD2, ARSJ, KIAA1598, RBL1, HSPA8, PREX1, PHF19, LIF, STK4, MASTL, HN1, LSM14B, CHAF1A,</i>	-

			Perda (3)	PTPRJ, PFAH1B1, MYO10, GLO1, GCLC, MYO10, C18orf19, PTBP1, NEDD4L, FBXO31, RCCD1, NEDD4L, TRIM37, PFAH1B1, BRMS1L, NUP35, CENPQ, GMFB, ICMT, C5orf30, SIRT7, RRP15, MBOAT2, TIMM17A, CSNK2B C7orf42, PHF1, WFS1, HDAC4, MIER2, CNTNAP1, RABGAP1, ZBTB47, THRA, OSBPL5, DAB2, LAPTM4A, EFN1, KIAA0240, RAB11FIP2, C7orf41, ADAMTSL2, TCF4, PTP4A3, NAV2, CRY2, TRIB2, TACC1, PKD2, FAM102A, ST6GALNAC3, ZRANB1, DIP2C, SPG20, SNX21, VASP, TCF7L1, SNX21, CALD1, SOX4, PKD2, TACC1, NTRK3, LRIG1, MBNL1, OLFML2B, RPS6KA2, TBL1X, RAB11FIP2, OLFML2B, THRA, PGM2L1, HECTD2, ARHGAP1, AKAP13, CALD1, AKAP13, BNC2, ATXN1, ITPKB, CCND2, RGMA, FYCO1, FZD7, IL17RD, TXNIP, SYNE1, PDLIM5, ZBTB47, PDLIM5, ARHGEF3, JAZF1, AKAP13, TNS1, CYBRD1, YPEL2, TACC1, LRRFIP1, KIAA1462, TSHZ3, TP53INP1, RASL11B, PGM2L1, CCND2, TNS1, SYTL4, BNC2, YPEL4, PDLIM5, PTGER3, SH3PXD2A, CALD1, RASSF2, RASD1, PDGFRA, TCEAL7, APCDD1, PPP1R12B, NTRK2, SLC40A1, PLSCR4, GUCY1A3, IGF1, SLC2A4, IGF1, FGL2, EGR2, ADAMTSS5, IL17RD, C20orf103, BNC2, PPP1R12B, CTSK, ATP1A2, ESR1, SULF1, SMOC2	+
<i>hsa-mir-548o</i>	7q22.1	101,833,193-101,833,307	Ganho (9)	CXCL1, IGF2BP3, CEP55, MKI67, EIF1AY, CA8, NEBL, SNRPA1, TMEM33, KLHL7, MYBL1, HSPA4L, ASCL1, CYCS, SLMO2, CACYBP, PTPN2, RPE, DSN1, DERL1, NDFIP2, FAM122B, ZNF286A, PPAT, DDX21, ARL6IP1, ODC1, PSIP1, RER1, ZNF259, UBA52	-
			Perda (4)	USP47, SENP7, SEMA6A, FGFR2, GATM, KCNJ8, LPP, LPP, PCDH17, CYYR1, SULF2	+
<i>hsa-mir-661</i>	8q24.3	145,091,346-145,091,435	Ganho (5)	CDH1, EPN3, C9orf140, NCAPH, SCD, HOXB13, PCK2, CSAG1, ZNF587, SCD, ITPR3, PCK2, POLR2D, C19orf48, ABHD10, POLR2D, TNFRSF10B, MRPL11, TSN	-
			Perda (2)	NLGN2, TPPP3, WT1, LRRC32	+
<i>hsa-mir-939</i>	8q24.3	145,590,171-145,590,253	Ganho (6)	GPRC5A, TCF7	-
			Perda (2)	CSPG4, KCNMB1, MRV11, DACT3	+
<i>hsa-mir-1254</i>	10q21.3	70,189,080-70,189,177	Ganho (6)	FOXD1, MPP6, RNF125, POLR1B, FTSJ2,	-
			Perda (1)	PBXIP1	+
<i>hsa-mir-210</i>	11p15.5	558,088-558,198	Ganho (10)	ZNF462, SCARA3	+
			Perda (1)		
<i>hsa-mir-1237</i>	11q13.1	63,892,649-63,892,751	Ganho (5)	KIF11, DHCR24, NQO1, KIF11, DGAT2, C10orf47, CDCP1, C14orf142, C12orf32, C20orf112, CRELD2	-
			Perda (0)	HLA-C, ZFP36L1, NFATC1, ZBTB16, EDNRA, THSD4, BCL6B, MYOC, GPR20, EDNRA, SFRP4	+
<i>hsa-mir-192</i>	11q13.1	64,415,184-64,415,294	Ganho (4)	TYMS, IKZF2, C9orf100, CDC7, L2HGDH, RFWD3, SIX4, SEMA4D, C10orf90, ARHGAP19	-
			Perda (0)	CLSTN1, COL5A1, CDON, WSCD2, FABP3	+
<i>hsa-mir-1228</i>	12q13.3	55,874,553-55,874,626	Ganho (3)	MAL2, DARS2, BMP7, BMP7, EXOSC3, RABGEF1	-

<i>hsa-mir-616</i>	16q13	56,199,212-56,199,309	Perda (1)	<i>NFE2L1, NFE2L1, TNFAIP8L3, TNFAIP8L3, F13A1</i>	+
			Ganho (3)	<i>SERPINB5, SCGB1D2, GPR126, GPR126, CHAC2, TXK, ERO1L, HSD17B7, NUDT21, HSD17B7, ADPGK</i>	-
<i>hsa-mir-328</i>	16q22.1	65,793,724-65,793,799	Perda (1)	<i>KLHL4, THBS2, THSD7A, PDE8B, RGS5, STON1, RGS5</i>	+
			Ganho (6)	<i>GABPB2, EDARADD, H2AFX, ACSS2</i>	-
<i>hsa-mir-642</i>	19q13.32	50,870,025-50,870,122	Perda (0)	<i>HIC1, MAN1C1, NEXN, SFRP1, SYNPO2, LMOD1, PGM5, SYNPO2</i>	+
			Ganho (9)	<i>SMC2, GCET2, FAM83H, INA, STEAP3, AK2, AK2, NCBP2, FAM103A1, TMEM183A,</i>	-
<i>hsa-mir-330</i>	19q13.32	50,834,091-50,834,185	Perda (1)	<i>C10orf54, MSRB2</i>	+
			Ganho (9)	<i>IRX2, DLX1, KRT80, SCG3, ONECUT2, RNF43, C3orf57, TMEM48, LRP8, LRP8, RNF182, CD247, TMEM48, CD47, MRRF, TDRKH, TACC3, WNT7B, MYPN, SEZ6, SLC2A1, TMEM48, ADK, FAM110A, TRIM14, NRAS, TOMM34, CCDC21, FGD3, RMI1, GFM2, TDRKH, C22orf29, C1QBP, CCDC86, LDHB, ARPC5L, GRPEL2, TIMM13, DNM1L, TOMM34, CCDC86, DOLPP1, GFM2, PDXK, AUP1, ASB1, POLE3, XRCC5</i>	-
<i>hsa-miR-769-5p</i>	19q13.32	51,214,029-51,214,147	Perda (1)	<i>SPRYD3, PPP2R5B, TOM1L2, VAMP2, SMO, ATP8B2, ITPK1, ZNF706, KLF9, EPHB3, ENAH, FGFR1, EGFLAM, CDKL1, PDE1A, FGFR1, PARVA, CDKL1, SYNPO, SLC9A9, PALM, PRIMA1, FOXO1, TES, RNF150, HEYL, CTTNBP2, ACVRL1, CCDC8, TES, FOXO1, COL6A3, PDE1A, LONRF2, MYL9, GPRASP1, RAI2, GIMAP8, RASL12, GIMAP8, PRND, CD93</i>	+
			Ganho (9)	<i>TPD52, PACSIN1,</i>	-
<i>hsa-mir-220c</i>	19q13.32	53,755,340-53,755,423	Perda (1)	<i>FAM53B, CBX7, MFAP4</i>	+
			Ganho (9)	<i>MAL, YKT6</i>	-
<i>hsa-mir-150</i>	19q13.33	54,695,853-54,695,937	Perda (1)	<i>EREG, MYB, IGF2BP1, NEK2, C13orf34, IGF2BP1, GINS1, EREG, AICDA, MAPK13, C16orf63, AIFM2, CTH, PTP4A1, METTL2A, MPZL1, EIF5, MPZL1, SUV39H2, VTA1, NSUN4, FADS1, TBC1D20, C17orf39, IPO9, FOXK2</i>	-
			Ganho (8)	<i>FTO, TEAD1, FBXL5, CEP68, ZNF25FZD4, DDAH2, ZBTB4, ZEB1, PDCD4, GTF2IRD2, MYADM, MAMDC2, PRRT2, SLC8A1, PIK3R1, PTPRB, SLC8A1, MASP1, SLC8A1, ENTPD1, PRICKLE2, ITIH5, SCN2B, AOC3</i>	+
<i>hsa-mir-647</i>	20q13.33	62,044,427-62,044,523	Perda (1)	<i>FTO, TEAD1, FBXL5, CEP68, ZNF25FZD4, DDAH2, ZBTB4, ZEB1, PDCD4, GTF2IRD2, MYADM, MAMDC2, PRRT2, SLC8A1, PIK3R1, PTPRB, SLC8A1, MASP1, SLC8A1, ENTPD1, PRICKLE2, ITIH5, SCN2B, AOC3</i>	+
			Ganho (4)	<i>STIL, LCP1, HSPC159</i>	-
			Perda (1)	<i>C6orf89, FAM109B, XAF1, ZNF521, RHOJ,</i>	+

*Sinais positivos indicam expressão gênica aumentada e sinais negativos, indicam expressão diminuída (*log2ratio*, cut-off de 1.0). Nesta tabela, os miRNAs envolvidos em ganhos foram dispostos a frente dos genes preditos com diminuição de expressão e os miRNAs envolvidos em perdas, a frente dos genes com aumento de expressão. Esta ordem de apresentação representa o mecanismo mais provável de regulação de miRNA (correlação inversa), porém, mecanismos de regulação com correlação positiva e/ou negativa dos miRNAs e genes preditos já foram relatados.

4.5.3.3 Agrupamento hierárquico

Para verificar se os 75 genes moduladores estavam associados a casos específicos, foi realizado um teste de agrupamento hierárquico não supervisionado (teste de 1000 permutações aleatórias e correlação de *Pearson*), considerando o perfil genômico e transcriptômico das amostras (Figura 33A e B, respectivamente). Após a realização do teste, as amostras foram categorizadas de acordo com a fase do ciclo, número de amostras avaliadas e diagnóstico de tumores múltiplos ou únicos. De forma geral, tanto o perfil genômico quanto transcriptômico dos 75 genes moduladores não foram estaticamente associados com o agrupamento das amostras de acordo com as categorias selecionadas. De fato, houve uma tendência de agrupamento entre as amostras de tumores múltiplos avaliados da mesma paciente em relação aos casos que tiveram apenas uma amostra avaliada. Entre os casos com diagnóstico de tumores múltiplos, apenas as amostras 300T_A e B e 301T_B e C apresentaram maior similaridade de ganhos e perdas genômicas e de expressão gênica. Além disso, foi gerado um *heatmap* dos genes, para observação do perfil de ganhos e perdas e o padrão de expressão gênica dos mesmos genes distribuídos entre as mesmas amostras (Figura 34).

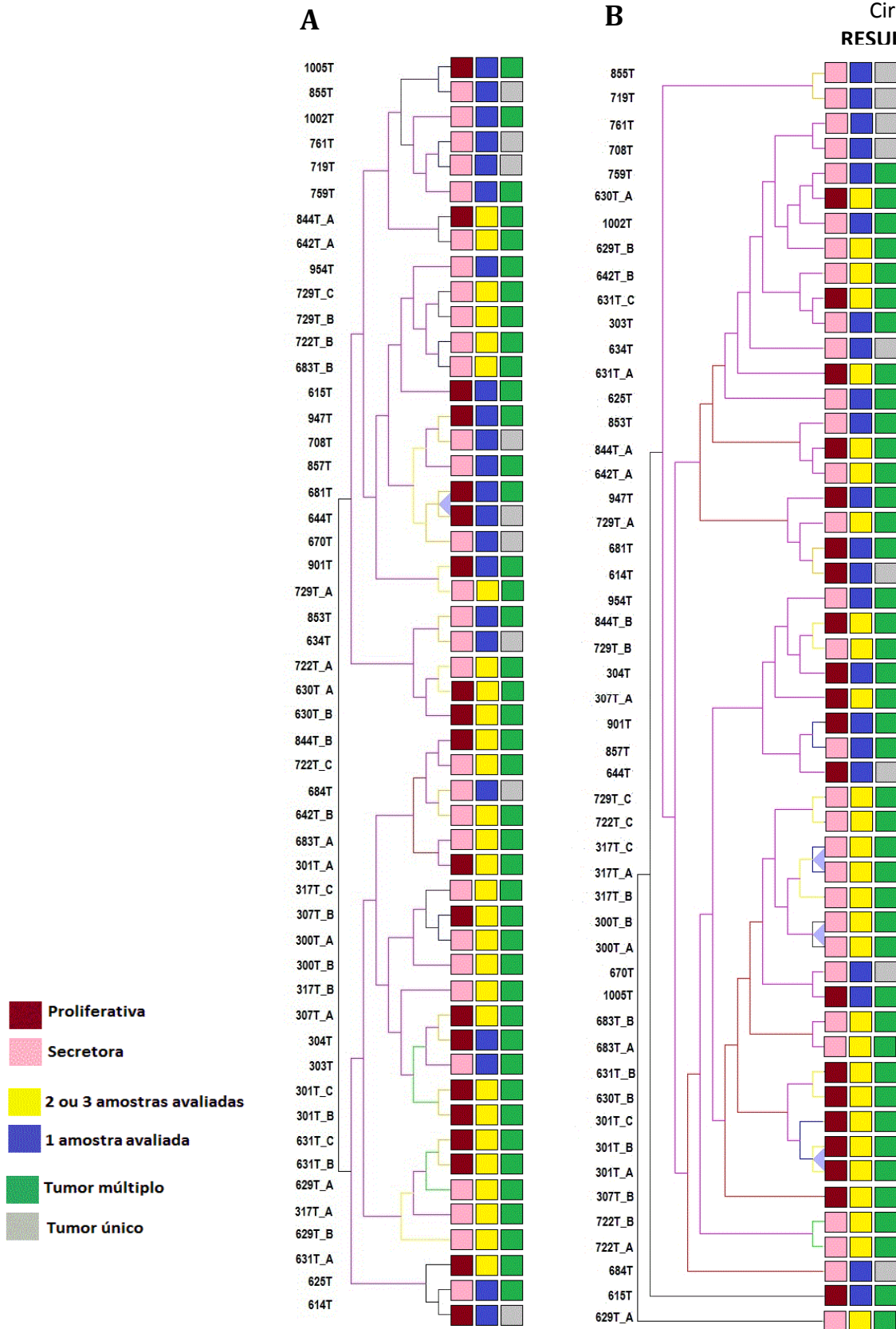



Figura 33. Agrupamento hierárquico não-supervisionado das amostras, considerando o perfil genômico (A) e transcritômico (B) dos 75 genes moduladores. A Figura mostra a categorização das amostras considerando fase do ciclo menstrual (proliferativa e secretora), número de amostras avaliadas (2 ou mais amostras ou uma amostra) e diagnóstico de tumores múltiplos ou únicos (*TMeV v.4.5*).

 perda genômica ou diminuição de expressão

 ganho genômico ou aumento de expressão

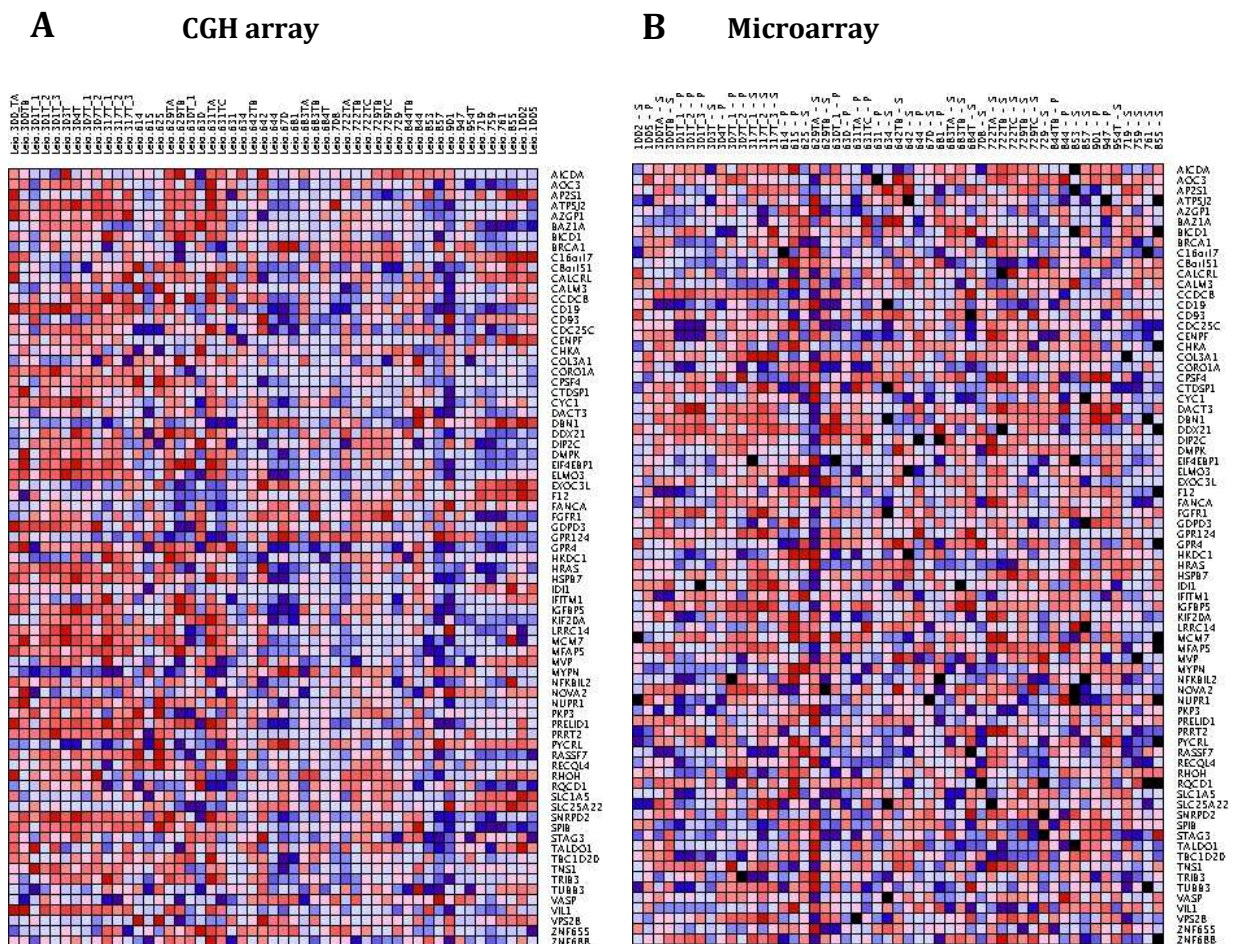


Figura 34. Genes moduladores identificados pelo CONEXIC. **A** - Padrão de ganhos e perdas dos 75 genes identificados por CGH array. **B** - Representação do perfil de expressão gênica dos mesmos 75 moduladores obtidos por análises globais de expressão gênica. Nas colunas, tanto em **A** como em **B**, as amostras estão dispostas na mesma ordem e nas linhas, estão os genes. Observar o perfil genômico e transcriptômico do mesmo gene na mesma amostra.

4.5.3.4 Frequência de localização cromossômica

Após a anotação das 25 regiões genômicas onde os 75 genes moduladores estavam mapeados, foi obtida a frequência de localização destas regiões ao longo dos cromossomos. Esta análise é crucial para determinar se os ganhos e perdas identificados estão preferencialmente localizados em regiões frequentemente envolvidas em pontos de quebra cromossômicos. De acordo com a análise de frequência, aproximadamente 60% dos genes estão mapeados em regiões que *a priori* não são alvos frequentes de quebras. Outros 30% dos genes estão localizados em regiões telôméricas, que são alvos frequentes de instabilidade cromossômica (Figura 35).

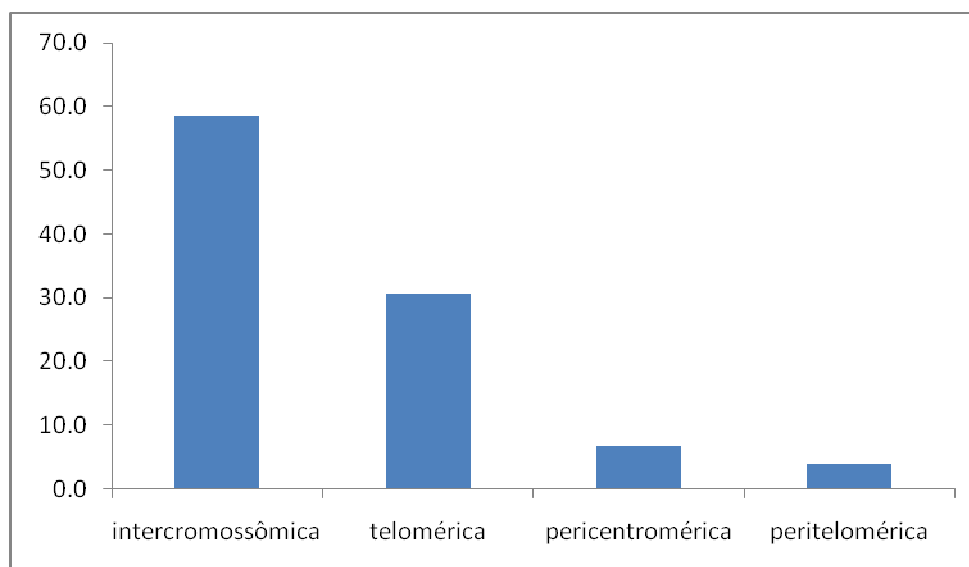


Figura 35. Frequência de localização de genes ao longo dos cromossomos. A maioria dos genes (~60%) está mapeada em regiões intercromossômicas e em regiões telôméricas (30%). Em torno de 6% estão mapeados em regiões próximas aos telômeros.

4.5.3.5 Meta-análise com dados públicos – GEO

Os 75 genes moduladores foram comparados ao banco de dados GEO (*Gene Expression Omnibus* - <http://www.ncbi.nlm.nih.gov/geo/>) onde estão depositados os resultados de estudos de expressão gênica em larga escala, incluindo cinco estudos em Leiomiomas Uterinos. Esta análise permitiu identificar genes que já foram associados com LU e também a descrição de novos transcritos. Estes estudos estão armazenados sob os seguintes códigos: GSE764 (Quade et al., 2004), GSE593 (Hoffman et al., 2004), GSE2152 (Vanharanta et al., 2005), GSE13319 (Crabtree et al., 2009) e GSE12814 (Hodge et al., 2009), respectivamente (Tabela 13).

Todos os 75 genes foram identificados em pelo menos três dos cinco estudos prévios. Estes genes apresentaram diferentes padrões de expressão, onde 52 deles (69%) apresentaram o mesmo padrão de expressão em pelo menos um estudo depositado no GEO.

Tabela 13. Meta-análise dos genes moduladores detectados em estudos prévios de micrarray em Leiomiomas Uterinos depositados no banco de dados *GEO* - *Genomic Expression Omnibus* (<http://www.ncbi.nlm.nih.gov/geo/>).

Gene	Estudo atual	Hoffman et al 2004	Quade et al 2004	Vanharanta et al 2005	Crabtree et al 2009	Hodge et al 2009
<i>AICDA</i>	-2,73	0,79	ND	0,85	0,03	0,48
<i>AOC3</i>	6,50	1,59	0,30	0,51	1,04	0,84
<i>AP2S1</i>	-1,86	1,01	1,85	1,25	1,99	2,22
<i>ATP5J2</i>	-1,01	2,34	ND	1,67	0,42	-0,94
<i>AZGP1</i>	-2,71	0,51	-0,83	0,19	0,00	-0,15
<i>BAZ1A</i>	-1,66	1,46	ND	0,81	-0,61	1,37
<i>BICD1</i>	-2,23	1,37	ND	0,96	0,13	1,03
<i>BRCA1</i>	-2,39	1,00	-1,71	1,12	1,27	1,19
<i>C16orf7</i>	0,89	1,84	ND	2,06	1,91	1,96
<i>C8orf51</i>	-2,58	1,11	ND	1,30	1,12	0,72
<i>CALCRL</i>	4,67	1,13	1,27	0,75	0,36	0,86
<i>CALM3</i>	-1,43	1,51	1,69	1,67	1,59	1,65
<i>CCDC8</i>	3,53	ND	ND	ND	0,90	0,93
<i>CD19</i>	-2,84	1,90	0,07	-0,09	0,55	0,79
<i>CD93</i>	7,31	2,14	ND	1,17	0,86	0,61
<i>CDC25C</i>	-4,12	0,93	0,08	1,08	-0,86	0,39
<i>CENPF</i>	-5,04	3,24	0,05	1,18	1,11	0,89
<i>CHKA</i>	-1,82	0,96	-2,39	0,90	0,43	1,02
<i>COL3A1</i>	4,68	2,55	1,63	1,73	1,62	2,57
<i>CORO1A</i>	-1,64	1,00	ND	0,03	0,23	0,74
<i>CPSF4</i>	-1,03	1,54	ND	2,02	2,03	1,78
<i>CTDSP1</i>	1,14	2,83	ND	1,79	2,25	1,77
<i>CYC1</i>	-1,35	2,65	-2,21	1,65	1,93	1,59
<i>DACT3</i>	6,05	ND	ND	ND	-0,98	1,84
<i>DBN1</i>	1,93	2,10	-0,08	1,64	1,31	1,63
<i>DDX21</i>	-2,20	1,92	0,41	1,32	2,47	0,90
<i>DIP2C</i>	2,23	1,60	ND	2,15	0,17	0,73
<i>DMPK</i>	2,07	3,83	0,84	0,79	1,28	0,34
<i>EIF4EBP1</i>	-1,67	0,85	ND	0,80	1,33	1,82
<i>ELMO3</i>	-1,58	-0,16	ND	0,43	1,48	0,59
<i>EXOC3L1</i>	1,98	ND	ND	ND	0,65	0,51
<i>F12</i>	-2,89	0,09	0,15	0,35	-0,83	0,22
<i>FANCA</i>	-3,26	0,29	ND	0,68	-1,39	1,13
<i>FGFR1</i>	2,36	1,22	1,28	1,12	1,57	0,38
<i>GDPD3</i>	-1,91	1,46	ND	1,75	1,54	0,75
<i>GPR124</i>	2,41	4,22	ND	1,16	1,29	1,35
<i>GPR4</i>	3,83	2,27	-1,89	1,00	0,59	0,72
<i>HKDC1</i>	-3,98	-0,75	ND	-0,20	-0,20	-0,58
<i>HRAS</i>	-1,30	1,42	-0,53	1,28	1,67	1,59
<i>HSPB7</i>	3,68	2,23	ND	0,50	0,95	1,09
<i>IDI1</i>	-2,21	3,45	ND	2,37	1,49	1,39
<i>IFITM1</i>	2,07	1,01	0,97	1,83	1,25	0,59
<i>IGFBP5</i>	3,12	0,53	0,23	0,24	-4,32	0,75
<i>KIF20A</i>	-3,70	0,42	ND	-1,62	-0,17	0,17
<i>LRRC14</i>	-1,18	0,20	0,43	0,13	1,55	1,42
<i>MCM7</i>	-1,77	1,64	1,19	1,35	0,68	1,19
<i>MFAP5</i>	5,92	2,26	0,56	0,23	1,33	1,04
<i>MVP</i>	1,56	1,37	-0,16	1,47	1,00	1,02
<i>MYPN</i>	-2,51	ND	ND	ND	-1,45	-0,53
<i>NOVA2</i>	1,79	0,82	ND	-1,54	0,02	-1,04
<i>NUPR1</i>	1,85	1,76	ND	0,62	1,65	1,67
<i>PKP3</i>	-2,73	0,45	ND	0,62	0,08	-0,59
<i>PRELID1</i>	-1,60	ND	ND	ND	2,05	2,10
<i>PRRT2</i>	3,23	ND	ND	ND	0,81	1,01
<i>PYCR1</i>	-0,98	0,34	ND	0,76	-0,62	-0,37
<i>RASSF7</i>	-2,57	1,60	-0,70	1,01	0,00	-1,10
<i>RECQL4</i>	-2,50	1,59	ND	0,17	-0,37	0,08
<i>RHOH</i>	-2,38	-1,19	0,67	1,06	0,30	0,26

RQCD1	-2,00	0,42	-0,51	1,17	-0,63	1,88
SLC1A5	-2,24	2,33	0,97	0,92	1,42	1,39
SLC25A22	-1,57	0,09	ND	1,12	1,35	1,24
SNRPD2	-0,66	3,68	1,00	2,18	2,32	2,10
SPIB	-2,43	0,66	1,00	1,11	-1,27	0,93
STAG3	-1,83	1,49	ND	0,33	1,01	0,70
TALDO1	-1,43	2,12	-2,05	0,08	2,27	1,46
TBC1D20	-1,26	ND	ND	ND	1,51	2,08
TNS1	3,27	-1,61	ND	-0,87	0,26	-0,60
TRIB3	-4,66	1,68	ND	1,72	-0,73	0,83
TONSL	-1,91	-0,56	-0,30	-0,10	-0,77	0,79
TUBB3	-2,23	1,30	1,85	1,49	1,70	1,78
VASP	2,32	2,44	0,27	1,34	1,51	0,62
VIL1	-3,94	1,34	ND	0,09	-0,36	0,81
VPS28	1,40	2,23	ND	2,12	1,78	1,90
ZNF655	-1,37	ND	ND	ND	0,89	0,64
ZNF688	1,25	2,19	ND	1,57	-0,60	1,52

Valores apresentados em \log_2 ratio. Valores negativos indicam diminuição de expressão e valores positivos indicam expressão aumentada. Em negrito, genes que apresentaram o mesmo padrão de expressão gênica em pelo menos um dos estudos depositados no GEO. ND=não identificado no estudo.

4.5.3.6 Análise de enriquecimento de genes - GSEA

Após a obtenção dos 75 genes moduladores, foi realizada a análise de enriquecimento para estes genes (*GSEA - Gene Set Enrichment Analysis* - <http://www.broadinstitute.org/gsea/index.jsp>) (Subramanian et al., 2005), baseada no valores de expressão gênica. Esta análise permite avaliar se cada um dos módulos associados a estes genes são enriquecidos com genes diferencialmente expressos identificados no estudo. Os resultados revelaram que entre os 75 genes moduladores, 32 foram enriquecidos entre 26 módulos e 43 genes não foram enriquecidos em qualquer um dos módulos (Tabela 14). Os genes *CORO1A*, *FGFR1*, *DDX21* e *DBN1* foram os mais frequentemente identificados entre os módulos e apenas *FGFR1* e *DBN1* apresentaram correlação positiva (ganho/aumento expressão). Outros genes também apresentaram correlação positiva como *NUPR1*, *COL3A1*, *AOC3*, *IFITM1*, *CALCRL*, *NOVA2*, *MVP*, *MFAP5*, *DIP2C* e correlação negativa como *CENPF* e *BICD1*. A Tabela 15 mostra os valores de enriquecimento para os 26 módulos que foram significativamente associados para os 32 genes. O módulo 8 apresentou o maior número de genes sobrepostos, com os nove genes incluindo *CORO1A*, *MCM7*, *SLC1A5*, *DDX21*, *NUPR1*, *DBN1*, *EIF4EBP1*, *VIL1*, *CHKA* (P valor $<10^{-3}$).

Tabela 14. Genes identificados nos módulos de enriquecimento gênico.

Gene	Nº de módulos enriquecidos	CGH array	Microarray	Região
<i>COR01A</i>	8	+	-	16p11.2
<i>FGFR1</i>	8	+	+	8p12
<i>DDX21</i>	8	+	-	10q21.3
<i>DBN1</i>	8	+	+	5q35.3
<i>CENPF</i>	7	-	-	1q41
<i>EIF4EBP1</i>	7	+	-	8p12
<i>NUPR1</i>	6	+	+	16p11.2
<i>COL3A1</i>	5	+	+	2q32.2
<i>AOC3</i>	4	+	+	17q21.31
<i>HRAS</i>	4	+	-	11p15.5
<i>MCM7</i>	3	+	-	7q22.1
<i>IFITM1</i>	3	+	+	11p15.5
<i>SPIB</i>	3	+	-	19q13.33
<i>CD19</i>	3	+	-	16p11.2
<i>SLC1A5</i>	2	+	-	19q13.32
<i>VIL1</i>	2	+	-	2q35
<i>CHKA</i>	2	+	-	11q13.2
<i>IDI1</i>	2	+	-	10p15.3
<i>TUBB3</i>	2	+	-	16q24.3
<i>CALCRL</i>	2	+	+	2q32.1
<i>STAG3</i>	2	+	-	7q22.1
<i>RECQL4</i>	2	+	-	8q24.3
<i>RHOH</i>	2	-	-	4p14
<i>NOVA2</i>	2	+	+	19q13.32
<i>CDC25C</i>	2	+	-	5q31.2
<i>AP2S1</i>	1	+	-	19q13.32
<i>SNRPD2</i>	1	+	-	19q13.32
<i>MVP</i>	1	+	+	16p11.2
<i>MFAP5</i>	1	+	+	12p13.31
<i>BICD1</i>	1	-	-	12p11.21
<i>DIP2C</i>	1	+	+	10p15.3
<i>FANCA</i>	1	+	-	16q24.3

Em negrito, genes que apresentaram correlação positiva ou negativa. Os sinais positivos (+) e negativos (-) representam o *status* dos genes em relação ao padrão de ganhos (+) e perdas (-) e ao aumento (+) e diminuição (-) de expressão gênica.

Tabela 15. Análise de enriquecimento de genes moduladores realizada pelo *GSEA - Gene Set Enrichment Analysis* (<http://www.broadinstitute.org/gsea/index.jsp>).

Módulo	Genes no Gene Set (K)	Genes sobrepostos (k)	k/K	P valor
533	42	<i>CENPF, CORO1A, AOC3</i>	0.0714	<10 ⁻³
8	406	<i>CORO1A, MCM7, SLC1A5, DDX21, NUPR1, DBN1, EIF4EBP1, VIL1, CHKA</i>	0.0222	<10 ⁻³
509	14	<i>AP2S1, IDI1</i>	0.1429	<10 ⁻³
315	15	<i>CENPF, TUBB3</i>	0.1333	<10 ⁻³
5	423	<i>CORO1A, AOC3, DDX21, NUPR1, FGFR1, CALCRL, COL3A1, MVP</i>	0.0189	<10 ⁻²
292	133	<i>IFITM1, SPIB, CD19, STAG3</i>	0.0301	<10 ⁻²
3	375	<i>MCM7, SLC1A5, DDX21, DBN1, EIF4EBP1, FGFR1, CALCRL</i>	0.0187	<10 ⁻²
118	395	<i>CENPF, CORO1A, DDX21, DBN1, EIF4EBP1, FGFR1, RECQL4</i>	0.0177	<10 ⁻²
86	43	<i>HRAS, RHOH</i>	0.0465	<10 ⁻²
137	531	<i>CORO1A, NUPR1, DBN1, EIF4EBP1, FGFR1, HRAS, NOVA2, BICD1</i>	0.0151	<10 ⁻²
12	351	<i>DDX21, NUPR1, DBN1, FGFR1, COL3A1, DIP2C</i>	0.0171	<10 ⁻²
9	114	<i>EIF4EBP1, VIL1, HRAS</i>	0.0263	<10 ⁻²
430	52	<i>CORO1A, FANCA</i>	0.0385	<10 ⁻²
1	361	<i>AOC3, DDX21, NUPR1, DBN1, FGFR1, COL3A1</i>	0.0166	<10 ⁻²
105	194	<i>CHKA, IDI1, FGFR1, COL3A1</i>	0.0206	<10 ⁻²
57	54	<i>MCM7, CDC25C</i>	0.0370	<10 ⁻²
345	122	<i>IFITM1, SPIB, CD19</i>	0.0246	<10 ⁻²
254	58	<i>NUPR1, IFITM1</i>	0.0345	<10 ⁻²
337	60	<i>CENPF, DDX21</i>	0.0333	<10 ⁻²
53	391	<i>CENPF, CORO1A, DDX21, DBN1, EIF4EBP1, RECQL4</i>	0.0153	<10 ⁻²
129	217	<i>DBN1, FGFR1, HRAS, NOVA2</i>	0.0184	<10 ⁻²
119	138	<i>SPIB, CD19, STAG3</i>	0.0217	<10 ⁻²
122	138	<i>AOC3, COL3A1, MFAP5</i>	0.0217	<10 ⁻²
308	67	<i>CORO1A, CDC25C</i>	0.0299	<10 ⁻²
438	67	<i>CENPF, CORO1A</i>	0.0299	<10 ⁻²
32	232	<i>SNRPD2, EIF4EBP1, TUBB3, RHOH</i>	0.0172	<10 ⁻¹

K = moduladores em sobreposição associados com os genes nos módulos; k = genes diferencialmente expressos associados aos moduladores; k/K = razão entre genes presentes nos módulos (genes diferencialmente expressos associados aos moduladores)/genes em sobreposição (moduladores em sobreposição associados com os genes nos módulos).

4.5.4 Análise funcional dos genes moduladores

Os 75 genes moduladores foram submetidos a análises funcionais pelo sistema IPA. Os resultados obtidos mostraram 56 genes elegíveis presentes em cinco redes gênicas com valores de *scores* variando de 17 a 25. Valores de *scores* mais altos indicam maior associação entre os genes importados com as moléculas do banco de dados IPA. A Tabela 16 mostra as funções biológicas, doenças e funções moleculares, associadas com as moléculas em cada rede. A rede com *score* mais alto (25) associou 13 moléculas elegíveis a partir dos dados obtidos com 22 outras moléculas do IPA. As funções significativas associadas a estas moléculas foram sinalização celular, metabolismo de ácidos nucleicos e bioquímica de pequenas moléculas. Entre todas as funções identificadas, moléculas envolvidas com ciclo celular foram as mais frequentes, estando presentes nas Redes 2, 3 e 4, enquanto que a Rede 5, associou moléculas envolvidas com expressão gênica, desenvolvimento do organismo e desenvolvimento e função do sistema nervoso. As Figuras 36A e B mostram uma representação esquemática das conexões entre os genes presentes nas Redes 1 e 4, respectivamente. Estas redes foram selecionadas para apresentação, pois melhor representam as funções associadas aos genes alterados.

As doenças e funções moleculares associadas aos genes estão representadas na Figura 37A e B, respectivamente. As doenças que apresentaram os três maiores valores de $-\log(P \text{ valor})$ foram câncer (~ 4.25), doença gastrointestinal (~ 4.25) e doença do sistema reprodutivo (3.25). Doença genética (3.00) e doença muscular e esquelética (2.50) foram classificadas na terceira e quarta posições, respectivamente. As cinco funções moleculares que apresentaram os maiores valores de $-\log(P \text{ valor})$ foram ciclo celular (5.5), morfologia celular (3.8), proliferação e crescimento celular (~ 3.5), manutenção e função celular (3.3) e organização e montagem celular (3.2).

As vias canônicas descritas, plotadas em valores da razão sobre $-\log(P \text{ valor})$, foram principalmente associadas com sinalização do câncer de mama hereditário (2.75), via de

ativação da protrombina (2.20), proteínas CHK no controle do *checkpoint* do ciclo celular (2.15), regulação do *checkpoint* G2/M do ciclo celular em resposta a danos no DNA (1.78) e fibrose hepática (1.70). A via clivagem e poliadenilação de pré-mRNA teve o maior valor de razão (Figura 38). Outras vias já descritas em LU foram geradas com menores valores de associação, como sinalização FGF, IGF-1, EGF e TGF-beta (1.25, 1.20, 0.95 e 0.50, respectivamente). A Figura 39 mostra a via canônica associada com fibrose hepática que possui grande similaridade com as moléculas associadas ao Leiomiomas Uterinos.

Tabela 16. Redes gênicas com os *scores* gerados a partir da associação dos 75 genes moduladores com o banco de dados *Ingenuity Pathway Analysis* (IPA).

Rede	Moléculas da rede*	Score	Moléculas em foco	Principais funções
1	<i>ADCYAP1R1, ADRB2, ↓AP2S1, ↓BICD1, CALCR, ↑CALCRL, ↓CORO1A, CRHR1, DISC1, EPB49, F Actin, GLP1R, GLP2R, Gpcr, ↑GPR4, ↑GPR124, GRB2, GRK4, GRK5, ↓HRAS, OPRM1, PAK1, PCNT, PFN, PIK3CG, RAMP3, ↓RASSF7, RGL2, ↓RHOH, SCTR, ↓SNRPD2, ↑TNS1, ↑VASP, ↓VIL1, VIPR1</i>	25	13	Sinalização Celular, Metabolismo de Ácidos Nucléicos, Bioquímica de pequenas Moléculas
2	<i>ATP5F1, ↓ATP5J2, CAMKK2, CASP4, ↑CD93, CRK, DBF4, ↓DDX21, ↑DMPK, F3, ↑FGFR1, FLNA, FRS2, GJA1, HSPB2, ↑HSPB7, ↑IGFBP5, JUN, KRT1, ↓MCM7, MYB, NEK6, PNO1, RPL35, SHB, ↓SLC1A5, ↓SLC25A22, SMC3, SMC1A, SP1, SPARC, ↓SPIB, ↓STAG3, ↓TALDO1, TRAF6</i>	25	13	Ciclo Celular, Proliferação e Crescimento Celular, Desenvolvimento e Função do Sistema Muscular e Esquelético
3	<i>ARRB2, ↓BRCA1, BRCC3, ↓CD19, ↓CDC25C, Cdk, CDK4, ↓CPSF4, CSDA, DYRK1A, ↓EIF4EBP1, FAM175A, ↓FANCA, FANCF, HMGA1, ↓IDI1, IFI6, IFI27, IFIT1, IFIT2, IFIT3, ↑IFITM1, MARK3, MVD, MX1, PAK7, ↓PKP3, ↓PYCRL, RPL18, RPL31, SFN, ↓TUBB3, UIMC1, WEE1, ↓ZNF655</i>	22	12	Ciclo Celular, Doenças e Condições Dermatológicas, Doença Imunológica
4	<i>↓AICDA, ↓AZGP1, ↓BAZ1A, BAZ1B, C12orf11, ↓CALM1, CDC14A, ↑COL3A1, ↓CYC1, ↑DBN1, FAM125A, FGF7, GTF2E1, GTF2H4, Histone h3, HNF4A, HTT, KCNN2, ↓KIF20A, MSC, ↓MYPN, PNO1, POLE3, PPP1CA, RIOK3, ROCK1, RQCD1, SUPT5H, TCF12, TNFSF11, UQCRC1, UQCRC2, UQCRH, ↑VPS28, VPS37B</i>	20	11	Montagem e Organização Celular, Duplicação do DNA, Recombinação e Reparo, Metabolismo de Carboidratos
5	<i>AKT2, CCL5, ↓CENPF, COPS6, CSF2, ↑CTDSP1, CYP19A1, CYP1A1, DLL1, ENO1, EP300, ESR1, ↓F12, HES1, JAG1, JAG2, KPNA1, KPNA2, ↓LRRRC14, MED24, ↑MFAP5, MNAT1, MTA1, ↑MVP, MYOD1, NFKB1, NOTCH1, ↑NUPR1, PTEN, RBPJ, ↓RECQL4, SMAD1, SMARCE1, ↓TONSL, ↓TRIB3</i>	17	10	Expressão Gênica, Desenvolvimento do Organismo, Desenvolvimento e Função do Sistema Nervoso

*Setas para cima (vermelhas) indicam genes mapeados em regiões de ganhos genômicos e setas para baixo (verdes) indicam genes mapeados em regiões de perdas genômicas. Em negrito, moléculas identificadas no estudo atual. Em negrito, moléculas identificadas no estudo atual.

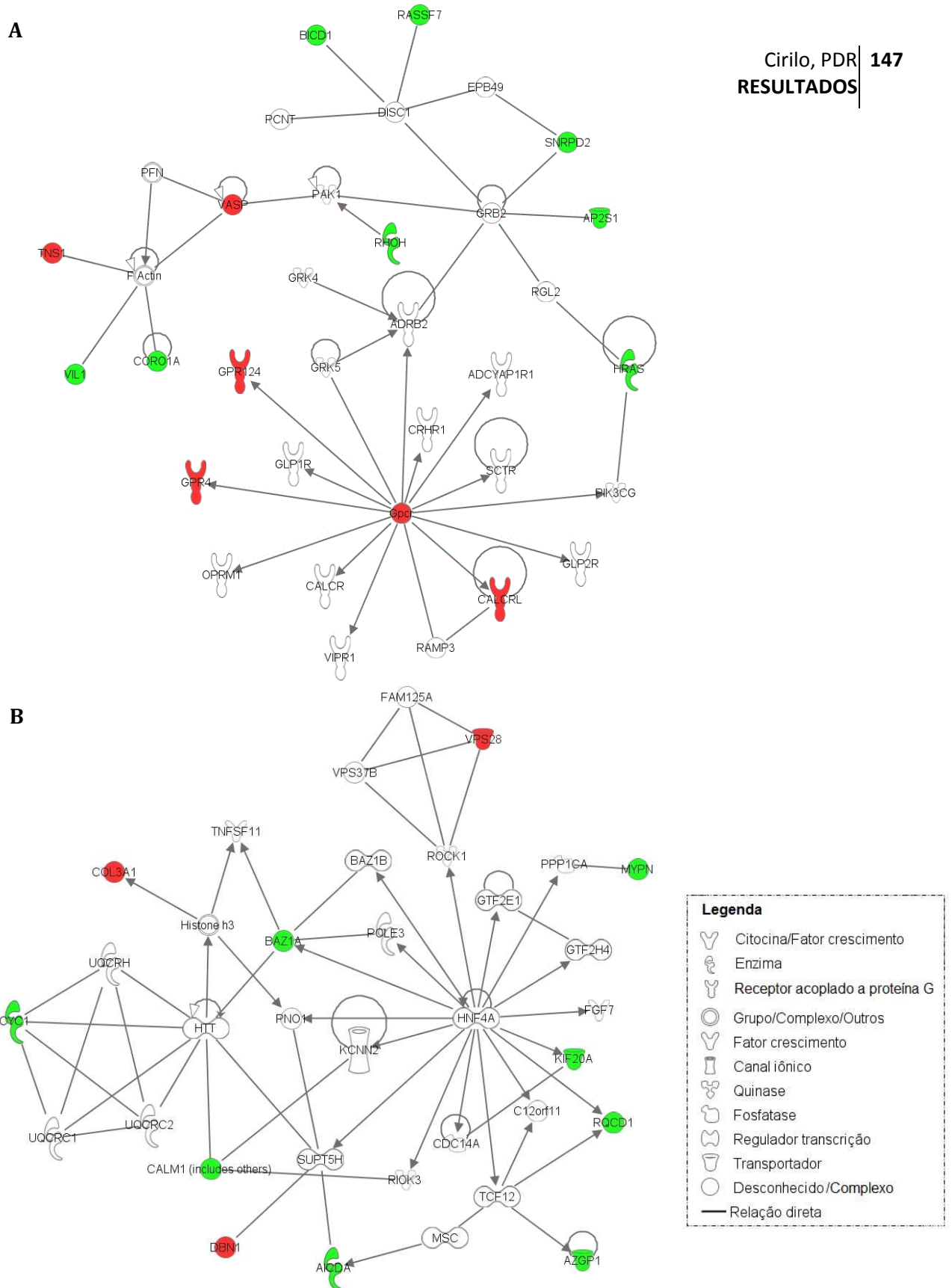


Figura 36. Redes geradas a partir das interações entre os 75 genes moduladores e o banco de dados IPA. **A - Rede 1:** genes associados com sinalização celular, metabolismo de ácidos nucléicos e bioquímica de pequenas moléculas; **B - Rede 4:** genes envolvidos com ciclo celular, doença hematológica e metabolismo de carboidratos. Em verde, genes com expressão diminuída e em vermelho, genes com expressão aumentada (IPA).

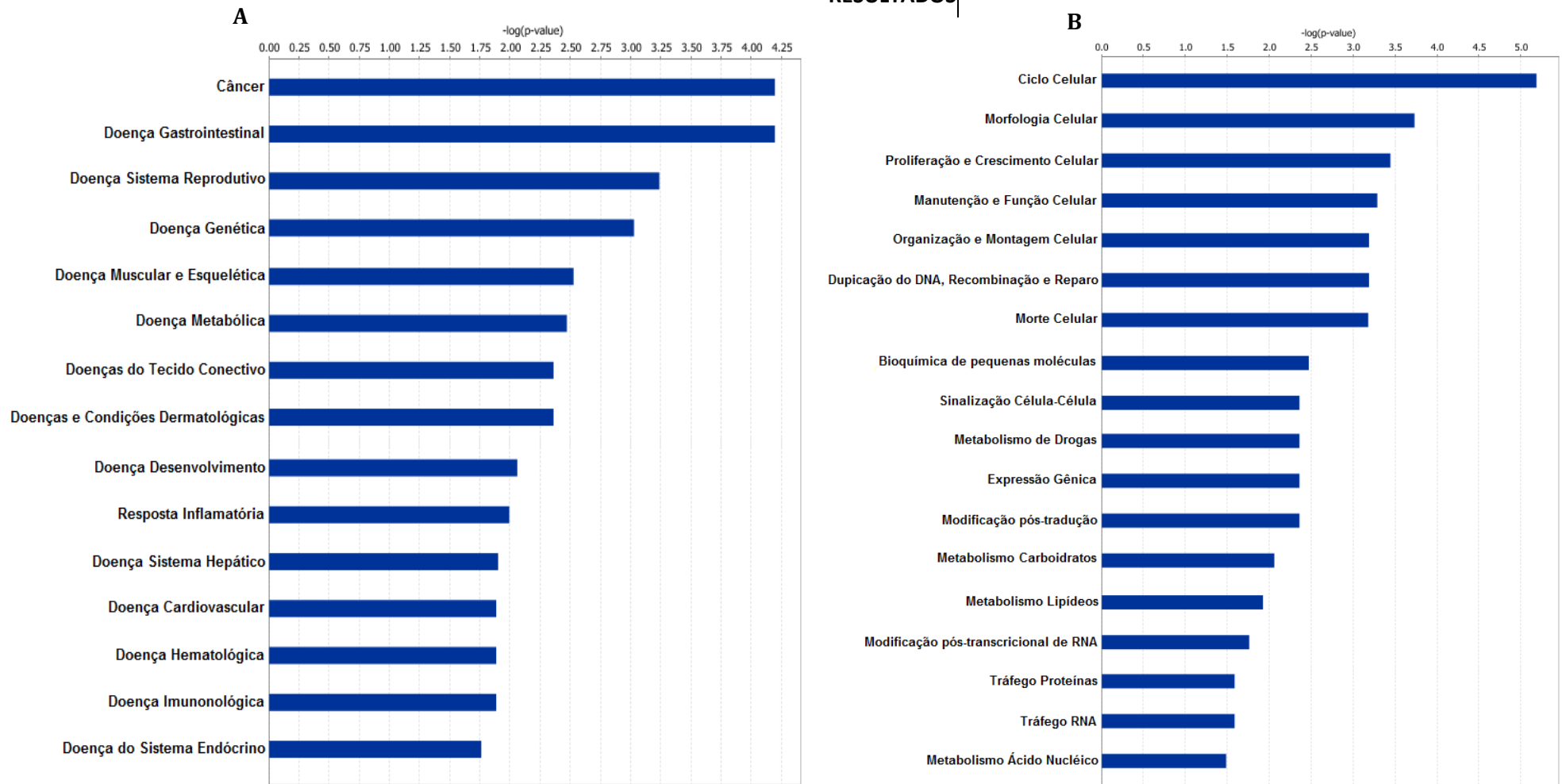


Figura 37. Doenças e funções associadas aos genes moduladores, em valores de $-\log(P \text{ valor})$. **A-** Câncer, doença gastrointestinal e do sistema reprodutivo aparecem com valores próximos ao P valor máximo observado. **B** – Ciclo celular, morfologia celular e proliferação e crescimento celular foram as funções com maior $-\log(P \text{ valor})$ (IPA).

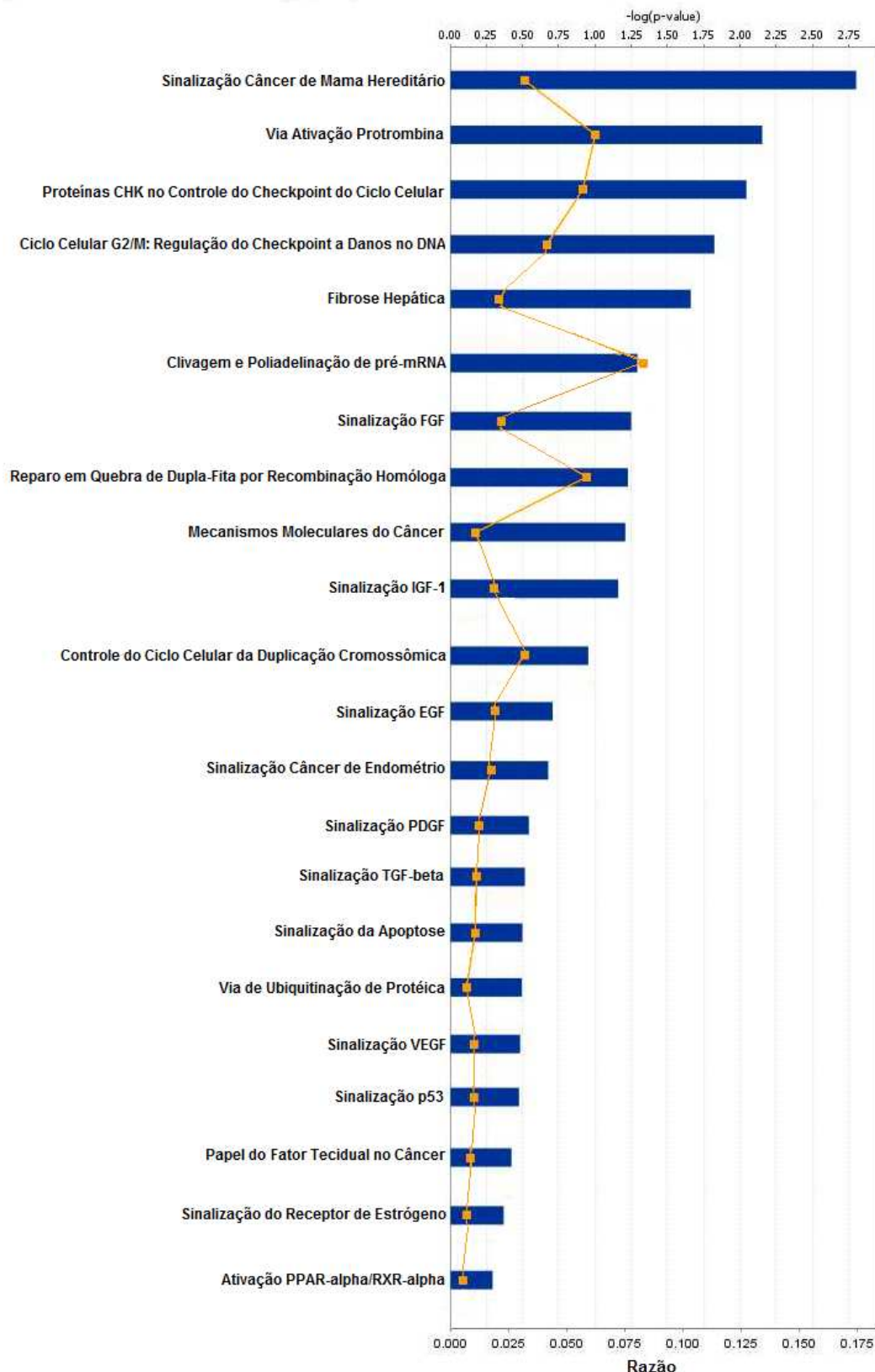


Figura 38. Vias canônicas associadas com os genes moduladores. As cinco principais vias descritas foram sinalização do câncer de mama hereditário, via de ativação da protrombina, proteínas CHK no controle do *checkpoint* do ciclo celular, regulação do *checkpoint* G2/M do ciclo celular em resposta a danos no DNA e fibrose hepática (valores da razão sobre $-\log P$ valor) (IPA).

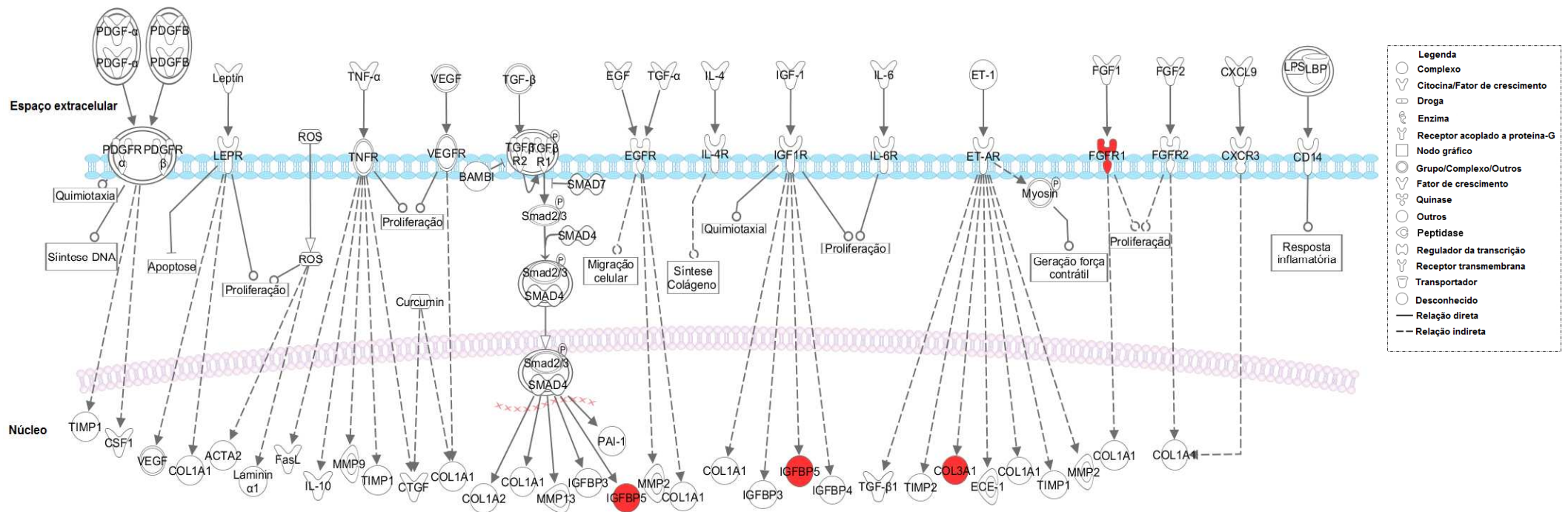


Figura 39. Sinalização da fibrose hepática semelhante aos processos proliferativos e fibróticos dos LU. O aumento da atividade tirosina quinase de FGFR1, juntamente com o aumento de IGFBP5 em resposta a fatores de crescimento poderia desencadear a proliferação celular identificada nestes tumores, porém, o aumento da atividade de COL3A1, favoreceria o acúmulo de matriz extracelular e deposição de colágeno nas células do miométrio (IPA). Em vermelho, moléculas com expressão aumentada (IPA).

5. Discussão

Os recentes achados sobre a identificação de alterações no número de cópias do DNA (CNVs) em indivíduos saudáveis têm gerado novas discussões sobre a origem genética das doenças humanas (Redon et al., 2006; Wheeler et al., 2008; Mills et al., 2011). Estas alterações, originadas pelos mecanismos celulares de recombinação cromossômica (Hastings et al., 2009), podem ter o envolvimento de poucos nucleotídeos até mais de 1Kb de sequências genômicas (Mills et al., 2011). Resta-nos saber, qual delas são variações normais do genoma, quais possuem associação com fenótipos específicos, qual o modelo de herança genética e quais possuem efeitos deletérios. Estudos atuais em tumores humanos têm como objetivo caracterizar o perfil de alterações genômicas e associá-los com achados de expressão gênica global, na tentativa de descrever quais CNVs possuem relação direta com o fenótipo tumoral (alterações funcionais) (Akavia et al., 2010). Esta estratégia pode contribuir para o conhecimento da etiologia dos tumores humanos.

Os Leiomiomas Uterinos são os tumores benignos mais comuns do trato genital que afeta mulheres em idade reprodutiva e com etiologia pouca conhecida (Canevari e Rogatto, 2007). Estes tumores constituem um importante problema de saúde pública tanto no Brasil quanto em outros países, como os EUA (Hodge et al., 2008). Aproximadamente 40-50% dos LU apresentam anormalidades citogenéticas não randômicas, envolvendo os cromossomos 3, 6, 7, 13, trissomia do cromossomo 12, translocação recíproca entre os cromossomos 12 e 14 e monossomia do cromossomo 22 (para revisão El-Gharib e Elsobky, 2010). Portanto, parece que metade dos LU são originados por outros mecanismos ou por alterações genômicas submicroscópicas. Neste estudo, 80 amostras de LU foram investigadas para a descrição de alterações no número de cópias de DNA. Os resultados obtidos demonstraram a presença de CNVs em todas as amostras (100% dos casos). Os cromossomos mais frequentemente alterados concordam com os achados descritos em literatura, além do envolvimento dos cromossomos 1, 11, 16,

17 e 19 que apresentaram altas frequências de alterações genômicas. Em adição, um subconjunto de amostras também foi avaliado para o perfil global de expressão gênica, demonstrando concordância parcial com os resultados descritos previamente em literatura. Os dados genômicos e transcriptômicos foram também integrados, utilizando uma abordagem que, para o nosso conhecimento, ainda não foi descrita. Os resultados revelaram potenciais marcadores moleculares para essa doença frequente e cuja etiologia é pouco conhecida. A discussão, em detalhes de cada um destes itens está apresentada a seguir.

5.1 Alterações no número de cópias de DNA

A metodologia utilizada para caracterizar as alterações genômicas foi a CGH array, que possui um maior poder de resolução para a identificação de CNVs, em relação às outras técnicas previamente utilizadas para avaliar os LU, como bandamento GTG e CGH cromossômica (Packenham et al., 1997; Rikala et al., 1998; Mantovani et al., 1999; Levy et al., 2000). Além disso, para o nosso conhecimento, esta é a maior casuística já avaliada utilizando esta abordagem.

Os dados de CGH array revelaram quatro amostras com perdas homozigotas (0,16%) e 79 amostras (3,27%) com alto nível de ganhos ou ampliações genômicas. Em geral, as ampliações estão associadas com falha na resposta a tratamento e estes genes amplificados podem ser alvos para o delineamento de terapias específicas. Como exemplo clássico de terapia molecular está a amplificação do gene *ERBB2* (17q12) em tumores de mama. Esta amplificação é observada em aproximadamente 20% dos tumores de mama e tem sido alvo de tratamentos eficazes utilizando principalmente o fármaco *Trastuzumab* (Gajria e Chandarlapaty, 2011). Entre as ampliações genômicas observadas neste estudo, foi identificado o envolvimento da região 17q12 apenas na amostra 1015T (Anexo

3), com o envolvimento dos genes *AATF*, *ACACA*, *DHRS11*, *LHX1*, *MRM1*. Esta amplificação, dista 3Mb do gene *ERBB2*, que não apresentou alto nível de ganhos neste estudo.

Outra particularidade deste estudo foi a avaliação de diferentes amostras obtidas de uma mesma paciente, totalizando 41 amostras de 17 pacientes. A avaliação da frequência de similaridade entre os diferentes tumores de uma mesma paciente foi baseada nas regiões descritas com os critérios utilizados para a identificação de CNVs excluindo-se aquelas presentes em >5% da população controle, ou seja, um critério menos restritivo de análise. Os resultados revelaram que a frequência de similaridade mais alta foi de 13,7% em pelo menos duas amostras da paciente 301T (301T_A, B e C). As amostras do caso 683T (683T_A e B), não apresentaram nenhuma alteração em comum. Quando foram consideradas as alterações similares presentes em mais de um caso diagnosticado com tumores múltiplos, foram observados ganhos em 2q35, 7q22.1, 8p12, 11q13.1-q13.2, 11p15.5, 12p13.31, 16p13.3 e perdas em 1p31.1, 3q28, 4q13.1, 5q31.1, 9q34.2-q34.3, 11p15.5, 14q32.2, 15q25.3, 16q24.3, 18q22.1-q22.2. Em geral, os resultados revelaram que amostras de tumores múltiplos compartilharam baixa frequência de similaridade de CNVs, contudo parecem existir alterações cruciais para a formação dos LU, pois além de estarem presentes em diferentes amostras de uma mesma paciente, também foram presentes em mais de um caso diagnosticado com tumores múltiplos. Estes resultados confirmam a hipótese no qual LU múltiplos são monoclonais em origem e surgem independentemente no útero (para revisão Canevari et al., 2005a; Cai et al., 2007) e que além disso, apresentam diferentes anormalidades cromossômicas (Mantovani et al., 1999; Pakiz et al., 2010).

A presença de mais de um tumor primário é uma das características observadas em pacientes com tumores familiares (Pakiz et al., 2010). Neste estudo, sete pacientes relataram história de LU em parentes de primeiro e segundo grau e cinco delas foram diagnosticadas com mais de um tumor primário (Anexo 2), totalizando dez amostras (046T_A, 046T_B, 084T, 629T_A, 629T_B, 633T, 683T_A, 683T_B, 842T e 857T). Estas

informações permitiram que fosse realizada uma busca de alterações mínimas em comum compartilhadas por estas amostras, que poderiam portar genes candidatos ao desenvolvimento de LU com origem familiar. Os resultados revelaram duas regiões mínimas comuns envolvidas em perdas em 1p36.33 nas amostras 046T_A, 084T, 629T_B e 633T e em 14q24.3 entre as amostras 046T_A, 629T_B e 842T e os genes minimamente associados foram *PRKCZ* e *VSX2*, respectivamente. A sonda A_14_P131252 cobre o íntron 4 do gene *PRKCZ*, que possui 17 éxons e 16 íntrons (Figura 10A). Este gene codifica um membro da família PKC das serina/treonina quinases, que está envolvido em processos de proliferação, diferenciação e secreção celular. O *VSX2* possui cinco éxons e quatro íntrons e codifica um fator de transcrição específico da retina. Mutações deste gene foram associadas com microftalmia, catarata e anormalidades da íris. A sonda A_14_P200645, que mapeia o gene *VSX2*, está localizada no intron 3 deste gene. Estes resultados indicam que as deleções envolvendo estes genes podem não ser funcionais, pois foram localizadas em regiões intrônicas. Vale ressaltar que nenhuma destas deleções foram similares entre as amostras de tumores múltiplos, indicando que elas podem não estar diretamente associadas com LU familiares. Até o momento, apenas mutações germinativas no gene *FH* (1q42.1) foram associadas a predisposição aumentada à Síndrome de Leiomiomas Uterinos e Cutâneos Múltiplos ou a Síndrome de Leiomiomatose Hereditária e Câncer de Células Renais (Martinez-Mir et al., 2003; Lehtonen et al., 2004). Estudos adicionais são necessários avaliando todos os membros afetados das famílias para dar continuidade a busca de um novo gene candidato a predisposição familiar em Leiomiomas Uterinos.

Considerando que os LU são lesões benignas espera-se que não apresentem um grande número de alterações assim como não sejam observados cariótipos complexos. Contudo, 12 amostras tiveram grandes alterações estruturais (>10Mb) envolvendo os cromossomos 1 e 19. Duas amostras apresentaram um possível cromossomo 1 em anel: ?r(1)(p36.11q42.12) (amostra 084T) e ?r(1)(p31.1q32.2) (amostra 759T). O primeiro relato em literatura sobre a presença de cromossomos em anel envolvendo o cromossomo

1 em LU foi descrito por Nilbert et al. (1988). Os autores avaliaram duas amostras de LU por bandamento G e identificaram, entre outras anormalidades, alterações estruturais secundárias como r(1) nas duas amostras. Outros estudos utilizando citogenética clássica relataram em um limitado número de casos, a presença de r(1) (Pandis et al., 1990; Polito et al., 1999). Os autores identificaram estas alterações após cultura das células tumorais e por essa razão associaram este achado como alterações secundárias nos Leiomiomas Uterinos. Outros achados envolvendo deleções do braço curto do cromossomo 1 foram identificados em três amostras, incluindo os rearranjos del(1)(p36.33-p11.2) (amostra 046T_A), del(1)(p36.31-32.3) (amostra 629T_B) e del(1)(p36.12-p33) (amostra 1015T). Em 2006, Christacos et al. revisaram 800 LU cariotipados e identificaram um grupo de nove casos (1,1%) apresentando cariótipos com perda do braço curto do cromossomo 1. Utilizando CGH array, nós demonstramos que essa alteração foi observada em 6,3% dos casos, demonstrando o potencial desta metodologia em identificar rearranjos que não são detectados por métodos citogenéticos convencionais. Christacos et al (2006) também relataram que entre os casos com alterações envolvendo o cromossomo 1, quatro foram diagnosticados como LU com hiperplasia e atipia nuclear. O perfil transcricional destes tumores foi similar ao relatado em LMS, sugerindo um possível mecanismo patogénico em comum entre estas duas entidades. Antes da realização dos experimentos, todos os prontuários das pacientes foram revistos e os casos que desenvolveram LMS foram excluídas do estudo.

Outra alteração genômica identificada foram ganhos envolvendo o braço curto do cromossomo 19, com região mínima comum em +(19)(p13.3-p12). Esta anormalidade foi observada em dez amostras (300T_A, 301T_B, 301T_C, 304T, 317T_C, 629T_A, 631T_B, 643T_A, 733T_A, 733T_B). Levy et al (2000) já haviam descrito esta anormalidade em 3/12 amostras de Leiomiomas Uterinos. Ganhos em 19p parecem ser uma anormalidade comum entre os tumores benignos, já observadas em angiofibroma nasofaríngeo juvenil (Heinrich et al, 2007) e adenomas metanéfricos (Pan e Epstein, 2010). Contudo, foram

descritos ganhos nesta região também em tumores malignos, como carcinoma ovariano (Micci et al., 2004) e câncer de mama (Yu et al., 2009). O cromossomo 19 é rico em famílias gênicas duplicadas e agrupadas que pode torná-lo mais propenso a variações no número de cópias (Grimwood et al., 2004). Entretanto, segundo os autores, o cromossomo 19 apresenta pelo menos 97 genes com traços mendelianos, sendo que a maioria deles está envolvida em doenças genéticas raras, que foram identificadas em regiões específicas do cromossomo por mapeamento genético realizado em famílias.

Em relação à presença de alterações genômicas significativas, no qual são aplicados critérios estatísticos (picos significativos com $P < 0,05$; presença da alteração em $>10\%$ dos casos; exclusão de CNVs presentes em $>5\%$ da população referência), foram identificadas 45 CNVs com uma média de $6,76 \pm 6,0$ alterações por caso. Estes dados indicam que em relação a esta categorização, 10-25% das amostras deste estudo apresentaram alterações genômicas consistentes, com regiões mínimas comuns em pelo menos oito amostras. Este dado é concordante com a maioria dos estudos citogenéticos que mostra a presença de alterações consistentes entre 5 a 25% entre amostras avaliadas (Rein et al., 1991; Hodge et al., 2007). Entre as 45 alterações genômicas, 16 delas foram descritas previamente em estudos que utilizaram CGH cromossômica e CGH array (Packenham et al., 1997; Levy et al., 2000; Canevari e Rogatto et al., 2007; Vanharanta et al., 2007; Hodge et al., 2009). Portanto, neste estudo foi possível descrever 29 novas regiões associadas aos Leiomiomas Uterinos. Entretanto, nenhuma delas mostrou associação significativa com os dados clínicos. Provavelmente, estas alterações estejam mais diretamente associadas à biologia tumoral do que com fatores de risco associados a estes tumores.

Alguns genes que exercem importantes funções relacionadas às características dos LU foram selecionados para discussão. Primeiramente serão apresentados os achados identificados neste estudo e que também foram previamente reportados em literatura. Logo após a citação de cada alteração será apresentado em parênteses, o tamanho da

alteração, o número de genes e número de CNVs descritas no DGV. Para maiores informações, no Anexo 4 encontram-se os genes e miRNAs mapeados na região. Neste estudo, um subgrupo de amostras foi investigado para o perfil de expressão gênica, portanto, os genes que serão apresentados nesta seção foram indiretamente associados quanto ao seu padrão de expressão.

Relatos em literatura demonstram que um subgrupo de LU apresenta rearranjos envolvendo 6p21, sugerindo que o gene *HMGA1* é um relevante alvo de estudos nestes grupos de amostras (Kazmierczak et al., 1996). No presente estudo, foi verificado ganhos em 6p21.32 (169Kb; 17 genes; 23 CNVs) em 10/80 casos (12%). O gene *HMGA1* (codifica proteínas do tipo HMGA) está mapeado a aproximadamente 2Mb da região mínima comum, indicando que ele não é um potencial alvo de alterações no nosso estudo. Os LU com rearranjos envolvendo esta região apresentam expressão aumentada de *HMGA1* (Sornberger et al., 1999; Tallini et al., 2000; Nezhad et al., 2010). Em nosso estudo de expressão gênica, este gene não foi classificado entre os genes significantes. O *hsa-miRNA-1236* está mapeado nesta região e tem como genes alvos *CDK2NA*, o qual apresentou expressão gênica diminuída.

A região 11p15.5 (1Mb; 48 genes; 122 CNVs) foi envolvida em ganhos em 19/80 casos (24%). Estudos prévios de nosso grupo demonstraram alterações nessa região em amostras de Leiomiomas Uterinos (Canevari e Rogatto, 2007). O gene *HRAS* é um candidato mapeado nesta região pertencente à família de oncogenes RAS. A proteína codificada por este gene exerce funções chave nas vias de transdução de sinal (Lea et al., 2007). Níveis aumentados da p21 (proteína codificada por *HRAS*) foram encontrados em amostras de LU com ganhos em 11p15.5 (Fotiou et al., 1992). A região 11p15.5 é alvo frequente de instabilidade genômica, pois está localizada na região telomérica do cromossomo (Murnane et al., 2010). Expressão diminuída da p21 parece ser um evento característico de tumores benignos e já foi associada com leiomiomas gastrointestinais extra-uterinos (Cai et al., 1995). O gene *IFITM1*, também mapeado nesta região, codifica

uma proteína associada a diferenciação celular. Expressão aumentada deste gene foi associada com escape do sistema imune no carcinoma gástrico (Han et al., 2011). No presente estudo, este gene também apresentou expressão aumentada.

A região 1p36.33–p36.31 (3Mb; 32 genes; 233 CNVs) foi detectada como envolvida em perdas em 16/80 casos (20%). O gene supressor tumoral *TP73* está mapeado nesta região e codifica a proteína p73, que é um membro da família p53 de fatores de transcrição, em resposta ao estresse celular. Esta proteína tem sido testada farmacologicamente e gerado resultados promissores para terapia do câncer, pois coordena redes de sinalização associadas com proliferação, diferenciação e morte celular (Bisso et al., 2011). Segundo os autores, este gene pode atuar de forma ambígua, agindo tanto como supressor tumoral, como oncogene. O gene *TP73* não foi identificado como significativo entre os dados de expressão no presente estudo. Outro gene interessante mapeado nesta região é o *ARHGEF16*. Este gene codifica uma proteína envolvida em interações proteína-lipídeos. Em melanoma, foi descrita expressão diminuída do *ARHGEF16* associado a hipermetilação da sua região promotora (Bonazzi et al., 2009), se tornando um novo candidato a supressor tumoral. No presente estudo, este gene também apresentou expressão diminuída, sendo associado a perda da região 1p36.33–p36.31.

Outra região envolvida em perda foi 7q31.33 (1Mb; 2 genes; 60 CNVs), identificada em 13/80 casos (16,3%). Nesta região está mapeado o gene *POT1*, que atua na manutenção e proteção dos telômeros. A perda do telômero é acompanhada pela ativação do mecanismo de *checkpoint* mediado por ATR e Chk1, mas não é causada pela parada do *checkpoint* (Churikov e Price, 2008). Pitt e Cooper (2010) demonstraram que a perda da função de POT1 durante G1 leva a rápida erosão do telômero durante a progressão de S/G2. Disfunções teloméricas podem levar a anormalidades cromossômicas, resposta a danos ao DNA e até mesmo ao câncer (Xie et al., 2011). Bonatz et al. (1998) realizaram um estudo para investigar os mecanismos de encurtamento dos telômeros em LU, que possuem características benignas. Os autores identificaram encurtamento telomérico

significativo nos LU comparados ao miométrio adjacente, contudo, este mecanismo não foi associado com o tamanho do tumor. Outro achado destes autores foi que LU múltiplos apresentaram diferentes taxas de encurtamento telomérico, confirmando a origem clonal e independente destes tumores. Portanto, a perda da função do gene *POT1* poderia estar associada com mecanismos de encurtamento de telômeros em Leiomiomas Uterinos. No presente estudo, este gene não foi classificado como significativo nos estudos de expressão gênica. Embora estas regiões genômicas tenham sido previamente descritas como alteradas em LU, os potenciais genes candidatos ainda não foram descritos.

Entre os ganhos genômicos ainda não relatados em literatura, foi verificado ganhos envolvendo 8p12-p11.23 (1Mb; 21 genes; 11 CNVs) em 13/80 casos. O gene *STAR* está mapeado nesta região e codifica uma proteína que desempenha um importante papel na regulação da síntese dos hormônios esteróides, permitindo a conversão do colesterol em pregnenolona. Polimorfismos do gene *STAR* foram associados com tumores estrógeno-dependentes, como o carcinoma de endométrio (Terry et al., 2010). Outro gene candidato importante mapeado nesta região é o *FGFR1*. A proteína codificada por este gene (FGFR1) é uma tirosina-quinase que interage com o fator de crescimento de fibroblastos, desencadeando uma cascata de sinais e influenciando a divisão e a diferenciação celular (Acevedo et al., 2009). No útero, a expressão de FGFR1 é suprimida na primeira fase lútea (proliferativa) de mulheres normais, mas não em mulheres com sangramento relacionado à Leiomiomas Uterinos. Estes resultados reforçam o papel do sistema receptor-ligante bFGF na patogênese da hemorragia associada aos LU e podem ter implicações na fertilidade e contracepção (Anania et al., 1997). Recentemente, Bhide et al. (2010) relataram que o *brivanib alaninato* é um inibidor seletivo do VEGFR2 e do FGFR1 tendo uma atividade anti-tumoral em um grande espectro de modelos em xenoinxertos (carcinomas de células não pequenas de pulmão, carcinoma de mama e cólon). Os autores demonstraram que o *brivanib* é um agente anti-tumoral efetivo em modelos pré-clínicos em uma variedade de doses e a eficácia é acompanhada por alterações nas atividades

celulares e vasculares. Shiang et al. (2010) utilizaram CGH array e microarrays de expressão, para investigar o número de cópias e a expressão do transcrito *FGFR1* e a expressão da sua proteína em 21 linhagens de câncer de mama. Os autores demonstraram que em células de carcinoma mamário com amplificação do *FGFR1*, o *brivanib* diminui a auto-fosforilação do receptor, inibindo a atividade tirosina-quinase induzida pelo bFGF, e redução da fosforilação do ERK e AKT. As linhagens de câncer de mama com amplificação de *FGFR1* e expressão aumentada de sua proteína foram mais sensíveis a inibição do crescimento pelo *brivanib* do que as células sem amplificação deste gene. Os autores sugeriram que a amplificação do gene ou expressão aumentada da proteína FGFR1 em tumores de mama pode ser um indicador para o tratamento com *brivanib*, o qual teria um efeito anti-proliferativo em adição ao seu efeito anti-angiogênico. Na análise de expressão gênica global, este gene apresentou aumento de expressão. Assim, ganhos do gene *FGFR1* poderiam ter importantes implicações para o desenvolvimento de alvos terapêuticos que poderiam contribuir para um dos sintomas clínicos relevantes associados aos Leiomiomas Uterinos.

Outra região do cromossomo 8 envolvida em ganhos foi 8q24.3 (2Mb; 80 genes; 173 CNVs), observada em 12/80 casos (15%). O gene *CYP11B2*, mapeado nesta região, codifica uma enzima que cataliza reações envolvidas com o metabolismo de drogas, síntese de colesterol, esteróides e outros lipídeos. Alterações neste gene podem levar ao aumento da produção de hormônios esteróides associados com câncer de mama e de próstata, hipertensão arterial e fibrose cardíaca (Hakki et al., 2006). Este gene não foi detectado como significativo na análise de expressão gênica global. Entretanto, outros genes mapeados nesta região encontravam-se com diminuição da expressão (*C8orf51*, *PYCRL*, *FAM83H*, *CYC1*, *BOP1*, *NFKBIL2*, *MFSD3*, *RECQL4*, *LRR14*). Esta é outra região que apresenta grande instabilidade genômica por estar localizada em proximidade a telômeros (Huppi et al., 2008), o que explicaria em partes, as discrepâncias entre os achados.

O ganho de 11q12.3-q13.2 (~6Mb; 131 genes; 76 CNVs) foi o achado mais frequente entre as amostras, detectado em 20/80 casos (25%). O gene *ESRRA*, aí mapeado, codifica uma proteína que está intimamente relacionada com o receptor de estrógeno e tem sido associada com o câncer cervical (Choi et al., 2007). Esta proteína age como um regulador de transcrição sítio-específico e interage também com o fator de transcrição TFIIB por contato direto proteína-proteína. Este gene não foi classificado entre os genes significantes nos achados de expressão gênica global. O gene *CD248* codifica a proteína endosialina (CD248) que é considerada um marcador de superfície endotelial da vasculatura em tumores (Teicher, 2007). Recentemente, Smith et al. (2011) associaram a expressão de CD248 com a produção de matriz fibrótica em doença crônica dos rins (CKD). Os fibroblastos do estroma são as células primárias que atuam na produção da matriz fibrótica no rim. A expressão aumentada de CD248 por fibroblastos foi associada com a progressão desta doença. No presente estudo, este gene apresentou expressão aumentada, sendo um candidato para ensaios sobre o fenótipo fibrótico dos Leiomiomas Uterinos. Outro gene mapeado nessa região é o *PPP2R5B*. Este gene codifica a proteína PPP2R5B que regula negativamente as vias de resposta a danos no DNA, entre elas a via ATR e Chk1, já mencionadas acima (Marsh et al., 2007). Para nosso conhecimento, foram identificados apenas dois estudos avaliando mutações em *PPP2R5B*, nos quais falharam em associar este evento com neoplasia endócrina múltipla e câncer de mama familiar (Forbes et al., 1997; Marsh et al., 2007). No presente estudo, este gene apresentou expressão aumentada. Portanto, parece que em LU outros mecanismos, como ganhos genômicos, estariam associados a mudanças nas vias de resposta a danos no DNA.

Ganhos em 12q13.2-q13.3 (1.8Mb; 75 genes; 45 CNVs) foram detectados em 14/80 casos (17,50%). Esta provavelmente foi a região que mais apresentou genes que podem exercer funções importantes nos LU, tais como, processo metabólico de lipídeos (*APOF*), regulação positiva da transcrição (*CDK2*), ativadores da transcrição (*STAT2*, *STAT6*) e proteínas motoras do músculo liso (*MYL6*, *MYL6B*, *MYO1A*) entre outros que serão

apresentados, pois apresentaram concordância com achados de expressão dos mesmos transcritos. O gene *HSD17B6* codifica uma oxiredutase envolvida no metabolismo de andrógenos e estrógenos (Huang e Luu-The, 2000). Assim como os receptores de estrógeno e andrógeno, a proteína codificada por este gene também responde a ação destes hormônios (Baker, 2001). Portanto, a atividade aumentada deste gene poderia levar ao aumento da responsividade de sua proteína aos estrógenos, contribuindo para o desenvolvimento dos Leiomiomas Uterinos. Aqui foi detectado aumento do transcrito *HSD17B6*. Para o nosso conhecimento, apenas polimorfismos deste gene foram associados a características metabólicas em mulheres com Síndrome do Ovário Policístico (Jones et al., 2006; Ke et al., 2010). Outro gene interessante é o *ITGA7*. Este gene codifica a alfa(7)beta(1)-integrina, no qual sua expressão aumentada é geralmente encontrada no músculo esquelético em pacientes com Distrofia Muscular de Duchenne (DMD). Estudo *in vivo* revelou que a *laminina 111*, que se liga a alfa(7)beta(1)-integrina, é um potente alvo terapêutico em camundongos com DMD, pois foi associada com restauração da atividade normal deste gene no músculo (Rooney et al., 2009). O aumento da atividade deste gene poderia estar associado ao aumento da deposição da matriz extracelular em Leiomiomas Uterinos. No presente estudo, este gene também apresentou expressão aumentada. Outro gene associado à matriz extracelular nesta região é o *MMP19*, que também apresentou aumento de expressão no presente estudo. A proteína MMP19 é uma enzima que atua no remodelamento da matriz extracelular, no qual sua atividade aumentada foi associada com células proliferativas em glioma e melanoma (Lettau et al., 2010; Müller et al., 2010). Análises citogenéticas moleculares têm demonstrado que o gene *HMGA2*, mapeado em 12q14.3 (~10Mb desta região mínima comum), pode desempenhar um importante papel nos Leiomiomas Uterinos (Gross e Morton, 2001). Como apresentado acima, as proteínas da família HMG desempenham funções moleculares envolvidas na manutenção e arquitetura cromossômica, inclusive, modulando o acesso de fatores de transcrição ao DNA (Hess et al. 1998). A ativação do gene *HMGA2* em LU foi associada a translocação

entre os cromossomos 12 e 14 (Nilbert et al., 1990; Hennig et al., 1999). No presente estudo, este gene não foi classificado entre os diferencialmente expressos, porém, outros mapeados em 12q parecem contribuir para o desenvolvimento de LU além do gene *HGMA2*.

Ganhos 17p13.1 (74Kb; 6 genes; 20 CNVs) foi detectado em 8/80 casos. O gene *SHBG* aí mapeado codifica uma proteína de ligação a esteróides, que participa da regulação das respostas destes hormônios. Altos níveis do transcrito *SHBG* e expressão positiva da sua proteína foram associados com Leiomiomas Uterinos (Misao et al., 1995; Sendemir et al., 2008). Assim, é possível que este aumento de expressão ocorra por mecanismos de ganhos genômicos, como demonstrado no presente estudo. Pelos critérios utilizados, este gene não se encontrava entre os diferencialmente expressos. O gene *ATP1B2*, também mapeado nesta região, codifica uma K-ATPase que atua na motilidade do tecido muscular (Hundal et al., 1994). Floyd et al. (2010) relataram aumento da expressão do transcrito e da proteína codificada por este gene em amostras de Leiomiomas Uterinos. Este gene também não foi detectado entre os diferencialmente expressos na análise de expressão gênica global. Embora os níveis dos transcritos não tenham sido observados como alterados na análise de expressão dos transcritos na comparação com os ganhos detectados em 17p13.1, há a necessidade de validação dos resultados para confirmar o achado, pois a função deste gene parece ser importante para o aumento da motilidade do músculo liso durante o crescimento dos Leiomiomas Uterinos.

A região 20q13.33 (316Kb; 25 genes; 28 CNVs) foi envolvida em ganhos em 8/80 casos (10%). Nesta região, foi observado um *cluster* de miRNAs (*hsa-mir-941-1*, *hsa-mir-941-2*, *hsa-mir-941-3*, *hsa-mir-1914*, *hsa-mir-647*). Após uma busca nos bancos de dados de predição gênica por miRNA (*TargetScan* - <http://www.targetscan.org> e *PicTar* - <http://pictar.mdc-berlin.de/>), alguns genes alvos dos miRNAs mapeados em 20q13.33 parecem ter um importante papel em LU, no qual dois deles foram selecionados como candidatos. O gene *hsa-mir-941* têm como um dos alvos a sequência do gene *MLL2* (12q12-

q14). Portanto, ganhos genômicos de *hsa-mir-941* poderiam levar a diminuição da atividade do gene *MLL2*. A proteína *MLL2* atua como co-ativador do complexo transcricional do receptor de estrógeno alfa (Mo et al., 2006). Em nosso grupo, os transcritos dos genes receptores de estrógeno alfa, beta e de progesterona e a proteína codificada por estes genes, foram avaliados em amostras de LU e miométrio adjacente (Cirilo et al., em progresso). Treze de 47 amostras (~28%) apresentaram expressão diminuída de *ESR1* comparadas ao miométrio adjacente e 6/90 amostras (~7%) apresentaram marcação nula ou fraca para ER-alfa, enquanto que a maioria das amostras apresentou expressão normal à aumentada do transcrito (34/47 - 72%) e marcação moderada a forte da proteína (84/90 - 93%). Portanto, a perda da atividade de *MLL2* por seu miRNA regulador envolvido em ganhos genômicos, poderia ser um dos mecanismos indiretos associados a diminuição de *ESR1* em um subgrupo de Leiomiomas Uterinos. O gene *hsa-mir-647* têm como um dos alvos preditos a sequência do gene *ARNT*. O gene *ARNT* modula a expressão de genes envolvidos em proliferação e diferenciação celular (www.targetscan.org; www.ncbi.nlm.gov) e parece desempenhar um importante papel em Leiomiomas Uterinos. Khorram et al. (2002) investigaram a expressão deste gene em LU e observaram expressão diminuída em relação ao miométrio adjacente. Contudo, no presente estudo, este gene não foi classificado entre os diferencialmente expressos.

Perdas em 4q28.3 (3Mb; 2 genes; 121 CNVs) foram identificadas em 12/80 casos (15%), sendo um dos mais frequentes achados deste estudo. O gene *PCDH10* descrito nesta região, codifica uma protocaderina que atua na junção célula-célula (Wolverton e Lalande, 2001). A região promotora do gene *PCDH10* encontra-se hipermetilada em carcinomas nasofaríngeos, de esôfago, cervical e de próstata. A expressão diminuída deste gene foi associada a aumento da proliferação celular e progressão tumoral (Ying et al., 2006; Narayan et al., 2009; Li Z et al. 2011). Neste estudo, o gene *PCDH10* não foi classificado entre os genes diferencialmente expressos, contudo, a perda da sequência

deste gene poderia ser um dos eventos associados aos Leiomiomas Uterinos. A análise do padrão de metilação como um segundo evento no desenvolvimento destes tumores poderia estar associada a perda da capacidade de regulação da proliferação das células do miométrio. Estes dados precisam ser ainda validados por outros métodos.

Perdas em 5q31.1-q31.2 (1Mb; 19 genes; 35 CNVs) foram detectadas em 14/80 casos (17,5%). Nesta região, encontra-se o gene *TGFBI*, que codifica uma proteína que modula a adesão celular (Choi et al., 2010). A expressão alterada dos genes participantes da via de sinalização do TGF-beta parece ser um dos mecanismos centrais de doenças com características fibróides, incluindo os Leiomiomas Uterinos (Hoffman et al., 2004; Luo et al., 2005b). Para nosso conhecimento, não existem relatos da ação deste gene em Leiomiomas Uterinos. A expressão diminuída deste gene foi associada com insulinomas, que são tumores benignos do pâncreas (Nabokikh et al., 2007). Estes tumores são responsivos a ação do TGF-beta na indução da proliferação celular, diferenciação e apoptose, mesmo mecanismo identificado nos Leiomiomas Uterinos. Portanto, a atividade reduzida de *TGFBI* poderia estar associada a perda da indução da apoptose e aumento da proliferação celular em Leiomiomas Uterinos. Neste estudo, o gene *TGFBI* não foi identificado como diferencialmente expresso. Outro gene alterado nesta região é *CXCL14* que codifica uma citocina envolvida em respostas inflamatórias (Hromas et al., 1999). Sleeman et al. (2000) avaliaram a expressão desta proteína em tecidos tumorais e normais de útero (LU e miométrio normal), cérebro, intestino, mama e rim. Os autores identificaram que em carcinomas de mama a expressão da proteína *CXCL14* estava aumentada, enquanto que em LU estava reduzida. Os autores associaram este achado com a diferente biologia de tumores de mama e LU, no qual os tumores de mama possuem características invasivas, enquanto que LU apresentam um fenótipo benigno. No presente estudo, este gene não foi classificado como diferencialmente expresso pelos critérios de análise utilizados.

Desde 1991, os estudos citogenéticos de LU propuseram a existência de subgrupos de tumores com anormalidades cromossômicas específicas envolvendo principalmente os cromossomos 6, 7, 12 e 14 (Hu et al., 1991; Ligon e Morton, 2000; Hodge et al., 2007). Contudo, o presente estudo revelou outros subgrupos de tumores apresentando anormalidades genômicas não previamente descritas. A análise de agrupamento hierárquico não supervisionado dos dados genômicos permitiu a identificação de dois grupos de amostras com diferentes frequências de CNVs. Ganhos em 2q35, 7q22.1, 19p13.12-p13.11 e 19q13.32-q13.33 foram observados simultaneamente entre 11/42 amostras do Grupo 1 ($P < 0.01$) e perdas em 7q22.1-q22.2 foi observada exclusivamente entre 9/38 amostras do Grupo 2 ($P < 0.01$).

Na região 2q35 (2Mb; 26 genes; 27 CNVs) estão mapeados genes que desempenham importantes papéis no processo de desenvolvimento tumoral. O gene *IGFBP5*, codifica a proteína IGFBP5 que desempenha várias funções associadas a remodelação tecidual em resposta aos estrógenos, como crescimento, morte e mobilidade celular. A expressão aumentada deste gene foi associada com neoplasia intra-epitelial cervical e câncer de mama (Hou et al, 2009;. Taylor et al, 2010). No presente estudo, foi detectada expressão aumentada do gene *IGFBP5*. Os genes *CXCR1* e *CXCR2* codificam proteínas que agem em resposta a IL-8 e foram caracterizados em células tumorais e endoteliais (Waugh e Wilson, 2008). No entanto, Syssoev et al. (2008) avaliaram a expressão gênica de *CXCR1* e *CXCR2* em LU e não identificaram diferenças estatisticamente significativas entre os tumores e miométrio normal adjacente. No presente estudo, estes genes não foram classificados entre os diferencialmente expressos.

Ganhos em 7q22.1 (3Mb; 98 genes; 165 CNVs) foi um interessante achado deste estudo, uma vez que deleções nesta região têm sido amplamente relatada em Leiomiomas Uterinos (para revisão Hodge et al., 2009). Perdas envolvendo esta região também foram identificadas em outro subgrupo de amostras, as quais serão discutidas a seguir. Em 2003, foi publicada a sequência completa do cromossomo 7 (Scherer et al., 2003). Os autores

descreveram vários *amplicons* em 7q22.1, nos quais 29 foram originados por pontos de quebra cromossômica e 24 desses rearranjos foram envolvidos com doenças. Nesta região foi descrito o sítio frágil *FRA7F* que foi associado à pacientes com autismo. Este achado sugere que os ganhos identificados em 7q22.1 podem estar associados a essa região de instabilidade genômica. Entre os genes aí mapeados, oito foram diferencialmente expressos em nosso estudo, contudo, apresentaram expressão gênica diminuída. Contudo a literatura relata a associação de alguns genes com outros tumores. Em 7q22.1, estão mapeados diferentes *clusters* de genes, como por exemplo, os genes *ARPC1A* e *ARPC1B*, que atuam no ciclo celular. Laurila et al. (2009) utilizando CGH array relataram um *amplicon* recorrente em 7q22-q22 em tumores pancreáticos. Os autores demonstraram que o gene *ARPC1A*, que atua como regulador de migração celular pode ser um novo alvo terapêutico nestes tumores. Em LU, foi detectada a expressão diminuída do gene *ARPC1A*. Outros *clusters* foram identificados nesta região como de genes reguladores de transcrição (*ZNF789*, *ZNF394*, *ZNF498*, *ZNF655*) e de genes associados com o processo metabólico de esteróides (*CYP3A4*, *CYP3A5*, *CYP3A7*, *CYP3A43*). O gene *SERPINE1*, aí mapeado, parece ser crucial para o desenvolvimento dos LU, pois é um dos genes que responde a sinalização do TGF-beta (Leppert et al., 2006). Este gene codifica a proteína PAI1 que atua como o principal inibidor de ativação do plasminogênio tecidual (tPA) e uroquinase (UPA), inibindo a fibrinólise. Sourla et al. (1996) identificaram altos níveis de mRNA deste gene em amostras de LU nos casos com ausência de perda de heterozigose em 7q22. Estes dados sugeriram que as deleções em 7q22 não são os únicos eventos críticos associados aos Leiomiomas Uterinos. Na análise de expressão gênica, este gene não foi classificado entre os significantes. Entretanto, alterações neste gene parecem ter um importante papel nos LU, pois também tem sido alvo de estudos funcionais. Ding et al. (2004) trataram linhagens de LU e células do músculo liso com análogos de GnRH e TGF-beta para investigar a ação destas moléculas na regressão dos Leiomiomas Uterinos. Os autores relataram que a ação destas substâncias aumentaram os níveis de ERK1/2 fosforilada no

núcleo das células das duas linhagens, resultando na regulação de *FOS* e *JUN*. Este tratamento também reduziu a expressão de fibronectina, colágeno tipo 1 e PAI-1 nas duas linhagens. Os autores concluíram que terapias com análogo de GnRH e TGF-beta podem influenciar a redução do crescimento em Leiomiomas Uterinos. Outro gene identificado nesta região é o *PCOLCE*. Este gene codifica uma glicoproteína que direciona a clivagem enzimática do colágeno tipo I e aumenta a atividade de proteinase C. Ligon et al. (2002) avaliaram o número de cópias (FISH) e o nível do transcrito (RT-PCR) do gene *PCOLCE* em uma série de LU previamente cariotipados. Eles identificaram diferentes proporções de células apresentando a deleção em 7q, entretanto, a expressão do *PCOLCE*, não foi estatisticamente significativa. Este gene não foi classificado com diferencialmente expresso no presente estudo. Os ganhos descritos em 7q22.1, podem ser característicos de um subgrupo de Leiomiomas Uterinos merecendo validações futuras.

Os ganhos detectados em 19p13.12-p13.11 foram discutidos acima, quando apresentadas as anormalidade envolvendo grandes segmentos genômicos (>10Mb). O ganho em 19q13.32-q13.33 compreendeu uma região com ~6Mb, mapeando 315 genes e 178 CNVs. Esta região apresentou a maior taxa de sobreposição de CNVs identificadas neste estudo (100%). Isso indica que esta região pode ser um local de grande instabilidade genômica. Contudo, ganhos em 19q13 são frequentemente relatados em tumores humanos, como câncer colorretal (Sayagués et al., 2010), carcinoma ovariano (Bayani et al., 2011), carcinoma urotelial (Nishiyama et al., 2011) e até mesmo em neoplasias hematológicas, como linfoma não-Hodgkin (Slovak et al., 2011). Como discutido acima, o cromossomo 19 é propenso a variações no número de cópias. Portanto, juntas, as regiões 19p13.12-p13.11 e 19q13.32-q13.33 indicam que ganhos do cromossomo 19 são eventos comuns nas etapas da carcinogênese humana.

A perda envolvendo 7q22.1-q22.2 (855Kb; 9 genes; 24 CNVs) foi a mais frequente entre as amostras do Grupo 2. Entre as 15 amostras deste estudo que apresentaram deleções em 7q, nove pertenciam ao Grupo 2. Vários relatos indicam tamanhos variáveis

desta deleção (para revisão Ptacek et al., 2007). A Tabela 6 mostra todas as deleções identificadas no braço longo do cromossomo 7 distribuídas entre as amostras. Como já apresentado acima, deleções envolvendo 7q provavelmente é um dos eventos mais comuns em LU esporádicos (Hodge et al., 2009). Estudos de clonalidade em amostras de LU com deleções em 7q revelaram que estas alterações surgem durante a progressão tumoral, sugerindo que os genes mapeados nesta região são cruciais para o desenvolvimento ou manutenção dos Leiomiomas Uterinos (Sargent et al., 1994; Xing et al., 1997). Takahashi et al. (2001) mostraram que a deleção em 7q confere características biológicas distintas entre tumores com outras anormalidades cromossômicas. No presente estudo, não foi observada uma diferença significativa entre dados clínicos das amostras com deleção 7q e sem deleções em 7q (dados não mostrados). Vários estudos foram realizados com o objetivo de identificar um gene supressor tumoral nesta região diretamente associado com Leiomiomas Uterinos. Os estudos que utilizaram a estratégia de análise semi-integrada ou de validação por microarrays de expressão em LU foram feitos com o objetivo de descrever um gene crítico mapeado 7q (Vanharanta et al, 2005; Hodge et al., 2009). Hodge et al. (2009) identificaram os genes *HPB1* e *RINT1* como genes candidatos a perdas associados aos Leiomiomas Uterinos. Estes achados reforçaram a idéia da existência de um possível gene supressor de tumor associado a um subgrupo de LU com deleções em 7q. O *RINT1* também foi identificado como envolvido em perdas no presente estudo, contudo, não foi classificado entre os genes significantes nos resultados de expressão gênica. A proteína codificada por este gene forma complexos com as proteínas Rad50, envolvida na manutenção estrutural dos cromossomos, participando do reparo de quebras em dupla-fita de DNA, ativação do *checkpoint* do ciclo de celular e da manutenção dos telômeros (Xiao et al., 2001). Além do gene *RINT1*, o presente estudo revelou outros genes que serão apresentados a seguir. O gene *LHFPL3*, codifica um membro da família *lipoma HMGIC fusion partner* (LHFP), que são proteínas que atuam na diferenciação, proliferação e formação da matriz extracelular. Existem poucos estudos

sobre *LHFPL3*, porém, um deles (Ptacek et al., 2007) associou este gene com a deleção em 7q22 em Leiomiomas Uterinos. Ptacek et al. (2007) compilaram as informações de todas as deleções em 7q já descritas e definiram a região 7q22 como a mínima comum. Nesta região, os genes *LHFPL3* e *ORC5L* apareceram como preferencialmente alterados, sugerindo que são potenciais candidatos ao desenvolvimento dos Leiomiomas Uterinos. Estes genes não foram classificados como diferencialmente expressos na análise de expressão gênica global, entretanto, devido a sua função poderiam ser selecionados para validação futura.

Entre 27 amostras de LU avaliadas previamente por nosso grupo para alterações no número de cópias no cromossomo 7 utilizando a metodologia de FISH (Perez, 2004), 20 delas também foram avaliadas neste estudo atual. Os resultados da FISH revelaram que 9/27 amostras (33%) apresentaram a deleção em 7q22.3-q31.1. Quando os resultados de FISH e CGH array foram comparados, quatro amostras com a deleção revelada por FISH, também apresentaram esta deleção por CGH array (769T, 844T_A, 845T_A e 853T) e três amostras (711T, 729T_A e 947T) com a deleção revelada por FISH não apresentaram essa alteração por CGH array. Estas diferenças observadas entre as metodologias podem ser explicadas pelo próprio mapeamento da sonda de FISH. A região mínima comum identificada por CGH array foi a 7q22.1-q22.2 e a sonda RP11-46J20 mapeia a região 7q22.3-q31.1, que dista ~2Mb da região mínima alterada revelada por este estudo. Quando foram aplicados critérios menos estrigente na análise de CGH array, foram identificadas 29 regiões heterogêneas de deleções em 7q em 49 amostras (Tabela 6), contudo, os mesmos quatro casos (769T, 844T_A, 845T_A e 853T) se mantiveram concordantes com os achados de FISH. A metodologia de CGH array possui um maior poder de resolução para identificação de microdeleções (sondas de oligonucleotídeos), enquanto que as sondas de FISH (sondas de BACs) possuem menor resolução. Heterogeneidade no tamanho das deleções envolvendo 7q é um evento frequente, dificultando a delimitação de uma região mínima comum em 7q e a frequência deste

evento entre os Leiomiomas Uterinos (Vanharanta et al., 2005; Hodge et al., 2009). No entanto, mesmo em baixa frequência, a deleção em 7q22.3-q31.1 identificada por FISH, foi evidenciada por CGH array. Os genes mapeados nessa região (*COG5*, *DUS4L*, *BCAP29*, *SLC26A4*) também apresentaram perda de seqüências genômicas por CGH array. O gene *COG5* codifica uma das oito proteínas do complexo oligomérico de Golgi (COG), envolvido com o tráfego de moléculas intracelulares (Ungar et al., 2002). Velagaleti et al. (2010) relataram em uma amostra de LU a fusão entre os genes *HMGA2* e *COG5* resultante da translocação entre os cromossomos 7 e 12, t(7;12)(q31.2;q14.3). Os autores confirmaram por FISH o rearranjo envolvendo o locus *HMGA2* em 80% das células, assim como, por 3'RACE RT-PCR foi possível identificar a fusão do exon 4 do gene *HMGA2* no locus do gene *COG5*. O gene *BCAP29*, também mapeado em 7q22.3-q31.1 codifica a proteína Bap29 que está envolvida no tráfego de proteínas através da membrana celular. A Bap29 é um dos membros de uma grande família de proteínas conhecida como *ABC-transporters*, que já foram estabelecidas como o mecanismo mais comum de resistência ao tratamento por múltiplas drogas em células tumorais (Annereau et al., 2004). Este gene ainda não havia descrito em Leiomiomas Uterinos. Outro gene identificado nesta região é *SLC26A4* que codifica uma proteína de transmembrana envolvida com o transporte de sulfato. Este gene é encontrado frequentemente metilado em linhagens obtidas de pacientes com leucemia mielóide aguda (Kroeger et al., 2008). Uma meta-análise de estudos *genome-wide* confirmou a existência de um locus de suscetibilidade para a osteoartrose em 7q22, no qual foi confirmada por expressão gênica a ação dos genes *COG5*, *DUS4L*, *BCAP29* em tecidos articulares. Nos resultados de expressão gênica do presente estudo, estes genes não foram classificados como significantes.

De uma forma geral, os genes mapeados nas regiões onde foram detectadas CNVs significativas desempenham funções reconhecidamente associadas com os LU, como proliferação celular e processos fibróides. Contudo, quando as funções destes genes foram plotadas na forma de porcentagem, foi observada uma maior frequência de genes em

outras categorias funcionais, como regulação da transcrição, *splicing* de RNA, processo metabólico de lipídeos, regulação da transcrição da RNA polimerase II (RNA Pol II) e processo metabólico de esteróides. Na verdade, de uma forma indireta, genes envolvidos com estas funções poderiam também participar das etapas de proliferação celular e formação fibróide dos Leiomiomas Uterinos. Se a maioria dos genes identificados em regiões de CNVs está envolvida em processos transcricionais, como a regulação da transcrição da RNA polimerase II, é provável que a expressão de genes específicos envolvidos em vias clássicas de LU também esteja prejudicada. A RNA polimerase II é um complexo enzimático responsável pela síntese de mRNA em eucariotos. Sendo assim, a regulação dos níveis transcricionais dos genes que codificam as enzimas deste complexo, é crucial para o funcionamento global da expressão gênica. Portanto, é necessária a avaliação dos genes reguladores da RNA polimerase II.

Os genes envolvidos em processos metabólicos de lipídeos e esteróides parecem exercer um importante papel no desenvolvimento dos LU, uma vez que são tumores hormônio-dependentes (Walker e Stewart, 2005). Um estudo recente usando análise transcriptômica demonstrou que alterações das vias de metabolismo de lipídeos não são uma característica exclusiva dos tumores humanos e que outras doenças metabólicas e inflamatórias, como obesidade, aterosclerose e síndrome metabólica, compartilham esta mesma via (Hirsch et al., 2010). Particularmente, os dados sugerem que doenças inflamatórias e auto-imunes aumentariam o risco de desenvolvimento de câncer (Alderton, 2010). De fato, os autores chegaram a estas conclusões utilizando análise de sobreposição de vias, pelo sistema IPA, o mesmo usado no presente estudo. As análises funcionais que serão apresentadas a seguir foram úteis para a descrição de genes envolvidos nestas e em outras funções associadas aos Leiomiomas Uterinos.

As análises funcionais foram realizadas em duas etapas: para os genes mapeados nas regiões de CNVs significativas (análise total) e os genes mapeados nas regiões associadas aos Grupos 1 e 2, determinados por agrupamento hierárquico (análise de

grupos). Estas análises permitiram a descrição de redes gênicas, funções biológicas, doenças associadas e vias canônicas para cada lista de genes.

Entre as funções biológicas associadas com os genes da análise total, expressão gênica foi a função que apareceu entre sete das dez principais redes descritas, concordando com a análise prévia baseada na frequência dos genes presentes nas alterações genômicas. Para o nosso conhecimento, este é o primeiro relato demonstrando essa função associada ao Leiomiomas Uterinos.

Entre as sete redes detectadas como alteradas, três delas (Redes 3, 7 e 9) apresentaram conexões entre os genes regulados e reguladores do complexo da RNA Pol II. Nas três redes, o estrógeno apareceu como um dos principais fatores regulatórios da RNA Pol II. Em relação aos genes reguladores da RNA Pol II, na **Rede 3** foram observados os genes *ASH2L* e *MEN1*. Em nossa lista, o gene *ASH2L* (mapeado em 8p12-p11.23) foi envolvido em ganhos entre 13/80 casos. De acordo com o sistema IPA, este gene já havia sido descrito em LU, contudo não foi possível detectar esse estudo em literatura. Na **Rede 7**, os genes reguladores foram *CDK2*, *CDK8*, *GTF2B*, *POLR2D*, *POLR2G*, *POLR2I*, *POLR2L*, *RDBP*, *MED1* e *MED21*. Entre estes genes, quatro estavam envolvidos em ganhos genômicos: *CDK2* (12q13.2-q13.3-14/80), *POLR2G* (11q12.3-q13.2-20/80 casos) e *RDBP* (6p21.32 – 10/80 casos). Na **Rede 9** os genes que regulam a RNA Pol II foram *MED1* e *MED14*. Por fim, as moléculas reguladas pela RNA Pol II foram *ESR1*, *DHFR*, *POMC*, *SCGB1A1*, *TERC*, *KITLG*, *IL8*, *MAVS*, *HSPA4*, *TP53*, *MUCL1*, *SPTLC1*, *INS* e *CAT*. Estes genes estão preferencialmente associados com proliferação celular, resposta imune e ciclo celular e nenhum deles foi envolvido em ganhos e perdas genômicas. Juntos, estes dados mostram que genes mapeados em regiões de ganhos genômicos podem exercer um importante papel na regulação transcricional dos LU, incluindo genes regulados e reguladores do complexo da RNA Pol II em resposta ao estrógeno. A análise integrada dos dados poderá revelar quais os principais genes afetados pela alteração da regulação deste complexo.

Entre as outras redes gênicas, não foi observada relação direta entre os genes envolvidos no metabolismo de lipídeos. No sistema IPA, a categoria metabolismo de lipídeos foi observada em oitavo lugar [-log(*P* valor 2,6)] na classificação de funções moleculares (Figura 16B). Contudo, genes envolvidos com ciclo celular, duplicação do DNA, recombinação e reparo e sinalização celular foram os mais observados entre as conexões gênicas (Tabela 7). Estas são vias clássicas das múltiplas etapas da carcinogênese (Hoeijmakers, 2001; Mitelman et al., 2007; Holthausen et al., 2010). Quando avaliada a categoria doenças associadas, o câncer apareceu em primeiro lugar [-log(*P* valor 2,6)] (Figura 16A). As doenças seguintes foram doenças do sistema reprodutivo, doença genética, doenças musculares e doenças do tecido conjuntivo. Juntas, as doenças associadas apresentam similaridades aos Leiomiomas Uterinos.

Entre as vias canônicas geradas, as que apresentaram maiores valores de associação foram metabolismo de hormônios esteróides, sinalização de quimiocinas e reparo de quebra em dupla-fita de DNA por recombinação homóloga (Figura 17). A via de metabolismo de hormônios esteróides possui várias moléculas associadas com expressão gênica, metabolismo de drogas e bioquímica de pequenas moléculas. No presente estudo, os genes associados a estas moléculas, foram envolvidos em ganhos genômicos incluindo *CYP11B2* (8q24.3 – 12/80 casos), *CYP21A2* (6p21.32 – 10/80 casos), *HSD3B7* (16p11.2 – 13/80 casos). Como apresentado acima, é possível que a perda da regulação desta via esteja associada a doenças metabólicas prévias ao surgimento dos LU entre as pacientes. A maioria das pacientes deste estudo (41%) apresentou sobrepeso, que indica um risco aumentado de desenvolver obesidade. Contudo, individualmente estas CNVs não foram associadas a nenhum dado clínico, inclusive com IMC. No entanto, polimorfismos nestes genes já foram previamente associados com hipertensão, hiperplasia adrenal congênita e colestatose, respectivamente (Cheng e Xu, 2010; Carvalho et al., 2010; Mizuochi et al., 2010). Portanto, alterações nestes genes, como mutações ou polimorfismos poderiam aumentar o risco de desenvolvimento de doenças metabólicas.

A segunda via canônica identificada foi a de sinalização de quimiocinas. As quimiocinas são citocinas pró-inflamatórias que executam seus processos por meio dos receptores acoplados as proteínas G (GPCRs), regulando uma série de processos celulares (IPA). Dez genes foram associados a estas moléculas e identificados em regiões de ganhos genômicos: *CALM3* e *RRAS* (19q13.32-q13.33 – 15/80 casos), *HRAS* (11p15.5 – 19/80 casos), *MAPK3* (16p11.2 – 13/80 casos), *MAPK11* e *MAPK12* (22q13.33 – 9/80 casos), *MYL2* e *PPP1CC* (12q24.11-q24.12 – 12/80 casos), *PLCB3* e *PPP1CA* (11q12.3-q13.2 – 20/80 casos). Litovkin et al. (2008) identificaram o gene *MAPK12* com expressão aumentada em LU em relação ao miométrio adjacente. Nos resultados de expressão gênica de nosso estudo, este gene não foi classificado entre os significantes. Para o nosso conhecimento, não há relatos dos demais genes associados com Leiomiomas Uterinos.

A terceira via canônica identificada está envolvida com o reparo de quebra em dupla-fita de DNA por recombinação homóloga. As quebras em dupla-fita (DSB) são lesões causadas por genotóxicos, que alteram os mecanismos da duplicação, segregação e transcrição do DNA. O processamento impróprio de DBSs gera instabilidade genômica, que pode estimular as etapas da carcinogênese pela ativação de proto-oncogenes ou inativação de genes supressores de tumor. Os genes identificados nessa categoria estavam envolvidos em ganhos incluindo *BRCA1* (17q21.31 – 12/80 casos) e *LIG1* (19q13.32-q13.33 – 15/80 casos).

Três vias canônicas identificadas já foram associadas ao desenvolvimento dos Leiomiomas Uterinos: sinalização IGF-1, TGF-beta e do receptor de estrógeno. O IGF-1 é um polipeptídeo que promove a proliferação e crescimento celular. O IGF-1 forma um complexo com as proteínas IGFBP1 (1 a 6) que se ligam ao receptor de IGF-1 (IGF1-R) na membrana celular e desencadeia a cascata de sinalização. Esta ligação ativa moléculas da via Ras/Raf/MEK/ERK que induz a transcrição de genes promotores do crescimento e diferenciação celular. No presente estudo, foram identificados dez genes associados a esta via, no qual sete (*BAD*, *HRAS*, *IGFBP5*, *MAPK3*, *PXN*, *RPS6KB2*, *RRAS*) estavam mapeados em

regiões de ganhos e três (*GRB10*, *PRKACB*, *PRKCZ*) em regiões de perdas genômicas (Figura 18). Burroughs et al. (2002) testaram a hipótese sobre a perda da regulação da sinalização IGF-1 em LU e a proliferação celular característica destes tumores. Os autores avaliaram LU de ratos em diferentes fases do ciclo e identificaram que, independente do período, a expressão do gene *IGF-1* sempre se manteve aumentada nos tumores em relação ao miométrio. Estudos de expressão gênica global têm revelado expressão aumentada do gene *IGF-2* (Weston et al., 2003; Catherino et al., 2004a; Arslan et al., 2005) e do gene *IGF-1* em resposta aos estrógenos (Swartz et al., 2005) em Leiomiomas Uterinos. No presente estudo, o gene *IGFBP5* (2q35 – 15/80), que é uma das proteínas que se liga ao IGF-1 desencadeando a sinalização, foi envolvido em ganhos genômicos e apresentou expressão gênica aumentada. O aumento da atividade desta proteína poderia estar associado com o aumento da proliferação celular observada em células do miométrio, induzindo o crescimento dos Leiomiomas Uterinos. O gene *GRB10* (7p12.2 – 10/80) é um dos poucos genes envolvidos em perdas genômicas associados a via IGF-1. A proteína codificada por este gene interage com os IGFs e IGFs e acredita-se que ela desempenhe um efeito regulatório negativo da proliferação celular (Giovannone et al. 2003).

A família de fatores de crescimento transformantes, tipo beta (TGF-beta) são citocinas que controlam a proliferação, diferenciação, migração e apoptose de muitos tipos de células (para revisão Paiva et al., 2010). O sistema básico da sinalização TGF-beta é composto por seus receptores de membrana (TGFBR I e II) e uma família de receptores de substratos diretos (SMADs) que se movem para o núcleo sob ativação da fosforilação. O TGF-beta também pode exibir seus efeitos biológicos independentemente dos SMADs, por meio da ligase TRAF6. A perda da regulação da sinalização do TGF-beta foi diretamente associada com a transcrição de genes envolvidos na reorganização da matriz extracelular e aumento da deposição de colágeno em Leiomiomas Uterinos (Chegini et al., 1999 e 2002; Ding et al., 2004; Moore et al., 2010). No presente estudo, foram identificados cinco genes envolvidos em regiões de ganhos genômicos (*FOXH1*, *HRAS*, *INHBC*, *MAPK3*, *RRAS*) e um

gene envolvido em perda (*SMAD5*) entre os associados a esta via (Figura 19). O gene *INHBC* (12q13.2-q13.3 – 14/80) codifica a cadeia C da inibina, a qual é um importante regulador do sistema reprodutivo feminino, além de ser um membro da família TGF-beta (Ushiro et al. 2006). A expressão diferencial do gene *INHBC* e de sua proteína foi relatada em endométrio normal nas diferentes fases do ciclo menstrual, a qual pode estar associada com a permissividade da implantação do embrião durante a fase proliferativa intermediária do ciclo (Mylonas et al., 2010). Contudo, a expressão aumentada de *INHBC* também foi associada com adenocarcinoma de endométrio (Kimmich et al., 2010; Käufel et al., 2010). Na via gerada neste estudo, o gene *INHBC* encontra-se como ativador da sinalização TGF-beta. O *SMAD5* (5q31.1-q31.2- perdas detectadas em 14/80) foi associado a via de sinalização paralela desencadeada por BMP, que ligada a *INHBC*, poderia estar relacionada com a transcrição de genes envolvidos em Leiomiomas Uterinos. A perda da atividade de *SMAD5* já foi associada ao desenvolvimento tumoral. Camundongos *knockout* para *SMAD5* desenvolveram um tipo de tumor ovariano hormônio dependente, conhecido como tumor de células da granulosa, por meio da ativação de TGF-beta (Middlebrook et al. 2009).

Finalmente, a via de sinalização do receptor de estrógeno foi associada a oito genes mapeados em regiões de ganhos genômicos (*ERCC2*, *G6PC*, *HRAS*, *MAPK3*, *POLR2G*, *POLR2L*, *RRAS*, *TAF6L*) (Figura 20). Estes genes também foram associados às vias de sinalização IGF-1, TGF-beta e com o complexo da RNA Pol II, sugerindo que os estrógenos são os principais desencadeadores destas vias. Basicamente, a sinalização de estrógenos é mediada por seus receptores (ER-alpha e beta), que são fatores de transcrição. A ativação dos receptores pelo estrógeno recruta uma variedade de co-reguladores de ER para formar um complexo que regula positivamente ou negativamente a transcrição da RNA Pol II. Na via gerada, foram detectados ganhos genômicos dos genes da via Ras/ERK que ativam diretamente o ER-alfa e paralelamente, ganhos genômicos em genes do complexo

da RNA Pol II, que sob a ação do estrógeno, regulam a transcrição de genes alvos envolvidos na proliferação celular.

Na análise funcional de grupos foram comparadas as redes, doenças e funções do Grupo 1 (subgrupo de LU com ganhos em 2q35, 7q22.1, 19p13.12-13.11 e 19q13.32-q13.33) e do Grupo 2 (subgrupo de LU com perdas em 7q22.1). Esta análise mostrou que as redes com funções associadas a expressão gênica, proliferação e crescimento celular estava presente apenas no Grupo 1, enquanto que o Grupo 2, apresentou sobreposição de funções com o Grupo 1 como aquelas associadas ao ciclo celular e modificação pós-transcricional de RNA (Figura 21A). Contudo, o Grupo 2 apresentou maiores valores de $-\log(P \text{ valor})$ em relação ao Grupo 1, indicando que os genes do Grupo 2 teriam uma associação mais estreita com as funções relacionadas a organização celular e os genes do Grupo 1 a processos transcricionais e de crescimento. Isto confirma os dados prévios da análise funcional no qual os genes envolvidos em ganhos foram preferencialmente associados a proliferação celular. Em relação às doenças associadas, o Grupo 1 apresentou genes exclusivamente associados ao câncer e o Grupo 2 foi associado as doenças gastrointestinais, genéticas e inflamatórias, assim como o Grupo 1 (Figura 21B). A análise de sobreposição de redes resultou na detecção de redes órfãs, ou seja, que não se associaram entre redes do mesmo grupo e/ou do outro grupo (Figura 22). As Redes 18 a 25 apresentaram funções como metabolismo de lipídeos, doença metabólica, função e desenvolvimento do sistema muscular e transporte molecular, enquanto que as Redes 1 e 2 do Grupo 2, foram associadas ao desenvolvimento celular e interação e sinalização célula-célula. Esses achados indicam que as CNVs identificadas nos Grupos 1 e 2 compartilham genes que desempenham funções biológicas semelhantes, contudo após a análise de sobreposição de redes, algumas funções foram exclusivas ao Grupo 1, assim como ao Grupo 2.

Em resumo, os genes preferencialmente associados a estas vias foram descritos em regiões de ganhos genômicos, indicando que as CNVs podem ter um efeito fenotípico na regulação da transcrição e proliferação celular em Leiomiomas Uterinos.

5.2 Perfil de expressão gênica global

Quando foi descrita por Schena et al. (1995), as análises de expressão gênica global ou microarray de expressão, prometia ser uma poderosa ferramenta para estudos moleculares. Esta metodologia permite quantificar milhares de transcritos em tecidos específicos e, quando comparados com tecidos controle, permite o entendimento das bases moleculares da doença (Catherino et al., 2003). Atualmente, os dados destes estudos têm sido utilizados para análises integradas com os dados genômicos, proteômicos e de metilação. Esta nova abordagem tem contribuído para o entendimento de doenças com base em alterações no número de cópias de DNA, independente do tecido controle (Akavia et al., 2010). Os experimentos de microarrays de expressão realizados no estudo atual tiveram como objetivo principal gerar dados transcriptômicos para a correlação com os dados genômicos. Contudo, quando comparados aos dados de miométrio adjacente, foi possível também obter os genes diferencialmente expressos que alteram o fenótipo normal das células do miométrio (Catherino et al. 2003). Na verdade, a maioria dos estudos utiliza esta abordagem em LU (para revisão Dimitrova et al., 2009). A primeira estratégia do presente estudo foi a obtenção dos genes diferencialmente expressos em relação ao miométrio normal. Inicialmente, foi selecionado um subgrupo de pacientes com informações clínicas completas, incluindo dados da fase do ciclo menstrual. Considerando que os LU são hormônio-dependentes, estes tumores poderiam ser diferencialmente responsivos a oscilação dos estrógenos e progesteronas durante o ciclo menstrual, influenciando os níveis de expressão de genes alvos (Critchley et al., 2006). Esta seleção resultou em 51 amostras obtidas de 34 pacientes, 20 pacientes encontravam-se na fase

proliferativa e 31 na fase secretora do ciclo menstrual. Após a obtenção dos dados, a análise de agrupamento hierárquico não supervisionado revelou que os miométrios adjacentes foram agrupados, porém, as amostras não se agruparam de acordo com a fase do ciclo, indicando que os tumores possuem perfil de expressão gênica semelhantes, independente da fase do ciclo ao momento da cirurgia. Na verdade, durante o período menstrual, o endométrio responde a diferentes picos hormonais, contudo, ainda não está claro se o miométrio responde a estes estímulos da mesma forma. Noe et al. (1999) demonstraram por IHQ que o miométrio sub-endometrial apresenta um padrão de expressão de ER (receptor de estrógeno) e PR (receptor de progesterona) cíclico semelhante ao endométrio, contudo, a porção mais interna do miométrio não apresentou este padrão. Mertens et al. (2001) utilizaram IHQ para avaliar a expressão proteica de AR, ER e PR em diferentes cortes do útero. As células epiteliais glandulares do endométrio apresentaram um padrão cíclico em relação a expressão dos receptores, porém, os autores não observaram este aspecto nas células do miométrio, que apresentaram o mesmo padrão de expressão dos receptores, independente da fase do ciclo. Tsuchida et al. (2007) utilizaram a técnica FES-PET que quantifica *in vivo* a concentração de moléculas alvo. Os autores avaliaram a concentração de ER no endométrio e no miométrio de 16 mulheres saudáveis. No endométrio, foi identificada uma maior quantidade de ER na fase proliferativa do que na secretora; em contraste, não houve diferença estatística entre a quantidade de ER e a fase do ciclo. Entre os estudos de expressão gênica em LU, cinco deles utilizaram amostras de pacientes em diferentes fases do ciclo, contudo, nenhum deles identificou um perfil de expressão gênica diferencial associada a esta categorização, confirmando nossos achados (Tsibris et al., 2002; Catherino et al., 2003; Hoffman et al., 2004; Quade et al., 2004; Lee et al., 2005).

Também foi avaliado se os tumores múltiplos apresentavam um perfil transcriptômico característico. As amostras apresentaram uma tendência a se agruparem entre os casos com mais de uma amostra avaliada de uma mesma paciente em relação aos

casos com LU único (Figura 23), semelhante ao detectado pela análise de CGH array. Três casos apresentaram um perfil similar de transcritos entre as diferentes amostras da mesma paciente (300T_A e B; 301T_A, B e C; 307T_A e B e 317T_A, B e C). Dimitrova et al. (2009) investigaram o perfil de expressão gênica em três amostras de LU de uma única paciente e validaram o achado em outro *dataset*. Os autores revelaram uma alta similaridade do perfil de expressão entre as diferentes amostras de uma mesma paciente, identificando genes com expressão aumentada associados com ativação da proliferação e celular e inibição da apoptose e os genes com expressão diminuída, associados com invasão e metástase, os quais geralmente estão com expressão aumentada em carcinomas. O presente estudo demonstrou uma grande similaridade do perfil de expressão gênica entre os diferentes LU de uma mesma paciente.

Três pacientes relataram história de LU em parentes de primeiro e segundo grau (casos 629T, 683T e 857T). No entanto, estas amostras não apresentaram tendência a agrupamento e expressão gênica diferencial em relação aos outros tumores. Vanharanta et al. (2006) compararam o perfil de expressão gênica de sete LU com mutação germinativa no gene *FH* com 15 LU com o alelo selvagem deste gene. O gene *FH* têm sido associado com predisposição aumentada ao desenvolvimento de LU em síndromes hereditárias de câncer. Entre os tumores com a mutação em *FH*, os genes com expressão aumentada estavam associados com metabolismo de carboidratos e glicólise e entre os genes com expressão diminuída, foram relatados genes associados com matriz extracelular, adesão celular, desenvolvimento muscular e contração celular. No presente estudo, os resultados de CGH array não revelaram um novo gene candidato ao desenvolvimento de LU em famílias, portanto, não foi possível associar o achado com os dados de expressão gênica. Estudos adicionais com vários membros das famílias (afetados e não afetados) são necessários para a investigação de outros possíveis genes candidatos associados com LU familiares.

Entre os 206 genes diferencialmente expressos em relação ao miométrio adjacente, a frequência de genes com expressão diminuída foi maior (66%) em relação aos genes com expressão aumentada (34%). Esta preponderância é uma característica observada entre os achados de microarrays de expressão em Leiomiomas Uterinos (para revisão Dimitrova et al., 2009).

Uma proporção significativa de genes que apresentam expressão aumentada em tumores malignos, encontraram-se com expressão diminuída entre os LU, outro achado também já previamente descrito (Ahn et al., 2003). Foram detectados 13 genes (~6%) diferencialmente expressos que foram previamente relatados em estudos de microarrays de expressão em Leiomiomas Uterinos. A comparação de listas de genes obtidas de diferentes estudos de microarrays é um grande desafio pois as amostras são provenientes de pacientes de diferentes etnias, são utilizadas plataformas distintas assim como são variáveis os métodos de seleção dos genes diferencialmente expressos (Quade et al., 2004). Entre os estudos de microarrays de expressão em LU, este é o único que utiliza a plataforma *Agilent*, os demais utilizaram as plataformas da *Affymetrix*.

Comparando as listas de genes obtidos dos estudos em LU utilizando plataformas *Affymetrix*, apenas 15% deles são concordantes, revelando a variabilidade experimental destes estudos (Dimitrova et al., 2009). Após uma busca detalhada no banco de dados do *NCBI* e entre as tabelas dos estudos de microarrays de expressão, as principais funções identificadas entre os genes deste estudo foram associadas com regulação do ciclo celular, apoptose, proliferação e crescimento celular.

Alterações entre os níveis de transcritos em moléculas que atuam em qualquer uma das etapas do ciclo celular poderiam levar a célula a se tornar quiescente ou desencadear uma proliferação anormal, características dos tumores. Portanto, genes que atuam nestas vias podem ser responsáveis pela transformação neoplásica das células do miométrio. Estas vias parecem realmente exercer um importante papel nos LU, pois foram também identificadas em LU de modelo animal. Crabtree et al. (2009) avaliaram o perfil de

expressão gênica em LU de ratos e humanos e revelaram que os genes alterados em ambas as espécies participam de vias associadas com regulação do ciclo celular, produção de matriz extracelular e metabolismo de aminoácidos.

A análise funcional de redes e vias canônicas aplicada aos genes deste estudo parece confirmar esta hipótese. Esta análise revelou 13 redes de interação gênica e metade delas apresentou associação com genes que atuam nas etapas do ciclo celular. Os achados das Redes 1 e 2 serão discutidos em detalhes pois eles resumem os principais achados de todas as outras redes.

Entre os genes da Rede 1, que mostrou interações entre genes do ciclo celular, câncer, doença do sistema reprodutivo (Figura 24A), os genes *JUN* (1p32.1) e *FOS* (14q24.3) apareceram como centrais de rede e com expressão diminuída. Como já apresentado, o estrógeno e a progesterona, assim como os fatores de crescimento, influenciam a proliferação das células do miométrio. Os genes *JUN* e *FOS* atuam como reguladores da transcrição de genes alvos, em resposta a estes estímulos (Gustavsson et al., 2000). A expressão reduzida destes genes, já foi descrita previamente em LU por microarrays de expressão e confirmada por outras metodologias (Gustavsson et al., 2000; Skubitz e Skubitz, 2003a; Hoffman et al., 2004). Estes resultados indicam que em LU, um tumor essencialmente benigno, os esteróides parecem não modular a expressão de *JUN* e *FOS* e, portanto, outros mecanismos estariam associados à expressão diminuída destes genes.

Um dos genes ligados a *JUN* é o *CDKN2A* (9p21.3) que também apresentou expressão diminuída. Este gene é um supressor tumoral que codifica a proteína p16 que se liga especificamente a CDK-4, inibindo a atividade catalítica do complexo CDK4-ciclina D. A proteína p16 atua como um regulador negativo do ciclo celular, levando a parada do ciclo em G1 (Bodner-Adler et al., 2005). Skubitz et al. (2003b) identificaram a redução de expressão deste gene em LU e a expressão reduzida da p16 também foi relatada em LU comparados aos LMS (O'Neill et al., 2007; Gannon et al., 2008). Markowski et al. (2010)

propôs um mecanismo para explicar a expressão diminuída de *CDKN2A* em LU. A *senescência induzida por oncogenes* (OIS) é um fenômeno frequente em lesões pré-malignas, que leva a parada do crescimento celular principalmente pela ativação de duas potentes vias, representadas por p16 e p19. A relevância de OIS para o desenvolvimento de LU ainda não está muito clara, porém a proteína HMGA2, codificada por um gene alvo associado a anormalidades cromossômicas recorrentes, parece estar associada com a repressão do *locus Ink4a/Arf* (*CDKN2A*). Portanto, o gene *HMGA2* poderia contribuir para o crescimento de LU pela repressão de *CDKN2A*. No estudo atual, o gene *HMGA2* não apareceu entre os diferencialmente expressos. Dessa forma, é possível que outros mecanismos contribuam para a diminuição da expressão do gene *CDKN2A*.

O gene *SDC1* (2p24.1) codifica uma molécula de adesão (Syndecan-1) que se liga a células da matriz extracelular. A expressão aumentada deste gene foi associada a progressão e metástase em câncer de colo de útero (Numa et al., 2002). No presente estudo, o gene *SDC1* foi descrito com expressão diminuída. Por meio de *FOS*, *SDC1* está ligado com os genes de colágeno, no qual *JUN* também está diretamente conectado.

Como já apresentado, os colágenos são os principais componentes associados com a característica fibróide dos Leiomiomas Uterinos. No presente estudo, o gene *COL3A1* (2q32.2) apresentou expressão aumentada e está diretamente ligado a *JUN*. Este gene codifica o colágeno III encontrado em tecidos conectivos, como o útero. Utilizando microarrays de expressão, Behera et al. (2007) identificaram expressão aumentada deste gene em Leiomiomas Uterinos. Outros dois genes codificadores de colágeno também foram identificados nesta rede e apresentaram expressão aumentada, o *COL4A1* (13q34) e o *COL17A1* (10q25.1). Contudo, estes genes ainda não haviam sido associados aos Leiomiomas Uterinos e são importantes candidatos para validação.

Outra observação interessante nesta rede foi a conexão de RAR, receptor do ácido retinóico, com os genes *JUN* e *FOS*. Embora no presente estudo o gene *RARA* não tenha sido identificado como diferencialmente expresso, há influencia de outros genes da rede com

expressão alterada. Vários membros da via RAR estão entre os mais consistentemente identificados com expressão alterada identificados por estudos de microarray em literatura (para revisão Arslan et al., 2005). Estes genes codificam enzimas envolvidas na síntese de RAR, como a álcool desidrogenase 1 (*ADH1*), aldeído desidrogenase 1 (*ALDH1*) e proteína celular de ligação a RAR 2 (*CRABP2*). Outros genes incluem o retinol desidrogenase (*RODH*) (Tsibris et al, 2002;. Skubitz e Skubitz, 2003a) e proteína celular de ligação a RAR 1 (*CRBP1*) (Vanharanta et al., 2005;. Zaitseva et al., 2006). A via do ácido retinóico controla uma série de processos biológicos, incluindo proliferação, diferenciação e apoptose e está implicada no desenvolvimento de vários tipos de tumores (Napoli, 1996). Dada a evidência da expressão aberrante de vários genes envolvidos nesta via em LU, quando comparado com o miométrio normal, é possível que alterações na síntese de RAR e de sinalização de ácido retinóico desempenham um papel importante no desenvolvimento destes tumores. No presente estudo foi identificada a expressão diminuída do gene *ALDH4A1* (1p36.13), da via do ácido retinóico, porém, não foi associada nesta rede.

A Rede 2 (Figura 24B) apresentou interações entre genes que exercem funções envolvidas com ciclo celular, doença hematológica e metabolismo de carboidratos. Um dos genes centrais desta via que se encontra com expressão diminuída, é o *PCNA* (20p12.3), que codifica uma proteína que atua como um “homotrímero”, auxiliando na síntese de DNA. Em resposta a danos no DNA, esta proteína é ubiquitinada e é associada a via de reparo a danos dependente de RAD6. A proteína PCNA geralmente encontra-se com atividade aumentada em LU (Gao Z et al., 2001; Zasławski et al., 2001). Tratamento com GnRH e CDB-2914 (modulador do receptor de progesterona) parecem reduzir a atividade desta proteína em células de Leiomiomas Uterinos (Wang et al., 2002; Chen et al., 2005; Xu et al., 2005), porém no presente estudo, este gene apresentou expressão diminuída e parece não ser um bom candidato a este tipo de terapia. Para nosso conhecimento, não há estudos que relataram a expressão diminuída de *PCNA* em Leiomiomas Uterinos. Quando

analisados os genes da rede que regulam o *PCNA*, verificou-se a ligação com RNA Pol II, que foi relatada anteriormente como um novo achado entre os resultados de CGH array. Genes que codificam as proteínas do complexo da RNA Pol II foram encontrados em regiões de ganhos genômicos. Portanto, um desequilíbrio do complexo da RNA Pol II poderia desencadear uma transcrição alterada de *PCNA*.

O gene *SFN* (1p36.11), que atua como regulador negativo do ciclo celular, também estava conectado com a RNA Pol II na Rede 2. A proteína codificada por este gene (14-3-3 ou estratifina) contribui para a parada do ciclo celular em G2. A sua expressão é induzida pela proteína p53 em resposta a danos do DNA. A perda de expressão deste gene foi observada em câncer de mama e de endométrio, associada principalmente com a hipermetilação de sua região promotora (Ito et al., 2005; Luo et al., 2010). Contudo, este gene ainda não havia sido associado aos Leiomiomas Uterinos. Assim, é possível que um dos mecanismos de inativação do gene *SFN* seja a deleção e o outro a hipermetilação.

Outro gene identificado como central da rede foi o *HIST2H3C* (1q21.2) que também está ligado ao *PCNA* e apresentou expressão diminuída. Este gene codifica uma das histonas do complexo do nucleossomo. Modificações neste complexo poderiam impedir o reposicionamento e ou montagem do nucleossomo, alterando processos de transcrição de RNA e reparo a danos no DNA (Manohar et al., 2009). Entre os genes ligados a *HIST2H3C*, o *PRLR* (5p13.2) também foi detectado com expressão diminuída. Este gene codifica o receptor de prolactina que pertence a família dos receptores de citocinas tipo 1. A expressão aumentada deste gene foi relacionada a pior prognóstico em câncer de mama e próstata (Jacobson et al., 2011). Para o nosso conhecimento, não há relatos do envolvimento deste gene em Leiomiomas Uterinos.

O gene *E2FA* (5p13.2) apresentou expressão aumentada e apareceu conectado ao *PCNA*. Este gene codifica um membro da família de fatores de transcrição que desempenha um importante papel na regulação do ciclo celular na transição G1/S. A expressão aumentada deste e outros genes que atuam na etapa G1/S foi associada a lesões

intraepiteliais de útero (Arvanitis e Spandidos, 2008). Juntas, as redes de interação evidenciaram genes que já haviam sido descritos em LU associados com genes ainda não descritos nestes tumores, porém, bem caracterizados em outros tipos de neoplasias hormônio-dependentes. As funções exercidas por estes genes, associadas principalmente com ciclo celular, revelam um dos possíveis mecanismos primordiais da transformação de células normais do miométrio em células tumorais.

Entre as principais vias canônicas geradas, na via de regulação do *checkpoint* G1/S do Ciclo Celular, os genes *CDC25A*, *CDKN2A*, *E2F5*, *NRG1*, *TFDP1* foram associados às moléculas do sistema IPA. Na representação gráfica desta via (Figura 28), é possível observar a interação das moléculas codificadas por estes genes. Em resposta a danos no DNA, o TGF-beta ativa a cascata de sinalização de moléculas envolvidas no *checkpoint* do ciclo celular e na transcrição de genes alvos. Nas células eucarióticas, a regulação do *checkpoint* garante a fidelidade da divisão celular. Este tipo de controle verifica se os processos em cada fase do ciclo celular foram corretamente preenchidos antes da progressão para a fase seguinte (IPA). A ação e função gene *CDKN2A* foi descrita acima.

O gene *CDC25A* (3p21.31), que apresentou expressão diminuída, codifica uma fosfatase (Cdc25A) crucial para a transição G1/S do ciclo celular. Em carcinomas, este gene tem sido alvo de agentes anti-tumorais (Lavecchia et al., 2010), contudo, em tecidos musculares, a expressão reduzida deste gene foi observada por mecanismos da regulação dos miRNAs-322/424 e miRNAs-503. A diminuição de expressão de *CDC25A* durante a diferenciação muscular seria ocasionada pela ação desses dois microRNAs, que levaria a repressão transcricional de *CDC25A*, induzindo a quiescência do ciclo celular (Sarkar et al., 2010). Estes miRNAs não foram encontrados nas regiões de ganhos e perdas significativas, contudo, este mecanismo de regulação não pode ser descartado.

O gene *NRG1* (8p12) codifica a neuregulina, que é um proteoglicano secretado na junção neuromuscular que inerva neurônios motores e sensoriais. A diminuição de expressão deste gene foi associada com rearranjos complexos em 8p, frequentemente

observados em câncer de mama, cólon e de pâncreas (Pole et al., 2006). Contudo, a análise de CGH array deste subgrupo de amostras não revelou deleções em 8p (dados não mostrados) apenas ganhos, portanto, outros mecanismos poderiam levar a diminuição da expressão deste gene, como por exemplo, hipermetilação, como já relatada em adenoma e carcinomas colorretal (Oster et al., 2011).

O gene *TFDP1* (13q34), também identificado com expressão diminuída, codifica a proteína DP-1, que forma dímeros com a proteína E2F, promovendo a transcrição de genes alvo. Este complexo controla a atividade transcricional de genes envolvidos na progressão G1/S do ciclo celular. Geralmente, em associação com E2F, este gene apresenta expressão aumentada em câncer de mama e de pulmão, induzindo a proliferação celular (Melchor et al., 2009; Castillo et al., 2010). Contudo, a falha do complexo E2F/TFDP1, poderia levar a parada do ciclo celular em G1, causada pela formação de heterodímeros defeituosos (Wu et al., 1996). Considerando a atividade aumentada de *E2F5* (que codifica uma das proteínas do complexo E2F), a expressão diminuída de *TFDP1* poderia contribuir para os mecanismos de proliferação celular anormal em um subgrupo de Leiomiomas Uterinos.

A outra via selecionada foi a de sinalização do IGF-1. O IGF-I exerce sua ação mitogênica por aumentar a síntese de DNA, acelerando a progressão do ciclo celular de G1/S e por inibir a apoptose. É um potente mitógeno em uma ampla variedade de linhagens de células neoplásicas, incluindo células estrógeno-dependentes como próstata, mama, ovário e útero (Yu e Rohan, 2000). Os genes diferencialmente expressos identificados nesta via foram *FOS*, *IGF1*, *JUN*, *SFN* e *YWHAE*. Esta via parece realmente exercer um efeito crucial no desenvolvimento dos LU, pois também foi identificada no estudo de CGH array (Figura 18). Quando a via da sinalização IGF-1 identificada por CGH array e pelos genes de expressão foram comparadas, foi verificada a sobreposição do gene *SFN* com atividade reduzida nas duas vias. Expressão aumentada do gene *IGF-1* (12q23.2) foi relatada em LU comparados ao miométrio adjacente, independentemente da fase do ciclo (Englund et al., 2000; Skubitz e Skubitz., 2003a). Burroughs et al. (2002) testaram a

hipótese sobre a perda da regulação da sinalização IGF-1 em LU e proliferação celular característica destes tumores. Os autores avaliaram LU de ratos em diferentes fases do ciclo e identificaram que independente do período do estro, a expressão do gene *IGF-1* sempre se manteve aumentada nos tumores em relação ao miométrio. Como já apresentado, estudos de microarrays de expressão demonstraram o aumento de expressão do gene *IGF-2* (Weston et al., 2003; Catherino et al., 2004a; Arslan et al., 2005) e do gene *IGF-1* em resposta aos estrógenos (Skubitz e Skubitz, 2003a; Swartz et al., 2005) em Leiomiomas Uterinos. O gene *YWHAE* (17p13.3), identificado com expressão diminuída, codifica a 14-3-3 *epsilon* que interage com CDC25 e atua na divisão celular. A expressão reduzida deste gene foi associada a meduloblastomas, indicando sua função como possível supressor tumoral (Cvekl et al., 2004).

Os resultados da análise de expressão gênica global demonstraram o envolvimento de genes que controlam os processos de recombinação cromossômica, reparo a danos no DNA e indução da apoptose associados com estímulos de proliferação e acúmulo da matriz extracelular.

5.3 Integração dos dados genômicos e transcriptômicos

Atualmente, o acúmulo do conhecimento sobre dados genômicos em tumores humanos tem permitido o desenvolvimento de uma nova área na pesquisa sobre câncer, a integração dos dados das alterações no número de cópias de DNA com os perfis de expressão gênica, metilação, miRNAs e proteínas (Jönsson et al., 2010; Zavadil et al., 2010; Alvarez et al., 2011; Su et al., 2011). Os resultados destes estudos aumentam a compreensão dos mecanismos patogênicos do câncer e permitem o desenvolvimento de novas terapias. Neste estudo, um dos objetivos propostos foi a realização da integração de dados genômicos e transcriptômicos, na tentativa de selecionar genes candidatos a

moduladores envolvidos com vias moleculares associadas à gênese dos Leiomiomas Uterinos.

Foram integrados os dados genômicos e transcriptômicos obtidos das análises de CGH array e microarray de expressão respectivamente, de 51 amostras de LU obtidas de 34 pacientes. Para a realização desta análise, os dados de CGH array foram analisados por outro método de bioinformática, o qual gera resultados compatíveis com o algoritmo utilizado para a integração dos dados, o CONEXIC. O CONEXIC não permite a utilização de dados de CGH array gerados por programas de análise “fechados”, como é o caso do programa *Nexus* (www.c2b2.columbia.edu/danapeerlab/html/software.html). Contudo, o algoritmo JISTIC (utilizado pelo programa *DNA Copy* para a análise de CGH array) é uma variável de outro algoritmo, o *Circular Binary Segmentation* (Olshen et al., 2004), que gerou resultados muito próximos aos obtidos pela análise do programa *Nexus*. Aproximadamente 80% das regiões identificadas pelo *Nexus* também foram identificadas pelo *DNA Copy* (dados não mostrados). A semelhança obtida entre os resultados provavelmente deve-se ao fato de o algoritmo utilizado pelo *Nexus*, o *Rank Segmentation*, também ser uma variável do algoritmo *Circular Binary Segmentation*, utilizado pelo pacote *DNA Copy*. Isto revela a acurácia de duas análises independentes que identificaram uma alta similaridade de CNVs, indicando que as alterações descritas podem ter um importante papel na formação dos Leiomiomas Uterinos.

A análise de CGH array identificou 1192 genes mapeados em regiões com *G-score* significativo (picos de alterações) com ganhos e perdas genômicas, enquanto que a análise de expressão gênica identificou 3325 genes significantes. Após a obtenção destes dados, foi aplicada a análise de integração dos dados. Até o momento não existe um consenso sobre qual o melhor algoritmo utilizar para integração destes tipos de dados, sendo assim, foi utilizado neste estudo uma ferramenta recentemente publicada por Akavia et al. (2010). Esta ferramenta tem como objetivo principal identificar genes moduladores relacionados com o fenótipo tumoral, o CONEXIC. Uma característica chave do algoritmo

utilizado neste estudo, é que ele vai além da identificação dos genes moduladores. Ao associar genes candidatos com módulos de genes e anotá-los usando dados da literatura, o CONEXIC fornece pistas sobre os papéis fisiológicos destes genes e dos genes ligados a eles. O CONEXIC difere de outros métodos em várias maneiras. Em primeiro lugar, ele usa principalmente as informações de expressão de um gene candidato (ao invés do padrão do número de cópias) como uma representação do *status* do gene. Em segundo lugar, ele associa um gene candidato a um módulo de genes cuja expressão corresponde a um gene modulador. Em terceiro lugar, associa o número de cópias de DNA com os dados de expressão gênica, que fornece maior sensibilidade para identificar regiões significativamente aberrantes que não seriam identificadas apenas em análises de DNA. Os métodos clássicos de análise integrada, baseados em dados de alterações no número de cópias estão limitados a detectar grandes regiões genômicas que contêm múltiplos genes, onde o gene modulador não é facilmente identificado entre eles (Akavia et al., 2010).

Esta análise permitiu a identificação de 75 genes moduladores que explicam o comportamento de expressão dos 3325 genes identificados na análise de expressão gênica. Estes genes estavam distribuídos em 25 regiões genômicas, nos quais 32% (8/25 regiões) já haviam sido descritas previamente em LU por estudos de HR-CGH, que são 1p36.13, 4p14, 7q22.1, 11p15.5, 14q13.2, 16p11.2, 19q13.32 e 19q13.33 (Packenham et al., 1997; Levy et al., 2000; Christacos et al., 2006; Canevari e Rogatto, 2007; Hodge et al., 2009). As novas regiões identificadas neste estudo contendo os genes moduladores foram 1q41, 2q32.1, 2q32.2, 2q35, 5q31.2, 5q35.3, 8p12, 8q24.3, 10p15.3, 10q21.3, 11q13.2, 12p11.21, 12p13.31, 16q24.3, 17q21.31, 20p11.2 e 20p13.

Outra peculiaridade do CONEXIC é a identificação de genes moduladores independente de seu padrão no número de cópias de DNA (Akavia et al., 2010). Apesar de a análise integrada associar o padrão de expressão gênica com alterações no número de cópias, para selecionar um gene modulador, o CONEXIC não leva em conta se houve correlação positiva (ganho genômico/aumento expressão) ou negativa (perda

genômica/diminuição de expressão) entre os dados, ou seja, o algoritmo identifica um gene modulador baseado nos dados de expressão (redes em módulos e GSEA) e depois associa com as regiões de instabilidade genômica (valores de *G*-score significativo). Neste estudo, entre os 75 genes moduladores, 46 apresentaram correlação inversa entre o padrão genômico e o perfil de expressão gênica. Este achado confirma que existem outros mecanismos de regulação gênica. Alterações no número de cópias de DNA são apenas um dos vários eventos que levam a alterações nos níveis de expressão do gene (Akavia et al., 2010). Mecanismos epigênicos, por exemplo, levam a alterações transcricionais (Watanabe e Maekawa, 2010). Entre estes mecanismos, a regulação da expressão de genes por miRNAs tem sido um dos principais focos em estudos de câncer atualmente, inclusive em Leiomiomas Uterinos (Ryan et al., 2010; Zavadil et al., 2010). Diante disso, foi realizada uma busca nos bancos de dados de predição de miRNAs que poderiam ser os possíveis reguladores de expressão destes genes que apresentaram correlação inversa (Tabela 11). Os genes *ATP5J2*, *AZGP1*, *CD19*, *CORO1A*, *CYC1*, *ELMO3*, *GDPD3*, *NFKBIL2*, *PKP3*, *PYCRL*, *RASSF7*, *RECQL4*, *SNRPD2*, *STAG3* e *TALDO1* não apresentaram predição por miRNAs, indicando que podem ser regulados por outros mecanismos. Estes resultados permitiram concluir que 46/75 (61,33%) dos genes moduladores apresentaram correlação inversa entre o padrão de ganhos e perdas genômicas e o perfil de expressão, indicando que alterações no número de cópias de DNA podem não ser o mecanismo primário de regulação destes genes. De fato, os genes *FANCA* e *TBC1D20* (identificados entre os 75 moduladores), tiveram expressão diminuída e de acordo com a lista de predição gênica, os miRNAs *hsa-miR-26b* e *hsa-miR-150* regulam estes genes, respectivamente. Interessantemente, os dois miRNAs estão mapeados em regiões envolvidas em ganhos genômicos neste estudo (2q35 e 19q13.33, respectivamente), o que explicaria a correlação inversa observada. O gene *FANCA* foi inicialmente associado com anemia de Fanconi, uma doença genética recessiva caracterizada por aumento de quebras cromossômicas, sensibilidade aumentada aos agentes que causam danos no DNA e defeitos no reparo a

danos do DNA (para revisão Collins et al., 2009). Atualmente, deleções heterozigotas da região promotora deste gene foram associadas com câncer de mama familiar (Solyom et al., 2011). Este gene foi identificado com expressão reduzida por Crabtree et al. (2009) em LU, contudo não havia sido associado com regulação por miRNAs. Para nosso conhecimento, não encontramos estudos associando o gene *TBC1D20* com doenças humanas. Portanto, é necessário validar os dados de expressão de miRNA e dos genes que são por eles regulados em uma série independente de casos, os quais explicariam os mecanismos de regulação da expressão gênica em Leiomiomas Uterinos.

Genes codificadores de miRNAs em regiões de ganhos e perdas genômicas podem levar a perda ou aumento de expressão de seus genes alvos, respectivamente. Esta hipótese é baseada no fato de que a maioria dos miRNAs relacionados ao câncer (~52,5%), estão mapeados principalmente em regiões de instabilidade cromossômica, como os sítios frágeis (Calin et al., 2004; Huppi et al., 2008; Caldas e Brenton, 2005; Laganà et al., 2010) e as alterações genômicas contribuem para essa instabilidade. Sendo assim, foi realizada uma busca dos miRNAs mapeados em regiões de ganhos e perdas genômicas descritas neste estudo. Essa busca totalizou 31 miRNAs, entre os quais, 22 foram elegíveis para regulação gênica. A lista de genes preditos (6635) foi comparada aos 3325 genes identificados por microarray de expressão, totalizando 780 genes sobrepostos, indicando que os miRNAs mapeados em regiões de instabilidade podem regular pelo menos 24% dos transcritos identificados neste estudo.

Para o nosso conhecimento existe apenas um estudo associando alterações no número de cópias em genes de miRNAs e o perfil de expressão de seus genes preditos em Leiomiomas Uterinos. Zavadil et al. (2010) utilizaram uma estratégia semelhante a utilizada no presente estudo. Os autores obtiveram o perfil de expressão de miRNAs e expressão gênica em 55 amostras de LU. Com o objetivo de avaliar se os miRNAs que estavam com expressão alterada estavam mapeados em regiões de instabilidade genômica, os autores realizaram CGH array em seis amostras de LU, nos quais

identificaram perda genômica de *hsa-miR-200a* e *hsa-miR-200b* (1p36.33-p34.3), *hsa-miR-206* (6p12.3-p11.1), *hsa-miR-15a*, *hsa-miR-16-1*, *hsa-miR-17*, *hsa-miR-19a*, *hsa-miR-20a*, *hsa-miR-92a-1* (13q12.12-q33.2). Estes miRNAs envolvidos em perdas genômicas também apresentaram diminuição de expressão. Os autores confirmaram através da comparação entre a lista de genes preditos com os dados de expressão e também por análises em modelos funcionais, que os genes alvos dos *hsa-miR-200a/b* (*TUBB*, *CYP1B1* e *CTBP2*) apresentavam expressão aumentada e foram associados com proliferação celular em Leiomiomas Uterinos. Os autores concluíram que a perda da função de miRNAs da família 200, identificada por perfil de expressão e CGH array, foi associada com a atividade aumentada de seus genes preditos, podendo contribuir para o crescimento tumoral. Em nosso estudo, não observamos as regiões genômicas identificada por estes autores e conseqüentemente, os miRNAs mapeados nelas também não foram identificados. Entretanto, nesta análise integrada, identificamos o gene *TUBB3* (16q24.3) com correlação inversa (ganho/diminuição expressão). Este gene também é regulado pelo *hsa-miR-200b* (Tabela 11) e pertence a mesma família do *TUBB*, identificado pelos autores. Foram descritas outras regiões portadoras de miRNAs que podem regular importantes genes associados aos Leiomiomas Uterinos. O *hsa-miR-26b* (2q35) envolvido em ganhos em seis casos regula genes que apresentaram expressão diminuída, como o *PCNA*. O gene *PCNA* foi apresentado acima como um dos importantes genes com expressão diminuída identificados na comparação com os miométrios adjacentes. Em literatura, há relatos da expressão aumentada do *PCNA* em LU (Gao et al., 2001; Zaslowski et al., 2001). No presente estudo, a atividade aumentada do *hsa-miR-26b* explicaria a expressão reduzida deste gene nestes casos.

O *hsa-mir-1236* (6p21.32) envolvido em ganhos em quatro casos, foi associado com a diminuição de expressão do gene *CDKN2A*. Este gene, também foi identificado no presente estudo com expressão reduzida associado ao miométrio adjacente. Skubitz et al.

(2003) identificaram diminuição da expressão do transcrito e sua proteína codificada em LU comparados aos LMS (O'Neill et al., 2007; Gannon et al., 2008).

Os miRNAs *hsa-mir-25*, *hsa-mi-106b*, *hsa-miR-548o*, mapeados em 7q22.1, parecem regular genes importantes já associados com Leiomiomas Uterinos. O *hsa-mir-25* envolvido em perda em três casos regula o gene *COL1A2*, identificado com expressão aumentada neste estudo. Aumento da atividade do gene *COL1A2*, foi associado com o aumento da deposição de colágeno, uma característica dos Leiomiomas Uterinos (Mesquita et al., 2010). O *hsa-miR-106b* envolvido em perdas em três casos regula genes que apresentaram expressão aumentada, como *ESR1* e *IGF-1*, reconhecidamente associados com regulação da transcrição e proliferação celular em Leiomiomas Uterinos (Swartz, 2005; Strissel et al., 2007). O *hsa-miR-548o*, envolvido em perdas em nove casos, foi associado com o aumento de expressão de *FGFR2*. Este gene tem sido um alvo terapêutico para tumores de mama e gástrico (Katoh, 2010; Asaoka et al., 2011). Em LU, a expressão protéica de *FGFR2* não diferiu estatisticamente entre tumores na fase proliferativa e secretora do ciclo menstrual (Wu et al., 2001).

Em geral, a descrição de miRNAs mapeados em regiões de ganhos e perdas genômicas não foi diretamente associada a uma correlação inversa de 46 dos 75 genes moduladores, contudo, este mecanismo poderia explicar a grande maioria das correlações inversas observadas neste estudo quando foram considerados todos os genes identificados na análise de expressão gênica global. Aproximadamente 60% dos 75 genes moduladores (Figura 35) estão mapeados em regiões que geralmente não são alvos frequentes de quebras cromossômicas, indicando que outros mecanismos poderiam estar associados com a correlação inversa observada entre os genes moduladores. Estudos adicionais são necessários para confirmar este achado.

O teste de agrupamento hierárquico não supervisionado dos genes moduladores indicou que entre os casos com diagnóstico de tumores múltiplos, apenas as amostras 300T_A e B e 301T_B e C apresentaram maior similaridade de ganhos e perdas genômicas

e também de expressão gênica. Como já discutido acima, Dimitrova et al. (2009) revelaram uma alta similaridade do perfil de expressão gênica entre diferentes amostras de uma mesma paciente. Este achado indica que um subgrupo de tumores é semelhante quanto ao padrão de ganhos e perdas genômicos assim como quanto ao perfil de expressão dos 75 genes moduladores.

A disponibilidade de dados públicos resultantes dos estudos de expressão gênica representam uma oportunidade sem precedentes para extrair informações genômicas relevantes e validação de hipóteses biológicas (Bisognin et al., 2009). Além disso, com estas informações é possível realizar testes de meta-análise para validação de um determinado conjunto de dados. No estudo atual, o perfil de expressão dos 75 genes moduladores foi comparado aos resultados de cinco estudos de expressão gênica em LU (Quade et al., 2004; Hoffman et al., 2004; Vanharanta et al., 2005; Crabtree et al., 2009; Hodge et al., 2009) que disponibilizaram estas informações no GEO (*Genomic Expression Omnibus*). Todos os 75 genes moduladores foram identificados em pelo menos três dos cinco estudos prévios, no qual 52 deles (69%) apresentaram o mesmo padrão de expressão em pelo menos um estudo (Tabela 13). Estes resultados indicam que os genes moduladores identificados são altamente confiáveis, quanto a sua associação aos Leiomiomas Uterinos e validam os dados obtidos no presente estudo.

A análise de enriquecimento de genes (*Gene Set Enrichimento Analysis -GSEA*) é um método computacional que compara um conjunto de dados de expressão gênica com grupos de genes relacionados a tumores (módulos), que compartilham as mesmas funções biológicas, regulação e localização cromossômica (Subramanian et al., 2005). Esta análise gerou o enriquecimento de 32 dos 75 genes moduladores dentro de 26 módulos gênicos. Os genes *CORO1A*, *FGFR1*, *DDX21* e *DBN1* foram os mais frequentemente identificados entre os módulos e apenas *FGFR1* e *DBN1* apresentaram correlação positiva (ganho/aumento expressão) (Tabela 14). Outros genes que foram enriquecidos em um menor número de módulos, porém não menos significantes, também apresentaram

correlação positiva como *NUPR1*, *COL3A1*, *AOC3*, *IFITM1*, *CALCRL*, *NOVA2*, *MVP*, *MFAP5*, *DIP2C* e correlação negativa *CENPF*, *RHOH* e *BICD1*. Entre estes genes, apenas *AOC3*, *CALCRL*, *FGFR1* e *COL3A1* foram previamente associados com Leiomiomas Uterinos (Anania et al., 1997; Nikitenko et al., 2006; Behera et al., 2007).

O ganho em *FGFR1* (8p12) foi um dos principais achados de CGH array para as 80 amostras de LU. A amplificação deste gene tem sido um potencial alvo terapêutico em cânceres de mama (Shiang et al., 2010). Como já apresentado, o sistema receptor-ligante bFGF na patogênese da hemorragia associada aos LU, podem ter implicações na fertilidade e contracepção (Anania et al., 1997). A descrição da correlação positiva deste gene na análise integrada revelou que tumores com amplificações do *FGFR1* podem ser alvos de estudos com drogas inibidoras de crescimento que atuam nas vias de proliferação *downstream* a sua ação de receptor de fator de crescimento associadas a um quadro clínico observado entre pacientes com Leiomiomas Uterinos.

O gene *DBN1* (5q35.3) codifica a proteína drebrina que atua no citoesqueleto, interagindo nos *gap junctions* em zonas de contatos célula-célula (Jin et al., 2002). A expressão aumentada deste gene, juntamente com *CACNB3*, *FLAD1* e *CCND2*, foi associada a predição de recorrência pós-cirúrgica em carcinoma de células não pequenas de pulmão (Mitra et al., 2011). Assim, ganhos e consequente aumento do transcrito do *DBN1* podem ser importantes na manutenção da estrutura dos LU como unidades bem delimitadas.

O gene *NUPR1* (16p11.2) que codifica a proteína p8, apresenta expressão aumentada em resposta ao estresse celular (Iovanna, 2002). A atividade oncogênica desta proteína foi associada com tumores de mama, tireóide, cérebro e pâncreas (Chowdhury et al., 2009). Os resultados de CGH array e microarrays de expressão deste estudo revelaram que vias do ciclo celular foram uma das mais frequentemente associadas com LU, portanto, a ação de *NUPR1* pode ter um importante papel na resposta a danos no DNA nestes tumores.

O gene *COL3A1* (2q32.2) codifica a cadeia pro-alfa I do colágeno tipo III, encontrado na pele, pulmão, útero, fígado, intestino e sistema vascular (Kuivaniemi et al., 1991). Neste estudo, este gene foi relatado com expressão aumentada em relação aos miométrios adjacentes. Como já apresentado, Behera et al. (2007) associaram o aumento de expressão deste gene com o aumento da deposição de colágeno em LU, mecanismo que contribui para o fenótipo fibróide observado nestes tumores.

O gene *AOC3* (17q21.31) codifica a proteína VAP-1 que participa principalmente de processos inflamatórios, porém, sua ação também foi observada em células do músculo liso. Jaakkola et al. (1999) avaliaram a expressão de VAP-1 em diferentes tecidos musculares normais e tumorais, entre eles, Leiomiomas Uterinos. Marcação positiva de VAP-1 foi observada em LU, enquanto que metade dos LMS avaliados não apresentaram marcação desta proteína. Os autores concluíram que a atividade de VAP-1 é perdida durante a transformação maligna e mantida em condições benignas. Neste estudo, ganho genômico do gene *AOC3* e aumento do transcrito poderiam estar associados a atividade de VAP-1 em Leiomiomas Uterinos.

O gene *IFITM1* foi apresentado acima como um dos possíveis genes candidatos associados com ganhos em 11p15.5 no estudo independente de CGH array. Como já discutido, este gene está associado com escape de células tumorais do sistema imune (Han et al., 2011). Provavelmente, este gene desempenhe um importante papel associado com sobrevivência das células do miométrio que exibem capacidade proliferativa diferenciada.

O gene *CALCRL* (2q32.1) codifica o receptor *like*-calcitonina (CALCRL) que forma complexos com a proteína-G. Receptores acoplados a proteína-G desempenham um importante papel como alvo de drogas (Young, 2005). Além disso, CALCRL medeia os efeitos da adrenomedulina (AM), que é um fator angiogênico induzido por hipóxia, associada com crescimento tumoral. Mesmo sendo tumores benignos, acredita-se que os LU possuem uma aberrante formação de novos vasos (angiogênese), essencial para o crescimento e proliferação destes tumores, pois apresentam atividade elevada de vários

fatores angiogênicos, como VEGF, bFGF, PDGF e AM (Harrison-Woolrych et al., 1995; Hague et al., 2000; Flake et al., 2003). Nikitenko et al. (2006) avaliaram o CALCRL como potencial alvo de terapia anti-angiogênica em Leiomiomas Uterinos. Os autores identificaram atividade aumentada de CALCRL no endotélio microvascular destes tumores, entre outros achados e concluíram que CALCRL poderia ser um potencial alvo terapêutico de inibidores de angiogênese em Leiomiomas Uterinos. Associados aos achados de nosso estudo, ganhos genômicos e aumento do transcrito de *CALCRL*, poderiam estar associados com o aumento da atividade de sua proteína em Leiomiomas Uterinos.

O gene *NOVA2* (19q13.32) codifica a proteína Nova-2 que se liga ao RNA produzido em neurônios e atua nas vias de drogas anti-epiléticas (Ueki et al., 1997; Heinzen et al., 2007). Para nosso conhecimento, este gene não possui associação direta com câncer, embora, possa estar associado com a Síndrome POMA (do inglês *paraneoplastic opsoclonus myoclonus ataxia*). A síndrome POMA, no qual os pacientes desenvolvem câncer de mama, tumores das trompas de Falópio, redução dos movimentos e demência, ocorre provavelmente devido às respostas auto-imunes desenvolvidas contra antígenos neuronais específicos pelas células tumorais (Luque et al., 1991; Pranzatelli, 1992; Buckanovich et al., 1993; Hormigo et al., 1994). As proteínas reconhecidas por seus anticorpos auto-imunes, Nova-1 e Nova-2, são proteínas que controlam os mecanismos de *splicing* de vários genes que atuam na função do circuito neuronal central (Jensen et al., 2000; Ule et al., 2005). Huang et al. (2005) estudando plasticidade sináptica em modelos *in vivo*, associaram a atividade aumentada de Nova-2 com a Síndrome POMA, tornando-se um alvo para esta doença auto-imune. Interessantemente, os tumores associados com esta síndrome são estrógeno-dependentes, portanto, o gene *NOVA2* poderia estar indiretamente associado com Leiomiomas Uterinos.

O gene *MVP* (16p11.2) codifica uma proteína (MVP) envolvida no transporte citoplasma-núcleo. A expressão aumentada desta proteína tem sido associada com

resistência à drogas e com a supressão do mecanismo de NHEJ (do inglês *non-homologous end-joining*), levando a instabilidade genômica em células tumorais (Lloret et al., 2009). Tumores cervicais que apresentam a expressão aumentada de MVP e aumento de angiogênese foram associados com resistência a maioria das drogas anti-câncer (Lara et al., 2009). Portanto, expressão aumentada deste gene em LU poderia não ser é um bom indicativo para tratamentos com anti-neoplásicos.

O gene *MFAP5* (12p13.31) codifica a glicoproteína MAGP2 que está associada com interações proteína-proteína na matriz extra celular e desempenha um papel na montagem de fibras elásticas (Penner et al., 2002). Em carcinomas de ovário, expressão aumentada de *MFAP5* foi associada com motilidade das células epiteliais via receptor de integrina alpha(v)beta(3) e correlacionada com densidade de microvasos, sugerindo um papel angiogênico nestes tumores (Mok et al., 2009; Spivey e Banyard, 2010). Aumento de expressão de *MFAP5* devido a ganhos genômicos, pode estar associado com a formação da vasculatura em LU, tornando este gene candidato a estudos de inibidores de angiogênese, associados com o gene *CALCRL*, também identificado neste estudo.

O gene *DIP2C* (10p15.3) codifica um fator de transcrição que atua no sistema nervoso central. Para nosso conhecimento, não foram encontrados estudos associando este gene com doenças humanas.

Entre os genes que apresentaram correlação negativa, encontra-se o gene *CENPF* (1q41) cuja proteína está associada ao complexo centrômero-cinetocóro (Testa et al., 1994). A expressão aumentada deste gene foi associada com invasão e metástase em carcinoma ductal pancreático (Grützmann et al., 2004). Invasão e metástase não são eventos comuns em LU, portanto, a atividade reduzida deste gene pode estar associada com a característica benigna destes tumores.

O gene *RHOH* (4p14) codifica uma proteína que é um dos membros da superfamília de GTPases que atuam na reorganização do citoesqueleto, adesão, motilidade e proliferação celular (Aspenström et al., 2007). A expressão normal deste gene é

característica de células hematopoiéticas, enquanto que expressão reduzida deste gene foi associada com leucemia HCL (do inglês *hairy cell leukemia*) e com leucemia mielóide aguda (Galiègue-Zouitina et al., 2008). O papel exato deste gene em tumores sólidos ainda não está claro, contudo, nosso estudo indica que a diminuição da atividade deste gene associada com perdas genômicas, pode estar associada com alterações no ciclo celular nas células dos Leiomiomas Uterinos.

O gene *BICD1* (12p11.21) codifica uma proteína envolvida com transporte vacuolar do retículo endoplasmático para o complexo de Golgi. A atividade reduzida deste gene foi associada com a variação do tamanho dos telômeros em humanos (Mangino et al., 2008). A disfunção dos telômeros está associada a um risco aumentado de desenvolvimento de câncer e outras doenças relacionadas a idade (Rudolph et al., 1999). Contudo, este gene ainda não havia sido associado com tumores humanos.

Neste estudo, também foi aplicada a análise funcional nos 75 genes moduladores. Como o CONEXIC e o GSEA consideram os valores de expressão gênica dos genes moduladores para a realização dos testes, estes valores também foram considerados na análise funcional. Em relação aos 75 genes, foram geradas cinco redes de interação gênica, nos quais os genes moduladores foram associados principalmente ao ciclo celular. De acordo com o sistema IPA, os genes com correlação positiva *AOC3*, *NOVA2* e *DIP2C* não foram elegíveis para a construção das redes gênicas, provavelmente, o sistema não identificou interações destes genes com os genes do banco de dados. Os outros genes com correlação positiva ou negativa apresentaram as seguintes associações: *FGFR1*: ciclo celular, proliferação e crescimento celular, desenvolvimento e função do sistema muscular e esquelético; *NUPR1*, *MVP*, *MFAP5* e *CENPF*: expressão gênica, desenvolvimento do organismo, desenvolvimento e função do sistema nervoso; *DBN1* e *COL3A1*: montagem e organização celular, duplicação do DNA, recombinação e reparo, metabolismo de carboidratos; *IFITM1*: ciclo celular, doenças e condições dermatológicas, doença imunológica; *CALCRL*, *RHOH* e *BICD1*: sinalização celular, metabolismo de ácidos nucleicos,

bioquímica de pequenas moléculas. Todas estas funções descritas na análise integrada também foram previamente associadas com os grupos de genes obtidos nas análises independentes de CGH array e expressão gênica, confirmando que estes genes desempenham funções biológicas cruciais para o desenvolvimento dos Leiomiomas Uterinos.

Outro achado interessante foi a associação dos genes moduladores com doenças como câncer, doença gastrointestinal, doença do sistema reprodutivo, doença genética, doença muscular e esquelética, entre outras (Figura 37A). Este achado também foi observado nas análises independentes de CGH array e expressão gênica. Dimitrova et al. (2009) descreveram a semelhança de perfis de expressão de genes identificados em sub-grupos de LU com outras doenças como câncer de mama, leucemia e câncer de pulmão. Outros estudos revelaram que doenças dermatológicas, como quelóides, apresentam perfis de expressão muito semelhantes aos LU, associados principalmente com o aumento da deposição de matriz extracelular (Leppert et al., 2006; Luo et al., 2007). No presente estudo foi possível verificar que novas doenças mostram semelhanças moleculares aos LU, como doenças hepáticas, cardiovasculares e hematológicas.

As vias canônicas geradas revelaram associação dos genes moduladores com as vias de sinalização já associadas com câncer, como também, refletiram alguns dos achados associados com as doenças. A via com o valor mais alto (2.75) (Figura 38) associou os genes *BRCA1*, *CDC25C*, *FANCA* e *HRAS* com a via de câncer de mama hereditário. Estes genes apresentaram correlação inversa (ganho genômico/perda expressão), porém, podem desempenhar importantes funções associadas com ciclo celular em Leiomiomas Uterinos e merecem ser melhor caracterizados nestes tumores. A via de sinalização do câncer de mama hereditário é desencadeada quando ocorrem danos ao DNA, onde *BRCA1* é fosforilado por sensores de resposta a danos (*ATM* e *ATR*). A proteína *BRCA1* também pode ser fosforilada por *CHK2*. A *BRCA1* fosforilada controla moléculas *downstream* do controle G2/M do ciclo celular (*CDC25C*, *FANCA* e *HRAS*). *BRCA1* também participa dos

processos de recombinação homóloga durante a meiose e no reparo de quebras em dupla-fita de DNA (*IPA Canonical Pathway Report*). Portanto, a perda da atividade destas moléculas permitiria que danos ao DNA não fossem eliminados, contribuindo para a instabilidade genômica.

A maioria dos estudos associando câncer de mama e LU relatam uma alta frequência de pacientes com carcinoma ductal invasivo ou lobular de mama com sítio metastático em Leiomiomas Uterinos (Minelli et al., 1998; Jashnani et al., 2007; Kondo et al., 2009; Ustaalioglu et al., 2009). Contudo, para nosso conhecimento, existe um estudo associando o gene *BRCA1* com câncer de mama e Leiomiomas Uterinos. Doridot et al. (2002) relataram um caso de uma paciente pré-menopáusia em tratamento com Tamoxifeno para câncer de mama, a qual apresentou um LU com alto índice mitótico. Além disso, esta paciente apresentou mutações em *BRCA1* e *BRCA2* que também confere um risco aumentado para o desenvolvimento de câncer de ovário. Os autores concluíram que o LU foi diagnosticado anteriormente ao tratamento com Tamoxifeno, porém, o uso dessa droga associada com mutações em *BRCA1* e *BRCA2* favoreceu o crescimento desencadeado do Leiomioma Uterino nesta paciente. Lehtonen et al. (2006) avaliaram 868 indivíduos de famílias com mutações no gene *FH* que desenvolveram a síndrome HLRCC (leiomiomatose hereditária e carcinoma renal) e o risco de desenvolvimento de outros tipos de câncer. Os autores identificaram uma alta frequência de câncer de mama nestes indivíduos. Investigação de alterações no gene *BRCA1* neste sub-grupo de pacientes que desenvolveram câncer mama e Leiomiomas Uterinos são necessárias para associação dos achados.

O gene *CDC25C* codifica a Cdc25 que atua na regulação da divisão celular. Atividade aumentada de Cdc25 geralmente está associada com invasão e metástase em vários tumores, incluindo carcinoma de endométrio, carcinoma nasofaríngeo e carcinoma de células não pequenas de pulmão (Brazelle et al., 2010; Cho et al., 2011; Su et al., 2011). Contudo, Zheng et al. (2011) relataram uma diminuição da atividade de Cdc25 em

carcinoma colorretal, devido á expressão aumentada de parafibromina, uma proteína associada com tumores de mama e carcinomas gástricos. Os autores associaram a redução de Cdc25 com prognóstico favorável em pacientes com carcinoma colorretal. No estudo atual, a redução da expressão do gene *CDC25C* pode estar associada com a perda da capacidade invasiva e metastática em Leiomiomas Uterinos.

O gene *HRAS* também foi associado com esta via. Interessantemente, o gene *HRAS* havia sido identificado como um gene candidato nos resultados da análise independente de CGH array, associado com ganhos em 11p15.5. Contudo, os estudos de microarrays de expressão revelaram diminuição de expressão deste gene em nosso estudo. Como discutido acima, 11p15.5 é alvo frequente de instabilidade genômica, pois está localizada na região telomérica dos cromossomos. Além disso, expressão diminuída da p21 (proteína codificada pelo *HRAS*) parece ser um evento característico de tumores benignos e já foi associada com leiomiomas gastrointestinais extra-uterinos (Cai et al., 1995). Outras vias associadas com ciclo celular também foram descritas, como as proteínas CHK associadas ao controle do ciclo celular, ciclo celular G2/M, reparo em quebra de dupla fita por recombinação homóloga e controle do ciclo celular da duplicação cromossômica, confirmando a importância dos genes modulares nas vias do ciclo celular.

Outra via canônica associada a doenças foi a fibrose hepática. Os genes associados a esta via foram *COL3A1*, *FGFR1* e *IGFBP5* e todos tiveram correlação positiva (Figura 39). Assim como o *FGFR1*, o *IGFBP5* foi identificado entre os achados independentes de CGH array. Como discutido, expressão aumentada do *IGFBP5* tem sido associada com neoplasia intra-epitelial cervical e câncer de mama, podendo desempenhar um papel no desenvolvimento de tumores estrógeno-dependentes (Hou et al, 2009; Taylor et al, 2010). A fibrose hepática é uma doença crônica do fígado associada com o acúmulo de matriz extracelular, desencadeada por hepatotoxinas como álcool, ácido biliar, glicose e ácidos graxos livres. Danos contínuos ao fígado levam ao acúmulo de células estreladas hepáticas ativas e miofibroblastos, que por sua vez sintetizam grandes quantidades de matriz

extracelular, principalmente colágeno, levando a fibrose tecidual (*IPA Canonical Pathway Report*). Estas características são muito similares às condições patológicas dos Leiomiomas Uterinos. Recentemente, foi descrito um inibidor (NP603) da atividade tirosina quinase de FGFR1, que inibe a proliferação de miofibroblastos associados com fibrose hepática em ratos (Lin et al., 2011). Em LU, o aumento da atividade tirosina quinase de FGFR1, juntamente com o aumento de IGFBP5 em resposta ao TGF-beta poderia desencadear a proliferação celular identificada nestes tumores, porém, o aumento da atividade de COL3A1, favoreceria o acúmulo de matriz extracelular e deposição de colágeno. Este mecanismo explicaria em partes, os baixos índices mitóticos e a característica fibróide identificada em Leiomiomas Uterinos.

Portanto, esse achado demonstrou uma grande semelhança molecular de LU com fibrose hepática para o qual já existem testes terapêuticos. Estudos adicionais com modelos *in vivo* são necessários para avaliar a taxa de resposta a drogas inibidoras de proliferação e crescimento celular em Leiomiomas Uterinos.

6. Conclusões

- ✓ A análise de alterações genômicas pela metodologia de CGH array permitiu identificar variações no número de cópias de DNA submicroscópicas em todos os casos de LU, sendo consideradas como relevantes aquelas presentes em, no mínimo, 10% dos casos.
- ✓ As regiões frequentemente envolvidas em ganhos genômicos foram 2q35, 6p21.32, 7q22.1 e 11q12.3-q13.2 e perdas em 4q28.3, 7q22.1-q22.2 e 7q31.33, revelando novas regiões alteradas e confirmando achados prévios descritos em literatura.
- ✓ A análise dos dados de CGH array por agrupamento hierárquico não supervisionado demonstrou uma tendência a separar dois grupos de amostras de acordo com a presença ou ausência de tumores múltiplos; revelando ganhos em 2q35, 7q22.1, 19p13.12-p13.11 e 19q13.32-q13.33 no primeiro grupo e perdas em 7q22.1-q22.2 no segundo grupo.
- ✓ Um subconjunto de genes envolvidos em regiões de ganhos genômicos foi associado à regulação da transcrição, incluindo genes regulados e reguladores do complexo da RNA Pol II em resposta ao estrógeno, sendo este, um novo achado do presente estudo.
- ✓ A análise funcional *in silico* revelou que genes mapeados em regiões de CNVs significativas desempenham funções associadas ao ciclo celular; duplicação do DNA; recombinação e reparo; sinalização celular envolvendo as vias de sinalização IGF-1, TGF-beta e do receptor de estrógeno.
- ✓ O perfil de expressão gênica global não demonstrou diferenças quanto à fase do ciclo menstrual, sugerindo ser um fator independente para o desenvolvimento dos Leiomiomas Uterinos.
- ✓ A análise funcional *in silico* dos dados de expressão gênica demonstrou que genes com expressão diminuída (66%) atuam nos processos de recombinação

cromossômica, reparo a danos no DNA e indução da apoptose, enquanto que genes com expressão aumentada (34%) atuam na proliferação e acúmulo da matriz extracelular. A ação conjunta destes grupos de genes pode contribuir para a transformação de células normais do miométrio em células neoplásicas.

- ✓ A integração dos dados genômicos e transcriptômicos aplicada neste estudo detectou 75 genes, classificados como potenciais modulares do fenótipo tumoral, entre os quais 26 tiveram correlação positiva e três apresentaram correlação negativa.
- ✓ A análise de enriquecimento de genes revelou que 32 dos 75 genes estavam associados a câncer. Os genes *FGFR1*, *DBN1*, *NUPR1*, *COL3A1*, *AOC3*, *IFITM1*, *CALCRL*, *NOVA2*, *MVP*, *MFAP5*, *DIP2C* apresentaram correlação positiva e *CENPF*, *RHOH* e *BICD1* apresentaram correlação negativa ($P < 0,05$).
- ✓ A análise funcional dos 75 genes moduladores demonstrou que *COL3A1*, *IGFBP5* e *FGFR1* atuam em vias de proliferação e da matriz extracelular, exibindo atividades indiretas na transcrição de genes alvos. Em adição, *FGFR1* foi previamente associado a outras doenças humanas. Este gene pode ser um importante alvo terapêutico para o tratamento dos Leiomiomas Uterinos.

7. Referências Bibliográficas

- Acevedo VD and Ittmann M. Paths of FGFR-driven tumorigenesis. *Cell Cycle*, 8, 580-588, 2009.
- Ahn WS, Kim KW, Bae SM et al. Targeted cellular process profiling approach for uterine leiomyoma using cDNA microarray, proteomics and gene ontology analysis. *Int J Exp Path*, 84, 267-279, 2003.
- Akavia UD, Litvin O, Kim J et al. An integrated approach to uncover drivers of cancer. *Cell*, 143, 1005-1017, 2010.
- Alderton GK. Common disease pathogenesis pathways. *Nat Rev Cancer*, 10, Research Highlights, 2010.
- Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large β -cell lymphoma identified by gene expression profiling. *Nature*, 403, 503-511, 2000.
- Alizadeh AA, Ross DT, Perou CM et al. Towards a novel classification of human malignancies based on gene expression patterns. *J Pathol*, 195, 41-52, 2001.
- Alvarez H, Opalinska J, Zhou L et al. Widespread Hypomethylation Occurs Early and Synergizes with Gene Amplification during Esophageal Carcinogenesis. *PLoS Genet*, 7, e1001356, 2011.
- Anania CA, Stewart EA, Quade BJ et al. Expression of the fibroblast growth factor receptor in women with leiomyomas and abnormal uterine bleeding. *Mol Hum Reprod*, 3, 685-691, 1997.
- Annereau JP, Szakacs G, Tucker CJ et al. Analysis of ATPbinding cassette transporter expression in drug-selected cell lines by a microarray dedicated to multidrug resistance. *Mol Pharmacol*, 66, 1397-1405, 2004.
- Arslan AA, Gold LI, Mittal K et al. Gene expression studies provide clues to the pathogenesis of uterine leiomyoma: new evidence and a systematic review. *Human Reproduction*, 20, 852-863, 2005.
- Arvanitis DA and Spandidos DA. Dereglulation of the G1/S phase transition in cancer and squamous intraepithelial lesions of the uterine cervix: a case control study. *Oncol Rep*, 20, 751-760, 2008.
- Asaoka Y, Ikenoue T, Koike K. New targeted therapies for gastric cancer. *Expert Opin Investig Drugs*. May;20, 595-604, 2011.
- Aspenström P, Ruusala A, Pacholsky D Taking Rho GTPases to the next level: the cellular functions of atypical Rho GTPases. *Exp Cell Res*, 15, 313, 3673-3679, 2007.
- Bahamondes L, Bahamondes MV, Monteiro I. Levonorgestrel-releasing intrauterine system: uses and controversies. *Expert Rev Med Devices*, 5, 437-445, 2008.
- Bakas P, Liapis A, Vlahopoulos S et al. Estrogen receptor alpha and beta in uterine fibroids: a basis for altered estrogen responsiveness. *Fertil Steril*, 90, 1878-1885, 2008.

- Baker ME. Evolution of 17 beta-hydroxysteroid dehydrogenases and their role in androgen, estrogen and retinoid action. *Mol Cell Endocrinol*, 171, 211-215, 2001.
- Barbieri RL and Andersen J. Uterine leiomyomas. The somatic mutation theory. *Sem Reprod Endocrinol*, 10, 301-309, 1992.
- Bayani J and Squire JA. Comparative genomic hybridization. *Curr Protoc Cell Biol*, Chapter 22, Unit 22.6, 2005.
- Bayani J, Marrano P, Graham C. Genomic instability and copy-number heterogeneity of chromosome 19q, including the kallikrein locus, in ovarian carcinomas. *Mol Oncol*, 5, 48-60, 2011.
- Behera MA, Feng L, Yonish B et al. Thrombospondin-1 and Thrombospondin-2 mRNA and TSP-1 and TSP-2 protein expression in uterine fibroids and correlation to the genes COL1A1 and COL3A1 and to the collagen cross-link ydroxyproline. *Reprod Sci*, 14, 63-76, 2007.
- Beheshti B, Park PC, Braude I et al. Microarray CGH. *Methods Mol Biol*, 204, 191-207, 2002.
- Beroukhim R, Getz G, Nghiemphu L et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. *PNAS*, 104, 20007-20012, 2007.
- Bhide RS, Lombardo LJ, Hunt JT et al. The antiangiogenic activity in xenograft models of brivanib, a dual inhibitor of vascular endothelial growth factor receptor-2 and fibroblast growth factor receptor-1 kinases. *Mol Cancer Ther*, 9, 369-378, 2010.
- Bisognin A, Coppe A, Ferrari F et al. A-MADMAN: annotation-based microarray data meta-analysis tool. *BMC Bioinformatics*, 10, 201, 2009.
- Bisso A, Collavin L, Del Sal G. p73 as a Pharmaceutical Target for Cancer Therapy *Curr Pharm Des*, Epub ahead of print, 2011.
- Bodner-Adler B, Bodner K, Czerwenka K et al. Expression of p16 protein in patients with uterine smooth muscle tumors: an immunohistochemical analysis. *Gynecol Oncol*, 96, 62-66, 2005.
- Bonatz G, Frahm SO, Andreas S et al. Telomere shortening in uterine leiomyomas. *Am J Obstet Gynecol*, 179, 591-596, 1998.
- Bonazzi VF, Irwin D, Hayward NK. Identification of candidate tumor suppressor genes inactivated by promoter methylation in melanoma. *Genes Chromosomes Cancer*, 48, 10-21, 2009.
- Borsari R, Bozzini N, Junqueira CR, et al. Genic expression of the uterine leiomyoma in reproductive-aged women after treatment with goserelin. *Fertil Steril*, 94, 1072-1077, 2010.
- Boveri T. *Zur Frage der Entstehung maligner Tumoren*. (Fischer, Jena, G.D.R.), 1914.
- Bowden W, Skorupski J, Kovanci E et al. Detection of novel copy number variants in uterine leiomyomas using high-resolution SNP arrays. *Mol Hum Reprod*, 15, 563-568, 2009.
- Bozzini N, Rodrigues CJ, Petti DA et al. Effects of treatment with gonadotropin releasing hormone agonist on the uterine leiomyomata structure. *Acta Obstet Gynecol Scand*, 82, 330-334, 2003.

- Brazelle W, Krehling JM, Gemmer J et al. Histone deacetylase inhibitors downregulate checkpoint kinase 1 expression to induce cell death in non-small cell lung cancer cells. *PLoS One*, 5, e14335, 2010.
- Bridges CB. The Bar's gene: a duplication. *Science*, 83, 210–211, 1936.
- Brosens I, Deprest J, Dal Cin P et al. Clinical significance of cytogenetic abnormalities in uterine myomas. *Fertil Steril*, 69, 232-235, 1998.
- Buckanovich RJ, Posner JB, Darnell RB. Nova, the paraneoplastic Ri antigen, is homologous to an RNA-binding protein and is specifically expressed in the developing motor system. *Neuron*, 11, 657–672, 1993.
- Burroughs KD, Howe SR, Okubo Y et al. Dysregulation of IGF-I signaling in uterine leiomyoma. *J Endocrinol*, 172, 83-93, 2002.
- Cai J, Jiang Y, Zhang Y. Quantitation of Ras p21 oncogene and DNA content of gastrointestinal smooth muscle tumors and prognostic significance. *Zhonghua Wai Ke Za Zhi*, 33, 225-228, 1995.
- Cai YR, Diao XL, Wang SF et al. X-chromosomal inactivation analysis of uterine leiomyomas reveals a common clonal origin of different tumor nodules in some multiple leiomyomas. *Int J Oncol*, 31, 1379-1389, 2007.
- Caldas C and Brenton JD. Sizing up miRNAs as cancer genes. *Nat Med*, 11, 712-714, 2005.
- Calin GA, Sevignani C, Dumitru CD et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci*, 2999-3004, 2004.
- Canevari RA, Pontes A, Rosa FE et al. Independent clonal origin of multiple uterine leiomyomas that was determined by X chromosome inactivation and microsatellite analysis. *Am J Obstet Gynecol*, 193, 1395–1403, 2005a.
- Canevari RA, Pontes A, Rogatto SR. Microallelotyping Defines Novel Regions of Loss of Heterozygosity in Uterine Leiomyomas. *Molecular Carcinogenesis*, 42, 177–182, 2005b.
- Canevari, RA and Rogatto SR. Uterine Leiomyoma: Updates in Cytogenetics and Molecular Analysis. *Asian Journal of Cancer*, 6, 15-28, 2007.
- Carter, NP. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat Genet*, 39, S16–S21, 2007.
- Carvalho B, Marques CJ, Carvalho D et al. Novel human pathological mutations. Gene symbol: CYP21A2. Disease: Adrenal Hyperplasia. *Hum Genet*, 127, 482-483, 2010.
- Castillo SD, Angulo B, Suarez-Gauthier A et al. Gene amplification of the transcription factor DP1 and CTNND1 in human lung cancer. *J Pathol*, 222, 89-98, 2010.
- Catherino WH, Prupas C, Tsibris JC et al. Strategy for elucidating differentially expressed genes in leiomyomata identified by microarray technology. *Fertil Steril*, 80, 282-290, 2003.

- Catherino W, Salama A, Potlog-Nahari C et al. Gene expression studies in leiomyomata: new directions for research. *Semin Reprod Med*, 22, 83–90, 2004a.
- Catherino WH, Leppert PC, Stenmark MH et al. Reduced dermatopontin expression is a molecular link between uterine leiomyomas and keloids. *Genes Chromosomes Cancer*, 40, 204-217, 2004b.
- Chegini N, Tang XM, Ma C. Regulation of transforming growth factor-beta1 expression by granulocyte macrophage-colony-stimulating factor in leiomyoma and myometrial smooth muscle cells. *J Clin Endocrinol Metab*, 84, 4138–4143, 1999.
- Chegini N, Ma C, Tang XM et al. Effects of GnRH analogues, ‘add-back’ steroid therapy, antiestrogen and antiprogestins on leiomyoma and myometrial smooth muscle cell growth and transforming growth factor-beta expression. *Mol Hum Reprod*, 8, 1071–1078, 2002.
- Chegini N, Verala J, Luo X, et al. Gene expression profile of leiomyoma and myometrium and the effect of gonadotropin releasing hormone analogue therapy. *J Soc Gynecol Investig*, 10, 161-171, 2003.
- Chen W, Yoshida S, Ohara N et al. Gonadotropin-releasing hormone antagonist cetrorelix down-regulates proliferating cell nuclear antigen and epidermal growth factor expression and up-regulates apoptosis in association with enhanced poly(adenosine 5'-diphosphate-ribose) polymerase expression in cultured human leiomyoma cells. *J Clin Endocrinol Metab*, 90, 884-892, 2005.
- Cheng X and Xu G. A systemic review of the relationship between aldosterone synthase - 344 C/T polymorphism and hypertension in Han. *Clin Exp Hypertens*, 32, 301-307, 2010.
- Cho JH, Lee JG, Yang YI et al. Eupatilin, a dietary flavonoid, induces G2/M cell cycle arrest in human endometrial cancer cells. *Food Chem Toxicol*, Epub ahead of print, 2011.
- Cho YL, Bae S, Koo MS et al. Array comparative genomic hybridization analysis of uterine leiomyosarcoma. *Gynecol Oncol*, 99, 545-551, 2005.
- Choi SH, Kim YT, Han KN. Dis Esophagus. Surgical management of the esophageal leiomyoma: lessons from a retrospective review, Epub ahead of print, 2011.
- Choi SI, Yoo YM, Kim BY et al. Involvement of TGF- β receptor- and integrin-mediated signaling pathways in the pathogenesis of granular corneal dystrophy II. *Invest Ophthalmol Vis Sci*, 51, 1832-1847, 2010.
- Choi YW, Bae SM, Kim YW et al. Gene expression profiles in squamous cell cervical carcinoma using array-based comparative genomic hybridization analysis. *Int J Gynecol Cancer*, 17, 687-696, 2007.
- Chowdhury UR, Samant RS, Fodstad O et al. Emerging role of nuclear protein 1 (NUPR1) in cancer biology. *Cancer Metastasis Rev*, 28, 225-232, 2009.
- Christacos NC, Quade BJ, Dal Cin P et al. Uterine Leiomyoma with deletions of 1p represent a distinct cytogenetic subgroup associated with unusual histologic features. *Genes Chromosomes Cancer*, 45, 304-312, 2006.

- Churikov D and Price CM. Pot1 and cell cycle progression cooperate in telomere length regulation. *Nat Struct Mol Biol*, 15, 79-84, 2008.
- Cirilo PDR, Rosa LAR, Drigo SA et al. Evaluation of gene and protein expression of estrogen receptors alpha, beta and progesterone receptors in uterine leiomyomas, in progress.
- Collins NB, Wilson JB, Bush T et al. ATR-dependent phosphorylation of FANCA on serine 1449 after DNA damage is important for FA pathway function. *Blood*, 113, 2181-2190, 2009.
- Conrad DF, Andrews TD, Carter NP et al. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet*, 38, 75-81, 2006.
- Cook H, Ezzati M, Segars JH et al. The impact of uterine leiomyomas on reproductive outcomes. *Minerva Ginecol*, 62, 225-236, 2010.
- Coronado GD, Marshall LM, Schwartz SM. Complications in pregnancy, labor, and delivery with uterine leiomyomas: a population-based study. *Obstet Gynecol*, 95, 764-769, 2000.
- Cotran RS, Kumar V, Robbins SL. Robbins. Pathologic basis of disease. Philadelphia: W.B. Saunders Company, 1989.
- Crabtree JS, Jelinsky SA, Harris HA et al. Comparison of human and rat uterine leiomyomata: identification of a dysregulated mammalian target of rapamycin pathway. *Cancer Res*, 69, 6171-6178, 2009.
- Cramer S and Patel A. The frequency of uterine leiomyomas. *Am J Clin Pathol*, 94, 435-438, 1990.
- Critchley HO, Robertson KA, Forster T et al. Gene expression profiling of mid to late secretory phase endometrial biopsies from women with menstrual complaint. *Am J Obstet Gynecol*, 195, 406, e1-16, 2006.
- Cvekl A Jr, Zavadil J, Birshtein BK et al. Analysis of transcripts from 17p13.3 in medulloblastoma suggests ROX/MNT as a potential tumour suppressor gene. *Eur J Cancer*, 40, 2525-2532, 2004.
- DATASUS - (<http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sih/cnv/qiuf.def>). Acesso em Março de 2011.
- Day-Baird D, Dunson DB, Hill MC et al. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. *Am J Obstet Gynecol* 188, 100-107, 2003.
- de Smith AJ, Tsalenko A, Sampas N et al. Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: Implications for association studies of complex diseases. *Hum Mol Genet*, 16, 2783-2794, 2007.
- Dimitrova IK, Richer JK, Rudolph MC et al. Gene expression profiling of multiple leiomyomata uteri and matched normal tissue from a single patient. *Fertil Steril*, 91, 2650-2663, 2009.
- Ding L, Xu J, Luo L et al. Gonadotropin releasing hormone and transforming growth factor β activate MAPK/ERK and differentially regulate fibronectin, type I collagen, and PAI-1 expression in

- leiomyoma and myometrial smooth muscle cells. *J Clin Endocrinol Metab*, 89, 5549–5557, 2004.
- Doridot V, This P, Clough KB. Explosive growth of uterine leiomyomas and carcinologic ovarian risk in a non-menopausal patient with BRCA1-BRCA2 mutation treated by tamoxifen. *J Gynecol Obstet Biol Reprod*, 31, 677-680, 2002.
- Douglas EJ, Fiegler H, Rowan A et al. Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. *Cancer Res*, 64, 4817-4825, 2004.
- Duhan N and Sirohiwal D. Uterine myomas revisited. *Eur J Obstet Gynecol Reprod Biol*, 152, 119-125, 2010.
- Eifel P, Axelson JA, Costa J et al. National Institutes of Health consensus development conference statement: adjuvant therapy for breast cancer. *J Natl Cancer Inst*, 93, 979–989, 2001.
- El-Gharib MN and Elsobky ES. Cytogenetic aberrations and the development of uterine leiomyomata. *J Obstet Gynaecol Res*, 36, 101–107, 2010.
- Englund K, Lindblom B, Carlstrom K et al. Gene expression and tissue concentrations of IGF-I in human myometrium and fibroids under different hormonal conditions. *Mol Hum Reprod*, 6, 915–920, 2000.
- Farquhar CM and Steiner CA. Hysterectomy rates in the United States 1990–1997. *Obstet Gynecol*, 99, 229–34, 2002.
- FEBRASGO - Manual de Orientação – Leiomioma Uterino. Federação Brasileira das Sociedades de Ginecologia e Obstetrícia, 117pp, 2004.
- Flake GP, Andersen J, Dixon D. Etiology and pathogenesis of uterine leiomyomas: a review. *Environ Health Perspect*, 111, 1037-1054, 2003.
- Floyd RV, Wray S, Quenby S et al. Expression and distribution of Na, K-ATPase isoforms in the human uterus. *Reprod Sci*, 17, 366-376, 2010.
- Forbes SA, Pannett AA, Bassett JH et al. Mapping of the gene encoding the B56 beta subunit of protein phosphatase 2A (PPP2R5B) to a 0.5-Mb region of chromosome 11q13 and its exclusion as a candidate gene for multiple endocrine neoplasia type 1 (MEN1). *Hum Genet*, 100, 481-485, 1997.
- Fotiou SK, Tserkezoglou AJ, Mahera H et al. Chromosome aberrations and expression of ras and myc oncogenes in leiomyomas and a leiomyosarcoma of the uterus. *Eur J Gynaecol Oncol*, 1, 340-345, 1992.
- Fusco A and Fedele M. Roles of HMGA proteins in cancer. *Nat Rev Cancer*, 7, 899-910, 2007.
- Gajria D and Chandarlapaty S. HER2-amplified breast cancer: mechanisms of trastuzumab resistance and novel targeted therapies. *Expert Rev Anticancer Ther*, 11, 263-275, 2011.
- Galiègue-Zouitina S, Delestré L, Dupont C et al. Underexpression of RhoH in Hairy Cell Leukemia. *Cancer Res*, 68, 4531-4540, 2008.

- Gannon BR, Manduch M, Childs TJ. Differential Immunoreactivity of p16 in leiomyosarcomas and leiomyoma variants. *Int J Gynecol Pathol*, 27, 68-73, 2008.
- Gao Z, Matsuo H, Wang Y et al. Up-regulation by IGF-I of proliferating cell nuclear antigen and Bcl-2 protein expression in human uterine leiomyoma cells. *J Clin Endocrinol Metab*, 86, 5593-5599, 2001.
- Garcia-Torres R, Cruz D, Orozco L et al. Alport syndrome and diffuse leiomyomatosis. Clinical aspects, pathology, molecular biology and extracellular matrix studies. A synthesis. *Nephrologie*, 21, 9-12, 2000.
- Gevaert O, De Smet F, Timmerman D et al. Predicting the prognosis of breast cancer by integrating clinical and microarray data with bayesian networks. *Bioinformatics*, 22, e184-190, 2006.
- Giovannone B, Lee E, Laviola L et al. Two novel proteins that are linked to insulin-like growth factor (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling. *J Biol Chem*, 278, 31564-31573, 2003.
- Gökaslan H, Türkeri L, Kavak ZN et al. Differential diagnosis of smooth muscle tumors utilizing p53, pTEN and Ki-67 expression with estrogen and progesterone receptors, *Gynecol Obstet Invest*, 59, 36-40, 2005.
- 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*, 65, 7127-7136, 2005.
- Goodison S, Sun Y, Urquidi V. Derivation of cancer diagnostic and prognostic signatures from gene expression data. 2, 855-862, 2010.
- Green MR, Aya-Bonilla C, Gandhi MK et al. Integrative genomic profiling reveals conserved genetic mechanisms for tumorigenesis in common entities of non-hodgkin's lymphoma. *Genes Chrom Cancer*, 50, 313-326, 2011.
- Grimwood J, Gordon LA, Olsen A et al. The DNA sequence and biology of human chromosome 19. *Nature*, 428, 529-535, 2004.
- Gross KL and Morton CC. Genetics and the development of fibroids. *Clin Obst Gynecol*, 44, 335-349, 2001.
- Grützmann R, Pilarsky C, Ammerpohl O et al. Gene expression profiling of microdissected pancreatic ductal carcinomas using high-density DNA microarrays. *Neoplasia*, 6, 611-622, 2004.
- Gustavsson I, Englund K, Faxén M et al. Tissue differences but limited sex steroid responsiveness of c-fos and c-jun in human fibroids and myometrium. *Mol Hum Reprod*, 6, 55-59, 2000.
- Hague S, Zhang L, Oehler MK et al. Expression of the hypoxically regulated angiogenic factor adrenomedullin correlates with uterine leiomyoma vascular density. *Clin Cancer Res*, 3, 2808-2814, 2000.

- Hakki T and Bernhardt R. CYP17- and CYP11B-dependent steroid hydroxylases as drug development targets. *Pharmacol Ther*, 111, 27-52, 2006.
- Han JH, Lee S, Park YS et al. IFITM6 expression is increased in macrophages of tumor-bearing mice. *Oncol Rep*, 25, 531-536, 2011.
- Harrison-Woolrych ML, Sharkey AM, Charnock-Jones DS et al. Localization and quantification of vascular endothelial growth factor messenger ribonucleic acid in human myometrium and leiomyomata. *J Clin Endocrinol Metab*, 80, 1853-1858, 1995.
- Hastings PJ, Lupski JR, Rosenberg SM et al. Mechanisms of change in gene copy number. *Nat Rev Genet*, 10, 551-564, 2009.
- Heim S, Nilbert M, Vanni et al. A specific translocation, t (12;14) (q14-15;q23-24), characterizes a subgroup of uterine leiomyomas. *Cancer Genet Cytogenet*, 32, 13-17, 1988.
- Heinrich UR, Brieger J, Gosepath J et al. Frequent chromosomal gains in recurrent juvenile nasopharyngeal angiofibroma. *Cancer Genet Cytogenet*, 175, 138-143, 2007.
- Heinzen EL, Yoon W, Tate SK et al. Nova2 interacts with a cis-acting polymorphism to influence the proportions of drug-responsive splice variants of SCN1A. *Am J Hum Genet*, 80, 876-883, 2007.
- Hennig Y, Deichert U, Bonk U et al. Chromosomal translocations affecting 12q14-15 but not deletions of the long arm of chromosome 7 associated with a growth advantage of uterine smooth muscle cells. *Mol Hum Reprod*, 5, 1150-1154, 1999.
- Hess JL. Chromosomes translocation in benign tumors: the HMGI proteins. *Am J Clin Pathol*, 109, 251-261, 1998.
- Hirsch HA, Iliopoulos D, Joshi A et al. A transcriptional signature and common gene networks link cancer with lipid metabolism and diverse human diseases. *Cancer Cell*, 17, 348-361, 2010.
- Hodge JC and Morton CC. Genetic heterogeneity among uterine leiomyomata: insights into malignant progression. *Hum Mol Genet*, 16, R7-R13, 2007.
- Hodge JC, Quade BJ, Rubin MA et al. Molecular and cytogenetic characterization of plexiform leiomyomata provide further evidence for genetic heterogeneity underlying uterine fibroids. *Am J Pathol*, 172, 1403-1410, 2008.
- Hodge JC, Park PJ, Dreyfuss JM et al. Identifying the molecular signature of the interstitial deletion 7q subgroup of uterine leiomyomata using a paired analysis. *Genes Chromosomes Cancer*, 48, 865-885, 2009.
- Hoeijmakers JHJ. Genome maintenance mechanisms for preventing cancer. *Nature*, 411, 366-374, 2001.
- Hoffman PJ, Milliken DB, Gregg LC et al. Molecular characterization of uterine fibroids and its implication for underlying mechanisms of pathogenesis. *Fertil Steril*, 636-649, 2004.

- Holthausen JT, Wyman C, Kanaar R. Regulation of DNA strand exchange in homologous recombination. *DNA Repair (Amst)*, 9, 1264-1272, 2010.
- Hormigo A, Dalmau J, Rosenblum MK et al. Immunological and pathological study of anti-Ri-associated encephalopathy. *Ann Neurol*, 36, 896-902, 1994.
- Hou XJ, Zhang YZ, Liu X et al. Expressions of IGFBP-5, cFLIP in cervical intraepithelial neoplasia, cervical carcinoma and their clinical significances: a molecular pathology. *J Exp Clin Cancer Res*, 28, 1-16, 2009.
- Hromas R, Broxmeyer HE, Kim C et al. Cloning of BRAK, a novel divergent CXC chemokine preferentially expressed in normal versus malignant cells. *Biochem Biophys Res Commun*, 255, 703-706, 1999.
- Hu J and Surti U. Subgroups of uterine leiomyomas based on cytogenetic analysis. *Hum. Pathol*, 22, 1009-1016, 1991.
- Huang CS, Shi SH, Ule J et al. Common molecular pathways mediate long-term potentiation of synaptic excitation and slow synaptic inhibition. *Cell*, 123, 105-118, 2005.
- Huang XF and Luu-The V. Molecular characterization of a first human 3(alpha->beta)-hydroxysteroid epimerase. *J Biol Chem*, 275, 29452-29457, 2000.
- Huang Y, Shen XJ, Zou Q et al. Biological functions of microRNAs. *Bioorg Khim*, 36, 747-752, 2010.
- Hundal HS, Maxwell DL, Ahmed A et al. Subcellular distribution and immunocytochemical localization of Na,K-ATPase subunit isoforms in human skeletal muscle. *Mol Membr Biol*, 11, 255-262, 1994.
- Huppi K, Volfovsky N, Runfola T et al. The identification of microRNAs in a genomically unstable region of human chromosome 8q24. *Mol Cancer Res*, 6, 212-221, 2008.
- Inoue K and Lupski JR. Molecular mechanisms for genomic disorders. *Annu Rev Genomics, Hum Genet*, 3, 199-242, 2002.
- Ingraham SE, Lynch RA, Surti U et al. Identification and characterization of novel human transcripts embedded within HMGA2 in t(12;14)(q15;q24.1) uterine leiomyoma. *Mutat Res*, 602, 43-53, 2006.
- Iovanna JL. Expression of the stress-associated protein p8 is a requisite for tumor development. *Int J Gastrointest Cancer*, 31, 89-98, 2002.
- Ishikawa H, Ishi K, Serna VA et al. Progesterone is essential for maintenance and growth of uterine leiomyoma. *Endocrinology*, 151, 2433-2442, 2010.
- Ito K, Suzuki T, Akahira J et al. 14-3-3sigma in endometrial cancer--a possible prognostic marker in early-stage cancer. *Clin Cancer Res*, 11, 7384-7391, 2005.
- Jaakkola K, Kaunismäki K, Tohka S et al. Human vascular adhesion protein-1 in smooth muscle cells. *Am J Pathol*, 155, 1953-1965, 1999.

- Jacobson EM, Hugo ER, Borcharding DC et al. Prolactin in breast and prostate cancer: molecular and genetic perspectives. *Discov Med*, 11, 315-324, 2011.
- Jacoby VL, Fujimoto VY, Giudice LC et al. Racial and ethnic disparities in benign gynecologic conditions and associated surgeries. *Am J Obstet Gynecol*, 202, 514-521, 2010.
- Jashnani KD, Agrawal VR, Naik LP. Metastasis of breast lobular carcinoma in a cervical leiomyoma. *Indian J Pathol Microbiol*, 50, 378-379, 2007.
- Jensen KB, Dredge BK, Stefani G et al. Nova-1 regulates neuron-specific alternative splicing and is essential for neuronal viability. *Neuron*, 25, 359-371, 2000.
- Jeong YJ, Noh EM, Lee YR et al. 17beta-estradiol induces up-regulation of PTEN and PPARgamma in leiomyoma cells, but not in normal cells. *Int J Oncol*, 36, 921-927, 2010.
- Jiang Y, Suo G and Sadarangani A. Expression profiling of protein tyrosine kinases and their ligand activators in leiomyoma uteri. *Syst Biol Reprod Med*, 56, 318-326, 2010.
- Jin M, Tanaka S, Sekino Y et al. A novel, brain-specific mouse drebrin: cDNA cloning, chromosomal mapping, genomic structure, expression, and functional characterization. *Genomics*, 79, 686-692, 2002.
- Jones MR, Italiano L, Wilson SG et al. Polymorphism in HSD17B6 is associated with key features of polycystic ovary syndrome. *Fertil Steril*, 86, 1438-1446, 2006.
- Jönsson G, Staaf J, Vallon-Christersson J et al. Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics. *Breast Cancer Res*, 12, 2010.
- Kallioniemi A, Kallioniemi OP, Sudar D et al. Comparative genomic hybridization for molecular cytogenetics analysis of solid tumors. *Science*, 258, 818-821, 1992.
- Kanamori T, Takakura K, Mandai M et al. PEP-19 overexpression in human uterine leiomyoma. *Mol Hum Reprod*, 9, 709-717, 2003.
- Kandalaft PL and Esteban JM. Bilateral massive ovarian leiomyomata in a young woman: a case report with review of the literature. *Mod Pathol*, 5, 586-589, 1992.
- Katoh M. Genetic alterations of FGF receptors: an emerging field in clinical cancer diagnostics and therapeutics. *Expert Rev Anticancer Ther*, 10, 1375-1379, 2010.
- Käufel SD, Makovitzky J, Kuhn C et al. Inhibin/activin-betaC subunit in human endometrial adenocarcinomas and HEC-1a adenocarcinoma cell line. *In Vivo*, 24, 695-698, 2010.
- Kawaguchi K, Fujii S, Konishi I et al. Ultrastructural study of cultured smooth muscle cells from uterine leiomyoma and myometrium under the influence of sex steroids. *Gynecol Oncol*, 21, 32-41, 1985.
- Kayisli UA, Berkkanoglu M, Kizilay G et al. Expression of proliferative and preapoptotic molecules in human myometrium and leiomyoma throughout the menstrual cycle. *Reprod Sci*, 14, 678-686, 2007.

- Kazmierczak B, Bol S, Wanschura S et al. Pac clone containing the HMGI(Y) gene spans the breakpoint of a 6p21 translocation in a uterine leiomyoma cell line. *Genes Chromosomes Cancer*, 17, 191-193, 1996.
- Kazmierczak B, Wanschura S, Rosigkeit J et al. Molecular characterization of 12q14-15 rearrangements in three pulmonary chondroid hamartomas. *Cancer Res*, 55, 2497-2499, 1995.
- Ke L, Che YN, Cao YX et al. Polymorphisms of the HSD17B6 and HSD17B5 genes in Chinese women with polycystic ovary syndrome. *J Womens Health*, 19, 2227-2232, 2010.
- Khorram O, Garthwaite M, Golos T. Uterine and ovarian aryl hydrocarbon receptor (AHR) and aryl hydrocarbon receptor nuclear translocator (ARNT) mRNA expression in benign and malignant gynaecological conditions. *Mol Hum Reprod*, 8, 75-80, 2002.
- Kimmich T, Brüning A, Käußl SD et al. Inhibin/activin-betaC and -betaE subunits in the Ishikawa human endometrial adenocarcinoma cell line. *Arch Gynecol Obstet*, 282, 185-191, 2010.
- Kondo NI, Yoshida S, Kajiyama H et al. Metastasis of breast cancer to a uterine leiomyoma. *Breast Cancer*, 16, 157-161, 2009.
- Kools PF and Van de Ven WJ. Amplification of a rearranged form of the high mobility group protein gene hmgic in osteosarcoma cells. *Cancer Genet Cytogenet*, 91, 1-7, 1996.
- Kroeger H, Jelinek J, Estécio MR. Aberrant CpG island methylation in acute myeloid leukemia is accentuated at relapse. *Blood*, 112, 1366-1373, 2008.
- Kuivaniemi H, Tromp G, Prockop DJ. Mutations in collagen genes: causes of rare and some common diseases in humans. *FASEB J*, 5, 2052-2060, 1991.
- Laganà A, Russo F, Sismeiro C et al. Variability in the incidence of miRNAs and genes in fragile sites and the role of repeats and CpG islands in the distribution of genetic material. *PLoS One*, 17, e11166, 2010.
- Lara PC, Lloret M, Clavo B et al. Severe hypoxia induces chemo-resistance in clinical cervical tumors through MVP over-expression. *Radiat Oncol*, 6, 1-14, 2009.
- Laser L, Lee P and Wei J. Cellular senescence in usual type uterine leiomyoma. *Fertil Steril*, 93, 2020-2026, 2010.
- Laurila E, Savinainen K, Kuuselo R et al. Characterization of the 7q21-q22 amplicon identifies ARPC1A, a subunit of the Arp2/3 complex, as a regulator of cell migration and invasion in pancreatic cancer. *Genes Chromosomes Cancer*, 48, 330-339, 2009.
- Lavecchia A, Di Giovanni C, Novellino E. Inhibitors of Cdc25 phosphatases as anticancer agents: a patent review. *Expert Opin Ther Pat*, 20, 405-425, 2010.
- Lea IA, Jackson MA, Li X et al. Genetic pathways and mutation profiles of human cancers: site- and exposure-specific patterns. *Carcinogenesis*, 28, 1851-1858, 2007.

- Lee C and Scherer SW. The clinical context of copy number variation in the human genome. *Expert Rev Mol Med*, 12, 1-29, 2010.
- Lee EJ, Kong G, Lee SH et al. Profiling of differentially expressed genes in human uterine leiomyomas. *Int J Gynecol Cancer*, 15, 146-154, 2005.
- Lee M and Kim Y. CHESS (CgHExpreSS): A comprehensive analysis tool for the analysis of genomic alterations and their effects on the expression profile of the genome. *BMC Bioinformatics*, 16, 1-8, 2009.
- Lehtonen HJ. Hereditary leiomyomatosis and renal cell cancer: update on clinical and molecular characteristics. *Fam Cancer*, Epub ahead of print , 2011.
- Lehtonen HJ, Kiuru M, Ylisaukko-Oja SK et al. Increased risk of cancer in patients with fumarate hydratase germline mutation. *J Med Genet*, 43, 523-526, 2006.
- Lehtonen R, Kiuru M, Vanharanta S et al. Biallelic inactivation of fumarate hydratase (FH) occurs in nonsyndromic uterine leiomyomas but is rare in other tumors. *Am J Pathol*, 164, 17-22, 2004.
- Leppert PC, Catherino WH, Segars JH. A new hypothesis about the origin of uterine fibroids based on gene expression profiling with microarrays. *Am J Obstet Gynecol*, 195, 415-420, 2006.
- Lettau I, Hattermann K, Held-Feindt J et al. Matrix metalloproteinase-19 is highly expressed in astroglial tumors and promotes invasion of glioma cells. *J Neuropathol Exp Neurol*, 69, 215-223, 2010.
- Levy B, Mukherjee T, Hirschhorn K. Molecular cytogenetic analysis of uterine leiomyomata and leiomyosarcoma by comparative genomic hybridization. *Cancer Genet Cytogenet*, 121, 1-8, 2000.
- Li Z, Li W, Xie J et al. Epigenetic inactivation of PCDH10 in human prostate cancer cell lines. *Cell Biol Int*, Epub ahead of print, 2011.
- Ligon AH and Morton CC. Genetics of uterine leiomyomata. *Genes Chromosomes Cancer*, 28, 235-245, 2000.
- Ligon AH and Morton CC. Leiomyomata: heritability and cytogenetic studies. *Hum Reprod Update*, 7, 8-14, 2001.
- Ligon AH, Scott IC, Takahara K et al. PCOLCE deletion and expression analyses in uterine leiomyomata. *Cancer Genet Cytogenet*, 137, 133-137, 2002.
- Lin N, Chen S, Pan W et al. NP603, a novel and potent inhibitor of FGFR1 tyrosine kinase, inhibits hepatic stellate cell proliferation and ameliorates hepatic fibrosis in rats. *Am J Physiol Cell Physiol*, Epub ahead of print, 2011.
- Litovkin KV, Ivanova OV, Bauer A et al. Microarray study of gene expression in uterine leiomyoma. *Exp Oncol*, 30, 106-111, 2008.

- Lloret M, Lara PC, Bordon E, Fontes F, Rey A, Pinar B, Falcon O. Major vault protein may affect nonhomologous end-joining repair and apoptosis through Ku70/80 and bax downregulation in cervical carcinoma tumors. *Int J Radiat Oncol Biol Phys*, 73, 976-979, 2009.
- Louhimo R and Hautaniemi S. CNAmets: an R package for integrating copy number, methylation and expression data. *Bioinformatics*, 27, 887-888, 2011.
- Luo J, Feng J, Lu J et al. Aberrant methylation profile of 14-3-3 sigma and its reduced transcription/expression levels in Chinese sporadic female breast carcinogenesis. *Med Oncol*, 7, 791-797, 2010.
- Luo X, Ding L, Xu J, et al. Leiomyoma and myometrial gene expression profiles and their responses to gonadotropin-releasing hormone analog therapy. *Endocrinology*, 146, 1074-1096, 2005a.
- Luo X, Ding L, Xu J et al. Gene expression profiling of leiomyoma and myometrial smooth muscle cells in response to transforming growth factor-beta. *Endocrinol*, 146, 1097-1118, 2005b.
- Luo X, Pan Q, Liu L et al. Genomic and proteomic profiling II: Comparative assessment of gene expression profiles in leiomyomas, keloids, and surgically-induced scars. *Reprod Biol Endocrinol*, 5, 1-12, 2007.
- Luque FA, Furneaux HM, Ferziger R et al. Anti-Ri: an antibody associated with paraneoplastic opsoclonus and breast cancer. *Ann Neurol*, 29, 241-251, 1991.
- MacConaill LE and Garraway LA. Clinical Implications of the Cancer Genome. *J Clin Oncol*, 28, 5219-5228, 2010.
- Mangino M, Brouillette S, Braund P et al. A regulatory SNP of the BICD1 gene contributes to telomere length variation in humans. *Hum Mol Genet*, 17, 2518-2523, 2008.
- Manohar M, Mooney AM, North JA et al. Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J Biol Chem*, 284, 23312-23321, 2009.
- Mantovani MS, Barbieri Neto J, Philbert PMP. Multiple uterine leiomyomas: Cytogenetic analysis. *Gynecol Oncol*, 72, 71-75, 1999.
- Mangioni S, Viganò P, Lattuada D et al. Overexpression of the Wnt5b gene in leiomyoma cells: implications for a role of the Wnt signaling pathway in the uterine benign tumor. *J Clin Endocrinol Metab*, 90, 5349-5355, 2005.
- Mao X, Barfoot R, Hamoudi RA et al. Allelotype of uterine leiomyomas. *Cancer Genet Cytogenet*, 114, 89-95, 1999.
- Markowski DN, von Ahsen I, Nezhad MH et al. HMGA2 and the p19Arf-TP53-CDKN1A axis: a delicate balance in the growth of uterine leiomyomas. *Genes Chrom Cancer*, 49, 661-668, 2010.
- Marsh A, Healey S, Lewis A. Mutation analysis of five candidate genes in familial breast cancer. *Breast Cancer Res Treat*, 105, 377-389, 2007.

- Martinez-Mir A, Glaser B, Chuang GS et al. Germline Fumarate Hydratase Mutations in Families with Multiple Cutaneous and Uterine Leiomyomata. *J Invest Dermatol*, 741-744, 2003.
- McCarroll SA, Kuruvilla FG, Korn JM et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet*, 40, 1166-1174, 2008.
- Meadows KL, Andrews DM, Xu Z et al. Genome-wide analysis of loss of heterozygosity and copy number amplification in uterine leiomyomas using the 100K single nucleotide polymorphism array. *Exp Mol Pathol*, Epub ahead of print, 2011.
- Melchor L, Saucedo-Cuevas LP, Muñoz-Repeto I et al. Comprehensive characterization of the DNA amplification at 13q34 in human breast cancer reveals TFDP1 and CUL4A as likely candidate target genes. *Breast Cancer Res*, 11, R86, 2009.
- Meloni AM, Surti U, Sandberg AA. Deletion of chromosome 13 in leiomyomas of the uterus. *Cancer Genet Cytogenet*, 53, 199-203, 1991.
- Mertens HJ, Heineman MJ, Theunissen PH et al. Androgen, estrogen and progesterone receptor expression in the human uterus during the menstrual cycle. *Eur J Obstet Gynecol Reprod Biol*, 98, 58-65, 2001.
- Mesquita FS, Dyer SN, Heinrich DA et al. Reactive oxygen species mediate mitogenic growth factor signaling pathways in human leiomyoma smooth muscle cells. *Biol Reprod*, 82, 341-351, 2010.
- Meza-Zepeda LA, Kresse SH, Barragan-Polania AH et al. Array comparative genomic hybridization reveals distinct DNA copy number differences between gastrointestinal stromal tumors and leiomyosarcomas. *Cancer Res*, 66, 8984-8993, 2006.
- Micci F, Teixeira MR, Haugom L et al. Genomic aberrations in carcinomas of the uterine corpus. *Genes Chrom Cancer*, 40, 229-246, 2004.
- Middlebrook BS, Eldin K, Li X et al. Smad1-Smad5 ovarian conditional knockout mice develop a disease profile similar to the juvenile form of human granulosa cell tumors. *Endocrinology*, 150, 5208-5217, 2009.
- Mills RE, Walter K, Stewart C et al. Mapping copy number variation by population-scale genome sequencing. *Nature*, 470, 59-65, 2011.
- Minelli L, Romagnolo C, Giambanco L et al. Uterine leiomyoma metastasis as a first sign of breast cancer. *J Am Assoc Gynecol Laparosc*, 5, 213-215, 1998.
- Misao R, Nakanishi Y, Fujimoto J et al. Expression of sex hormone-binding globulin mRNA in uterine leiomyoma, myometrium and endometrium of human subjects. *Gynecol Endocrinol*, 9, 317-323, 1995.
- Mitelman Database of Chromosome Aberrations and Gene Fusions in Cancer (2011). Mitelman F, Johansson B and Mertens F (Eds.), <http://cgap.nci.nih.gov/Chromosomes/Mitelman>.
- Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation *Nature Rev Cancer*, 7, 233-245, 2007.

- Mitra R, Lee J, Jo J et al. Prediction of Postoperative Recurrence-Free Survival in Non-Small Cell Lung Cancer by Using an Internationally Validated Gene Expression Model. *Clin Cancer Res*, 17, 2934-2946, 2011.
- Mittal KR, Chen F, Wei JJ et al. Molecular and immunohistochemical evidence for the origin of uterine leiomyosarcomas from associated leiomyoma and symplastic leiomyoma-like areas. *Mod Pathol*, 22, 1303-1311, 2009.
- Mittal KR and Joutovsky A. Areas with benign morphologic and immunohistochemical features are associated with some uterine leiomyosarcomas. *Gynecol Oncol*, 104, 362-365, 2007.
- Mittal KR, Popiolek D, Demopoulos RI. Uterine myxoid leiomyosarcoma within a leiomyoma. *Hum Pathol*, 31, 398-400, 2000.
- Mizuochi T, Kimura A, Ueki I et al. Molecular genetic and bile acid profiles in two Japanese patients with 3beta-hydroxy-DELTA5-C27-steroid dehydrogenase/isomerase deficiency. *Pediatr Res*, 68, 258-263, 2010.
- Mo R, Rao SM, Zhu YJ. Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *J Biol Chem*, 281, 15714-15720, 2006.
- Mok SC, Bonome T, Vathipadiekal V et al. A gene signature predictive for outcome in advanced ovarian cancer identifies a survival factor: microfibril-associated glycoprotein 2. *Cancer Cell*, 16, 521-532, 2009.
- Moore AB, Yu L, Swartz CD et al. Human uterine leiomyoma-derived fibroblasts stimulate uterine leiomyoma cell proliferation and collagen type I production, and activate RTKs and TGF beta receptor signaling in coculture. *Cell Commun Signal*, 10, 8-10, 2010.
- Moore SD, Herrick SR, Ince TA et al. Uterine leiomyomata with t(10;17) disrupt the histone acetyltransferase MORF. *Cancer Res*, 64, 5570-5577, 2004.
- Moorehead ME and Conard CJ. Uterine leiomyomas: a treatable condition. *Ann N Y Acad Sci*, 948, 21-29, 2001.
- Müller M, Beck IM, Gadesmann J et al. MMP19 is upregulated during melanoma progression and increases invasion of melanoma cells. *Mod Pathol*, 23, 511-521, 2010.
- Murnane JP. Telomere loss as a mechanism for chromosome instability in human cancer. *Cancer Res*, 70, 4255-4259, 2010.
- Mylonas I, Brüning A, Shabani N et al. Evidence of inhibin/activin subunit betaC and betaE synthesis in normal human endometrial tissue. *Reprod Biol Endocrinol*, 19, 143, 2010.
- Nabokikh A, Ilhan A, Bilban M et al. Reduced TGF-beta1 expression and its target genes in human insulinomas. *Exp Clin Endocrinol Diabetes*, 115, 674-682, 2007.
- Napoli JL. Biochemical pathways of retinoid transport, metabolism, and signal transduction. *Clin Immunol Immunopathol*, 80, 52-62, 1996.

- Narayan G, Scotto L, Neelakantan V et al. Protocadherin PCDH10, involved in tumor progression, is a frequent and early target of promoter hypermethylation in cervical cancer. *Genes Chromosomes Cancer*, 48, 983-992, 2009.
- Nezhad MH, Drieschner N, Helms S et al. 6p21 rearrangements in uterine leiomyomas targeting HMGA1. *Cancer Genet Cytogenet*, 203, 247-252, 2010.
- Nikitenko LL, Cross T, Campo L et al. Expression of terminally glycosylated calcitonin receptor-like receptor in uterine leiomyoma: endothelial phenotype and association with microvascular density. *Clin Cancer Res*, 12, 5648-5658, 2006.
- Nilbert M, Heim S, Mandahl N et al. Ring formation and structural rearrangements of chromosome 1 as secondary changes in uterine leiomyomas with t(12;14)(q14-15;q23-24). *Cancer Genet Cytogenet*, 36, 183-190, 1988.
- Nilbert M and Heim S. Uterine Leiomyoma Cytogenetics. *Genes Chrom Cancer*, 2, 3-13, 1990.
- Nilbert M, Heim S, Mandahl N et al. Characteristic chromosome abnormalities, including rearrangements of 6p, del(7q), +12, and t(12;14), in 44 uterine leiomyomas. *Hum Genet*, 85, 605-611, 1990.
- Nishiyama N, Arai E, Nagashio R et al. Copy number alterations in urothelial carcinomas: their clinicopathological significance and correlation with DNA methylation alterations. *Carcinogenesis*, 32, 462-469, 2011.
- Noe M, Kunz G, Herbertz M et al. The cyclic pattern of the immunocytochemical expression of oestrogen and progesterone receptors in human myometrial and endometrial layers: characterization of the endometrial-subendometrial unit. *Human Reproduction*, 14, 190-197, 1999.
- Nowak RA. Identification of new therapies for leiomyomas: What in vitro studies can tell us. *Clin Obstet Gynecol*, 44, 327-334, 2001.
- Noyes RW, Hertig AT, Rock J. Dating the endometrial biopsy. *Fertil Steril*, 1, 3-25, 1950.
- Numa F, Hirabayashi K, Kawasaki K et al. Syndecan-1 expression in cancer of the uterine cervix: association with lymph node metastasis. *Int J Oncol*, 20, 39-43, 2002.
- Olshen AB, Venkatraman ES, Lucito R et al. Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics*, 5, 557-572, 2004.
- O'Neill CJ, McBride HA, Connolly LE et al. Uterine leiomyosarcomas are characterized by high p16, p53 and MIB1 expression in comparison with usual leiomyomas, leiomyoma variants and smooth muscle tumours of uncertain malignant potential. *Histopathol*, 50, 851-858, 2007.
- Oster B, Thorsen K, Lamy P et al. Identification and validation of highly frequent CpG island hypermethylation in colorectal adenomas and carcinomas. *Int J Cancer*, Epub ahead of print, 2011.
- Ozisik YY, Meloni AM, Surti U et al. Inversion (X)(p22q13) in a uterine leiomyoma. *Cancer Genet Cytogenet*, 61, 131-133, 1992.

- Ozisik YY, Meloni AM, Surti U et al. Involvement of 10q22 in leiomyoma. *Cancer Genet Cytogenet*, 69, 132-135, 1993a.
- Ozisik YY, Meloni AM, Surti U et al. Deletion of 7q22 in uterine leiomyomas. A cytogenetic review. *Cancer Genet Cytogenet*, 71, 1-6, 1993b.
- Ozisik YY, Meloni AM, Stone JF et al. Spontaneous expression of the chromosome fragile site at 10q23 in leiomyoma. *Cancer Genet Cytogenet*, 74, 73-75, 1994.
- Packenhams JP, Du Manoir S, Schrock E et al. Analysis of genetic alterations in uterine leiomyomas and leiomyosarcomas by comparative genomic hybridization. *Mol Carcinog*, 19, 273-279, 1997.
- Paiva CE, Drigo SA, Rosa FE et al. Absence of transforming growth factor-beta type II receptor is associated with poorer prognosis in HER2-negative breast tumours. *Ann Oncol*, 21, 734-740, 2010.
- Pakiz M, Potocnik U, But I. Solitary and multiple uterine leiomyomas among Caucasian women: two different disorders? *Fertil Steril*, 94, 2291-2295.
- Pan CC and Epstein JI. Detection of chromosome copy number alterations in metanephric adenomas by array comparative genomic hybridization. *Mod Pathol*, 23, 1634-1640, 2010.
- Pan Q, Luo X, Chegini N. microRNA 21 (miR-21): Response to hormonal therapies and regulatory function in Leiomyoma, transformed Leiomyoma and Leiomyosarcoma cells. *Mol Hum Reprod*, 1-35, 2009.
- Pandis N, Heim S, Bardi G et al. Parallel karyotypic evolution and tumor progression in uterine leiomyoma. *Genes Chrom Cancer*, 2, 311-317, 1990.
- Park J, Chen L, Ratnashinge L et al: Deletion polymorphism of UDP-glucuronosyltransferase 2B17 and risk of prostate cancer in African American and Caucasian men. *Cancer Epidemiol Biomarkers Prev*, 15, 1473-1478, 2006.
- Parker WH. Etiology, symptomatology, and diagnosis of uterine myomas. *Fertil Steril*, 87, 725-36, 2007.
- Parker WH, Fu YS, Berek JS: Uterine sarcoma in patients operated on for presumed leiomyoma and rapidly growing leiomyoma. *Obstet Gynecol*, 83, 414-418, 1994.
- Peddada SD, Laughlin SK, Minera K et al. Growth of uterine leiomyomata among premenopausal black and white women. *PNAS*, 16, 105, 19887-19892, 2008.
- Pei-Chiang C, Takahashi A, Hosono N et al. A genome-wide association study identifies three loci associated with susceptibility to uterine fibroids. *Nat Genetics*, 43, 447-451, 2011.
- Peng Y, Laser J, Shi G et al. Antiproliferative effects by Let-7 repression of high-mobility group A2 in Uterine Leiomyoma. *Mol Cancer Res*, 6, 663-673, 2008.

- Penner AS, Rock MJ, Kielty CM et al. Microfibril-associated glycoprotein-2 interacts with fibrillin-1 and fibrillin-2 suggesting a role for MAGP-2 in elastic fiber assembly. *J Biol Chem*, 277, 35044-35049, 2002.
- Perez GD. Caracterização das deleções no cromossomo 7 em Leiomiomas Uterinos pela metodologia de hibridação *in situ* fluorescente. In: Dissertação de Mestrado, 81 pp, 2004.
- Perou CM, Sorlie T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature*, 406, 747-752, 2000.
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature*, 406, 747-752, 2000.
- Pinkel D, Seagraves R, Sudar D et al: High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. *Nat Genet*, 20, 207-211, 1998.
- Pinto D, Marshall C, Feuk L et al. Copy-number variation in control population cohorts. *Hum Mol Genet*, 16, 168-173, 2007.
- Pitt CW and Cooper JP. Pot1 inactivation leads to rampant telomere resection and loss in one cell cycle. *Nucleic Acids Res*, 38, 6968-6975, 2010.
- Pole JC, Courtay-Cahen C, Garcia MJ et al. High-resolution analysis of chromosome rearrangements on 8p in breast, colon and pancreatic cancer reveals a complex pattern of loss, gain and translocation. *Oncogene*, 25, 5693-5706, 2006.
- Polito P, Dal Cin P, Kazmierczak B et al. Deletion of HMG17 in uterine leiomyomas with ring chromosome 1. *Cancer Genet Cytogenet*, 108, 107-109, 1999.
- Pranzatelli, MR. The neurobiology of the opsoclonus-myoclonus syndrome. *Clin. Neuropharmacol*, 15, 186-228, 1992.
- Prenzel KL, Schäfer E, Stippel D et al. Multiple giant leiomyomas of the esophagus and stomach. *Dis Esophagus*, 19, 504-508, 2006.
- Przybytkowski E, Ferrario C, Basik M. The use of ultra-dense array CGH analysis for the discovery of micro-copy number alterations and gene fusions in the cancer genome. *BMC Med Genomics*, 27; 4-6, 2011.
- Ptacek T, Song C, Walker CL et al. Physical mapping of distinct 7q22 deletions in uterine leiomyoma and analysis of a recently annotated 7q22 candidate gene. *Cancer Genet Cytogenet*, 174, 116-120, 2007.
- Quade BJ, Wang TY, Sornberger K et al. Molecular pathogenesis of uterine smooth muscle tumors from transcriptional profiling. *Genes Chrom Cancer*, 40, 97-108, 2004.
- Quintana DG, Thome KC, Hou ZH et al. ORC5L, a new member of the human origin recognition complex, is deleted in uterine leiomyomas and malignant myeloid diseases. *J Biol Chem*, 273, 27137-27145, 1998.

- Redon R, Ishikawa S, Fitch KR et al. Global variation in copy number in the human genome. *Nature*, 444, 444-454, 2006.
- Reeves R. Structure and function of the HMGI(Y) family of architectural transcription factors. *Environ Health Perspect*, 108, 803-809, 2000.
- Rein MS, Barbieri RL, Friedman AJ. Progesterone: A critical role in the pathogenesis of uterine myomas. *Am J Obstet Gynecol*, 172, 14-18, 1995.
- Rein MS, Friedman AJ, Barbieri RL et al. Cytogenetic abnormalities in uterine leiomyomata. *Obstet Gynecol*, 77, 923-926, 1991.
- Rein MS, Powell WL, Walters FC. et al. Cytogenetic abnormalities in uterine myomas are associated with myoma size. *Mol Hum Reprod*, 4, 83-86, 1998.
- Rikala SM, El-Rifai W, Lahtinen T et al. Different patterns of DNA copy number changes in gastrointestinal stromal tumors, leiomyomas, and schwannomas. *Hum Pathol*, 29, 476-481, 1998.
- Rogatto, SR. *Citogenética sem risco: Biossegurança e garantia de qualidade*. Ed. FUNPEC, 170pp, 2000.
- Rooney JE, Gurpur PB, Burkin DJ. Laminin-111 protein therapy prevents muscle disease in the mdx mouse model for Duchenne muscular dystrophy. *Proc Natl Acad Sci U S A*, 106, 7991-7996, 2009.
- Roth TM, Klett C, Cowan BD. Expression profile of several genes in human myometrium and uterine leiomyoma. *Fertil Steril*, 87, 635-641, 2007.
- Rudolph KL, Chang S, Lee HW et al. Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell*. Mar, 96, 701-712, 1999.
- Ryan BM, Robles AI, Harris CC. Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer*, 10, 389-402, 2010.
- Sanchez-Garcia F, Akavia UD, Mozes E et al. Identification of Significant Targets in Cancer. *BMC Bioinfo*, 11, 189-198, 2010.
- Sargent MS, Weremowicz S, Rein MS et al. Translocations in 7q22 define a critical region in Uterine Leiomyomata. *Cancer Genet Cytogenet*, 77, 65-68, 1994.
- Sarkar S, Dey BK, Dutta A. MiR-322/424 and -503 are induced during muscle differentiation and promote cell cycle quiescence and differentiation by down-regulation of Cdc25A. *Mol Biol Cell*, 21, 2138-2149, 2010.
- Sayagués JM, Abad Mdel M, Melchor HB et al. Intratumoural cytogenetic heterogeneity of sporadic colorectal carcinomas suggests several pathways to liver metastasis. *J Pathol*, 221, 308-319, 2010.
- Schena M, Shalon D, Davis RW et al. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, 270, 467-470, 1995.

- Scherer SW, Cheung J, MacDonald JR et al. Human chromosome 7: DNA sequence and biology. *Science*, 300, 767-772, 2003.
- Schwartz SM. Epidemiology of uterine leiomyomata. *Clin Obst Gynecol*, 44, 316-326, 2001.
- Segal E, Shapira M, Regev A et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genetics*, 34, 166-176, 2003.
- Sell SM, Tullis C, Stracner D et al. Minimal interval defined on 7q in uterine leiomyoma. *Cancer Genet Cytogenetics*, 157, 67-69, 2005.
- Sendemir A, Sendemir E, Kosmehl H et al. Expression of sex hormone-binding globulin, oxytocin receptor, caveolin-1 and p21 in leiomyoma. *Gynecol Endocrinol*, 24, 105-112, 2008.
- Shaw-Smith C, Redon R, Rickman L et al. Microarray based comparative genomic hybridization (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *J Med Genet*, 41, 241-248, 2004.
- Shiang CY, Qi Y, Wang B et al. Amplification of fibroblast growth factor receptor-1 in breast cancer and the effects of brivanib alaninate. *Breast Cancer Res Treat*, 123, 747-755, 2010.
- Shozu M, Sumitani H, Segawa T et al. Overexpression of aromatase P450 in leiomyoma tissue is driven primarily through promoter I.4 of the aromatase P450 gene (CYP19). *J Clin Endocrinol Metab*, 87, 2540-2548, 2002.
- Shaik NA, Lone WG, Khan IA et al. Enhanced transcription of estrogen receptor α and mitochondrial cytochrome b genes in uterine leiomyomas. *Gynecol Endocrinol*, Epub ahead of print, 2011.
- Simon R. Challenges of microarray data and the evaluation of gene expression profile signatures. *Cancer Invest*, 26, 327-332, 2008.
- Singh RR, Cheung KJ, Horsman DE. Utility of Array Comparative Genomic Hybridization in Cytogenetic Analysis. *Methods Mol Biol*, 730, 219-234, 2011.
- Skubitz KM and Skubitz AP. Differential gene expression in uterine leiomyoma. *J Lab Clin Med*, 141, 297-308, 2003.
- Skubitz KM and Skubitz APN. Differential Gene Expression in Leiomyosarcoma. *Cancer*, 98, 5, 1029-1038, 2003.
- Sleeman MA, Fraser JK, Murison JG et al. B cell- and monocyte-activating chemokine (BMAC), a novel non-ELR alpha-chemokine. *Int Immunol*, 12, 677-689, 2000.
- Slovak ML, Bedell V, Hsu YH et al. Hodgkin and Reed-Sternberg cells at disease onset harbor distinct recurring DNA copy number alterations in chemosensitive versus primary refractory Hodgkin lymphoma. *Clin Cancer Res*, Epub ahead of print, 2011.
- Smith SW, Eardley KS, Croft AP et al. CD248(+) stromal cells are associated with progressive chronic kidney disease. *Kidney Int*, Epub ahead of print, 2011.

- Solyom S, Winqvist R, Nikkilä J et al. Screening for large genomic rearrangements in the FANCA gene reveals extensive deletion in a Finnish breast cancer family. *Cancer Lett*, 28, 113-118, 2011.
- Sorlie T, Tibshirani R, Parker J et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA*, 100, 8418-8423, 2003.
- Sornberger KS, Weremowicz S, Williams AJ et al. Expression of HMGIY in three uterine leiomyomata with complex rearrangements of chromosome 6. *Cancer Genet Cytogenet*, 114, 9-16, 1999.
- Sotiriou C, Neo SY, McShane LM et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci USA*, 100, 10393-10398, 2003.
- Sourla A, Polychronakos C, Zeng WR. Plasminogen activator inhibitor 1 messenger RNA expression and molecular evidence for del (7) (q22) in uterine leiomyomas. *Cancer Res*, 56, 3123-3128, 1996.
- Sozen I and Arici A. Interactions of cytokines, growth factors, and the extracellular matrix in the cellular biology of uterine leiomyomata. *Fertil Steril*, 78, 1-12, 2002.
- Spivey KA and Banyard J. A prognostic gene signature in advanced ovarian cancer reveals a microfibril-associated protein (MAGP2) as a promoter of tumor cell survival and angiogenesis. *Cell Adh Migr*, 4, 169-171, 2010.
- Steck PA, Pershouse MA, Jasser SA. Identification of a candidate tumour suppressor gene, *MMAC1*, at chromosome 10q23.3 that is mutated in multiple advanced cancers. *Nat Genet*, 15, 356-362, 1997.
- Steiner E, Holzmann K, Elbling L et al. Cellular functions of vaults and their involvement in multidrug resistance. *Curr Drug Targets*, 7, 923-934, 2006.
- Stewart EA and Nowak RA. New concepts in the treatment of uterine leiomyomas. *Obstet Gynecol*, 92, 24-27, 1998.
- Stewart EA, Friedman AJ, Peck K et al. Relative overexpression of collagen type I and collagen type III messenger ribonucleic acids by uterine leiomyomas during the proliferative phase of the menstrual cycle. *J Clin Endocrinol Metab*, 79, 900-906, 1994.
- Stewart L, Glenn GM, Stratton P et al. Association of germline mutations in the fumarate hydratase gene and uterine fibroids in women with hereditary leiomyomatosis and renal cell cancer. *Arch Dermatol*, 144, 1584-1592, 2008.
- Stovall DW. Clinical symptomatology of uterine leiomyomas. *Clin Obstet Gynecol*, 44, 364-371, 2001.
- Strissel PL, Swiatek J, Oppelt P et al. Transcriptional analysis of steroid hormone receptors in smooth muscle uterine leiomyoma tumors of postmenopausal patients. *J Steroid Biochem Mol Biol*, 107, 42-47, 2007.

- Su G, Burant CF, Beecher CW et al. Integrated metabolome and transcriptome analysis of the NCI60 dataset. *BMC Bioinformatics*, 12, S36, 2011.
- Su M, Chung HY, Li Y. 6-O-Angeloylenolin induced cell-cycle arrest and apoptosis in human nasopharyngeal cancer cells. *Chem Biol Interact*, 189, 167-176, 2011.
- Subramanian A, Tamayo P, Mootha VK et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS*, 102, 15545-15550, 2005.
- Swartz CD, Afshari CA, Yu L et al. Estrogen-induced changes in IGF-I, Myb family and MAP kinase pathway genes in human uterine leiomyoma and normal uterine smooth muscle cell lines. *Molecular Human Reproduction*, 11, 441-450, 2005.
- Syssoev KA, Kulagina NV, Chukhlovin AB et al. Expression of mRNA for Chemokines and Chemokine Receptors in Tissues of the Myometrium and Uterine Leiomyoma. *Bull Exp Biol Med*, 145, 84-89, 2008.
- Takahashi T, Nagai N, Oda H et al. Evidence for RAD51L1/HMGIC fusion in the pathogenesis of uterine leiomyoma. *Genes Chromosomes Cancer*, 30,196-201, 2001.
- Takahashi K, Kawamura N, Ishiko O et al. Shrinkage effect of gonadotropin releasing hormone agonist treatment on uterine leiomyomas and t(12;14). *Int J Oncol*, 20, 279-283, 2002.
- Tallini G, Vanni R, Manfioletti G et al. HMGI-C and HMGI(Y) immunoreactivity correlates with cytogenetic abnormalities in lipomas, pulmonary chondroid hamartomas, endometrial polyps, and uterine leiomyomas and is compatible with rearrangement of the HMGI-C and HMGI(Y) genes. *Lab Invest*, 80, 359-369, 2000.
- Taylor KJ, Sims AH, Liang L et al. Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res*, 12, R39, 2010.
- Teicher BA. Newer vascular targets: endosialin (review). *Int J Oncol*, 30, 305-312, 2007.
- Terry K, McGrath M, Lee IM et al. Genetic variation in CYP11A1 and StAR in relation to endometrial cancer risk. *Gynecol Oncol*, 117, 255-259, 2010.
- Testa JR, Zhou JY, Bell DW et al. Chromosomal localization of the genes encoding the kinetochore proteins CENPE and CENPF to human chromosomes 4q24-q25 and 1q32-q41, respectively, by fluorescence in situ hybridization. *Genomics*, 23, 691-693, 1994.
- Tsibris JCM, Segars J, Coppola D et al. Insights from gene arrays on the development and growth regulation of uterine leiomyomata. *Fertil Steril*, 78, 114-121, 2002.
- Tsuchida T, Okazawa H, Mori T et al. In vivo imaging of estrogen receptor concentration in the endometrium and myometrium using FES PET — influence of menstrual cycle and endogenous estrogen level. *Nuclear Medicine and Biology*, 34, 205-210, 2007.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *PNAS*, 98, 5116-5121, 2001.

- Ueki K, Ramaswamy S, Billings SJ et al. ANOVA, a putative astrocytic RNA-binding protein gene that maps to chromosome 19q13.3. *Neurogenetics*, 1, 31-36, 1997.
- Ule J, Ule A, Spencer J et al. Nova regulates brain-specific splicing to shape the synapse. *Nat Genet*, 37, 844-852, 2005.
- Ungar D, Oka T, Brittle EE et al. Characterization of a mammalian Golgi-localized protein complex, COG, that is required for normal Golgi morphology and function. *J Cell Biol*, 57, 405-415, 2002.
- Ushiro Y, Hashimoto O, Seki M et al. Analysis of the function of activin betaC subunit using recombinant protein. *J Reprod Dev*, 52, 487-495, 2006.
- Ustaalioglu BB, Bilici A, Seker M et al. Metastasis of lobular breast carcinoma to the uterus in a patient under anastrozole therapy. *Onkologie*, 32, 424-426, 2009.
- van Rijk A, Sweersa M, Huys E et al. Characterization of a recurrent t(1;2)(p36;p24) in human uterine leiomyoma. *Cancer Genet Cytogenet*, 193, 54-62, 2009.
- Vanharanta S, Wortham NC, Laiho P et al. 7q deletion mapping and expression profiling in uterine fibroids. *Oncogene*, 24, 6545-6554, 2005.
- Vanharanta S, Pollard PJ, Lehtonen HJ et al. Distinct expression profile in fumarate-hydratase-deficient uterine fibroids. *Hum Mol Genet*, 15, 97-103, 2006.
- Vanharanta S, Wortham NC, Langford C et al. Definition of a minimal region of deletion of chromosome 7 in uterine leiomyomas by tiling-path microCGH array and mutation analysis of known genes in this region. *Genes Chromosomes Cancer*, 46, 451-458, 2007.
- Velagaleti GVN, Tonkb VS, Hakim NM et al. Fusion of HMGA2 to COG5 in uterine leiomyoma. *Cancer Genet Cytogenet*, 202, 11-16, 2010.
- Vikhlyayeva EM, Khodzhaeva ZS, Fantschenko ND. Familial predisposition to uterine leiomyomas. *Int J Gynaecol Obstet*, 51, 127-131, 1995.
- Vincent-Salomon A, Lucchesi C, Gruel N et al. Integrated genomic and transcriptomic analysis of ductal in situ carcinoma in situ of the breast. *Clin Cancer Res*, 1956-1965, 2008.
- Viswanathan M, Hartmann K, McKoy N. et al. Management of uterine fibroids: an update of the evidence (Review). *Evid Rep Technol Assess*, 154, 1-122, 2007.
- Walker CL and Stewart EA. Uterine Fibroids: the elephant in the room. *Science*, 308, 1589-1592, 2005.
- Wang H, Mahadevappa M, Yamamoto K et al. Distinctive proliferative phase differences in gene expression in human myometrium and leiomyomata. *Fertil Steril*, 80, 266-276, 2003.
- Wang Y, Matsuo H, Kurachi O, Maruo T et al. Down-regulation of proliferation and up-regulation of apoptosis by gonadotropin-releasing hormone agonist in cultured uterine leiomyoma cells. *Eur J Endocrinol*, 146, 447-456, 2002.

- Watanabe Y and Maekawa M. Methylation of DNA in cancer. *Adv Clin Chem*, 52, 145-167, 2010.
- Waugh DJ and Wilson C. The interleukin-8 pathway in cancer. *Clin Cancer Res*, 14, 6735-6741, 2008.
- Wei LH, Torng PL, Hsiao SM et al. Histone Deacetylase 6 Regulates Estrogen Receptor {alpha} in Uterine Leiomyoma. *Reprod Sci*, Epub ahead of print, 2011.
- Weston G, Trajstman AC, Gargett CE et al. Fibroids display an anti-angiogenic gene expression profile when compared with adjacent myometrium. *Mol Hum Reprod*, 9, 541-549, 2003.
- Wheeler DA, Srinivasan M, Egholm M et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature*, 452, 872-876, 2008.
- WHO Consultation on obesity. Geneva: World Health Organ, 1997.
- Williams AJ, Powell WL, Collins, T et al. HMGI(Y) expression in human uterine leiomyomata. Involvement of another high-mobility group architectural factor in a benign neoplasm. *Am J Pathol*, 150, 911-918, 1997.
- Wolverton T and Lalande M. Identification and characterization of three members of a novel subclass of protocadherins. *Genomics*, 66-72, 2001.
- Wong KK, de Leeuw RJ, Dosanjh NS et al: A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet*, 80, 91-104, 2007.
- Wu CL, Classon M, Dyson N et al. Expression of dominant-negative mutant DP-1 blocks cell cycle progression in G1. *Mol Cell Biol*, 16, 3698-3706, 1996.
- Wu T, Chen X, Xie L. Selective estrogen receptor modulators (SERMs) for uterine leiomyomas. *Cochrane Database Syst Rev*, 17, CD005287, 2007.
- Wu X, Blanck A, Olovsson M et al. Expression of basic fibroblast growth factor (bFGF), FGF receptor 1 and FGF receptor 2 in uterine leiomyomas and myometrium during the menstrual cycle, after menopause and GnRHa treatment. *Acta Obstet Gynecol Scand*, 80, 497-504, 2001.
- Xiao HL, Eyden B, Yan XC et al. Intraparenchymal myofibromatosis of the brain in an adult: report of an unusual case. *Neuropathology*, 30, 288-293, 2010.
- Xiao J, Liu CC, Chen PL et al. RINT-1, a novel Rad50-interacting protein, participates in radiation-induced G(2)/M checkpoint control. *J Biol Chem*, 276, 6105-6111 2001.
- Xie Y, Yang D, He Q et al. Zebrafish as a model system to study the physiological function of telomeric protein TPP1. *PLoS One*, 6, e16440, 2011.
- Xing YP, Powell WL, Morton CC. The del(7) subgroup in Uterine Leiomyomata: genetic and biologic characteristics. *Cancer Genet Cytogenet*, 98, 69-74, 1997.
- Xu Q, Takekida S, Ohara N et al. Progesterone receptor modulator CDB-2914 down-regulates proliferative cell nuclear antigen and Bcl-2 protein expression and up-regulates caspase-3

and poly(adenosine 5'-diphosphate-ribose) polymerase expression in cultured human uterine leiomyoma cells. *J Clin Endocrinol Metab*, 90, 953-961, 2005.

Xu Y, Jing Y, Ma S et al. Primary angioleiomyoma in the sellar region: a case report and literature review. *Clin Neuropathol*, 29, 21-25, 2010.

Yen-Ping HJ, Man WC, Wen Y et al. Transforming growth interacting factor expression in leiomyoma compared with myometrium. *Fertil Steril*, 94, 1078-1083, 2010.

Ying J, Li H, Seng TJ et al. Functional epigenetics identifies a protocadherin PCDH10 as a candidate tumor suppressor for nasopharyngeal, esophageal and multiple other carcinomas with frequent methylation. *Oncogene*, 25, 1070-1080, 2006.

Yoshida S, Ohara N, Xu Q et al. Cell-type specific actions of progesterone receptor modulators in the regulation of uterine leiomyoma growth. *Semin Reprod Med*, 28, 260-273, 2010.

Young A. Receptor pharmacology. *Adv Pharmacol*, 52, 47-65, 2005.

Yu H and Rohan T. Role of the insulin-like growth factor family in cancer development and progression. *J Natl Cancer Inst*, 92, 1472-1489, 2000.

Yu W, Kanaan Y, Bae YK et al. Chromosomal changes in aggressive breast cancers with basal-like features. *Cancer Genet Cytogenet*, 193, 29-37, 2009.

Zaitseva M, Vollenhoven BJ, Rogers PAW. In vitro culture significantly alters gene expression profiles and reduces differences between myometrial and fibroid smooth muscle cells. *Mol Hum Reprod*, 12, 187-207, 2006.

Zaloudek CJ and Norris HJ. Mesenchymal tumors of the uterus. In: Fenoglio CM, Wolff M, eds. *Prog Surg Pathol* 3th ed. New York: Masson, 35pp, 1981.

Zasławski R, Surowiak P, Dziegiel P et al. Analysis of the expression of estrogen and progesterone receptors, and of PCNA and Ki67 proliferation antigens, in uterine myomata cells in relation to the phase of the menstrual cycle. *Med Sci Monit*, 7, 908-913, 2001.

Zavadil J, Ye H, Liu Z et al. Profiling and Functional Analyses of MicroRNAs and Their Target Gene Products in Human Uterine Leiomyomas. *PLoS ONE*, 5, e12362, 2010.

Zeng WR, Scherer SW, Koutsilieris M et al. Loss of heterozygosity and reduced expression of the CUTL1 gene in uterine leiomyomas. *Oncogene*, 14, 2355-2365, 1997.

Zhang XY, Shi L, Liu Y et al. ICPS: an integrative cancer profiler system. *Bioinformatics*, 26, 2649-2650, 2010.

Zheng HC, Wei ZL, Xu XY et al. Parafibromin expression is an independent prognostic factor for colorectal carcinomas. *Hum Pathol*, Epub ahead of print, 2011.

Zydan HLS, Fagundes DJ, Sória-Vieira S et al. Hysterectomy and benign gynecological diseases: what has been performed in Medical Residency in Brazil? *Rev Bras Ginecol Obstet*, 29, 67-73, 2007.

Anexo 1. Estudos prévios de expressão gênica em larga escala em Leiomiomas Uterinos.

Autores	Amostras (n) e fase do ciclo*	Resumo genes diferencialmente expressos	Validação
Tsibris JCM et al. Fertil Steril, 78, 114-121, 2002	LU (9) X MM (9) proliferativa e secretora	<i>DLK1, IGF2, IGFBP5, FZD2, RET, TGFB3, COL5A2, IGF1, IRS1</i> <i>ADH1A, THBS1, FOS, ALDH1, CYP21A2, IGFBP-6, PTGER3, MS4A2, PTGER2, IGFBP1, MAPK5, SERPINE1, TGFB2, RCAN1</i>	Western blot
Weston G et al. Mol Hum Reprod, 9, 541-549, 2003	LU (12) X MM (12) mulheres pré-menopáusicas	<i>IGF2, COL4A2 e EDNRA, CTGF, CYR61</i>	RT-PCR
Wang H et al. Fertil Steril, 80, 266-276, 2003	LU (7) X MM (7) proliferativa	<i>ADH1A, MAPKKK5, ALDH1, BCL6, LAC1, EDN1, GATA2, TGFB1, S100A13, RAB21</i> <i>HBR2, CRABP2, IGF2, ESR1.</i>	RT-PCR IHQ
Kanamori T et al. Mol Hum Reprod, 9, 709-717, 2003	LU (37) X MM (37) proliferativa	<i>PCP4, RGS19, LEFTY2, SFRP1, TNC, MMP11, VCAN</i> <i>COL6A3, IGF2, TGFB3, IGFBP5</i>	Northern blot, IHQ
Chegini N et al. J Soc Gynecol Investig, 10, 161-71, 2003	LU (6) X MM (6) secretora	LU x M não tratado = <i>IL15, SAP102, TRK-T3, SMAD2, HMG17 / VEGF, TRAP1, IL10, RAC2, TIMP1 e 3, MCL1, P48</i> LU x LU tratado = <i>FN1, LAMA4, CDH5, VIM, TIMP3, RAD51C.</i> M tratado x M não tratado = <i>CDH6, CCND1, HMG2 / COL6A3, TIMP1, CASP10, EGFRBP, GRB2, STAT5A e 5B, RAD23A</i>	Não validou
Catherino WH et al. Fertil Steril, 80, 2003	LU (5) X MM (5) proliferativa e secretora	<i>Sem diferença de expressão gênica DLK1, FZD2 e CD24</i> primeiro experimento com 12.000 sequências	RT-PCR, qRT-PCR, Western Blot, IHQ
Ahn WS et al. Int J Exp Path, 84, 267-279, 2003	LU (6) x MM (6)	<i>CDKN1A, COL4A6, IGF2, LAGY.</i> <i>ALDH1A2, LY96, BCL9, MAPK10, DUSP1, PDGFRL, APM2, PYGL, SERPINB1, FRZB, CTGF, RIG, ANXA1, ADH1C, DPT, ALDH1A1, ADH1B</i>	RT-PCR, espectrometria de massa, PDGE - 2D

Skubitz KM e Skubitz AP J Lab Clin Med, 141, 297-308, 2003	LMS (4) x LU (20) x MM (46) x outros (18)	LMS x outros tecidos - <i>CDKIN2A (p16)</i> mesmos genes em LU encontrado em LMS x MM - <i>DCX, CAPN6, IL7B, PLP1, ADH1A e 1B, IGF1, JUN, FOS, e TU3A.</i>	Não validou
Catherino WH et al. Genes Chromosomes Cancer, 40, 204- 217, 2004	LU (11) x MM (11)	<i>DPT, CALD1, TJP1, VEGF</i> <i>VCAN, P311, TGFB3</i>	RT-PCR, qRT- PCR, IHQ, Microscopia
Hoffman PJ et al. Fertil Steril, 83, 639-649, 2004	LU (5) x MM (5) proliferativa e secretora	<i>SFRP1 e 4, ESR1, IGF2, TGFB1, CREG, CYP1B1, CRABP2, SOX4</i> <i>MAP3K5, SERPINE1, TIMP3, AXL, ALDH1A1, DUSP1, TGFB2, ADAMTS1, VEGF, IGFBP6, CYR61, S100A4, ANXA1, EGR1</i>	qRT-PCR
Quade BJ et al. Genes Chrom Cancer, 40, 97- 108, 2004	MM (4) x LU (7) x LMS (2) proliferativa e secretora	LMxLMS - <i>CYP1A2, ESR1,</i> LMxMM - <i>CYP1A2, ESR1, TGFB3, CTGF, IGF,, ATF3, ADH1A, CYR61, GAS1, DUSP1, BTG3, NBL1, COL4A6, ATF3, GRIA2, IGF1?, JUN, TGFB3, e TGFB2.</i>	Não validou
Swartz CD et al. Mol Hum Reprod, 11, 441-450, 2005	Linhagens células de músculo liso e de LU	Linhagem LU - <i>Stromal cell derived factor-1, Acid phosphatase, IGF1, Pre B cell enhancing factor, heparin binding EGF-like growth factor, Myc</i> Linhagem céls músculo liso - <i>CD4 antigen, serine/threonine protein kinase SGK, tumour associated antigen 6</i> Ambas linhagens - <i>Nuclear protein N10, Fos, MAP kinase phosphatase 1</i>	Northern e Western blot, IHQ
Luo X et al. Endocrinology, 146, 1074-1096, 2005a	LU (6) x MM (6) mulheres pré- menopáusicas	LMxM (GnRHa) - <i>ADAM8, ERF, E2F1, IL4, NGFbeta, MAPKK6, FGFR3, ERH, SMAD5, STAT5B</i> LM GnRHa-tratado X não tratado - <i>EGF, FGF2</i> MM GnRHa- tratado X não tratado - <i>DUSP8, TPM1, GHRH, COL18A1, IGF1, ESRRG, JUN, TGFB1, EGR2 e 3, MAFF, CRY1, SERPINE1,</i>	qRT-PCR, Western blot, IHQ
Luo X et al. Endocrinology, 146, 1097-1118, 2005b	LU (6) x MM (6) mulheres pré- menopáusicas	Genes diferencialmente expressos em resposta a tratamento com TGF-beta: <i>IL11, TGFB1, EGR3, CITED2, CREB1, NR4A1, RUNX1, RUNX2, PSMD9, CDKN1C, GAS1 e GRK5</i>	qRT-PCR

Lee EJ et al. Int J Gynecol Cancer, 15, 146-154, 2005	LU (10) x MM (10) maioria proliferativa	<i>IGF2, NPTX2, DCX, ADH1B, CYR61, DPT, DKK3, EGR1</i>	RT-PCR, IHQ
Mangioni S et al. J Clin Endocrinol Metab, 90, 5349-5355, 2005	LU (22) x MM (22) mulheres pré-menopáusicas	Genes sinalização Wnt e comparação entre outros estudos	qRT-PCR, RT-PCR, Western blot
Arslan AA et al. Hum Reprod, 20, p852-863, 2005	LU (3) x MM (3) proliferativa	Identificadood em outros estudos - <i>ADH1, ATF3, CRABP2, CYR61, DUSP1, FOS, GAGEC1, IGFBP6, PTGDS, PTGER3, TYMS, ABCA, ANXA1, APM2, CCL21, CDKN1A, CRMP1, EMP1, ESR1, FY, MAP3K5, TGFBR2, TIMP3</i> Alterações em genes que regulam a síntese de esteróides, metabolismo de IGF, sinalização TGF- β e formação de matrix extracelular LM x MM - <i>IGF2, MMP11, CD24, PTK7, PEMT, ADH1, sinalização TGFB, IGFBP6, CYR61/IGFBP10, CTGF, S100A4</i>	Não validou
Zaitseva M et al. Mol Hum Reprod, 12, 187-207, 2006	LU (6) x MM (6) secretora	LM x MM - <i>IGF2, PTK7, KLF10, COL1A2, COL5A2, COL3A1, CYR61, DUSP1</i> , Em cultura - <i>ADAM9, ADAM10, RARB, TGFB3, ESR1, PgR</i>	qRT-PCR
Behera MA et al. Reprod Sci, 14, 63-76, 2007	LU (10) x MM (10)	IHC - <i>THBS1</i> (10 casos) e <i>THBS2</i> (3 casos) RT-PCR e Microarray – correlação positiva entre <i>MMP2</i> e <i>THSB2</i> em 5 pacientes <i>COL1A1</i> e <i>COL3A1</i>	qRT-PCR IHQ
Roth TM et al. Fertil Steril, 87, 635-641, 2007	LU de mulheres negras (4) x MM (4) proliferativa	<i>PTN, COL3A1, IGFBP2, IGF2R, AGT, RARB, TGFB2, PPAR</i>	qRT-PCR RT-PCR
Luo X et al. Reprod Biol Endocrinol, 5, 1-12, 2007	LU (3) x quelóides (4) e adesões peritoniais (3)	<i>ESR1</i> expressão aumentada em LU x tecidos com quelóides, mas não validou. <i>NR2F1, PNN, SMAD4, SMAD5, STAT5B, JUN, TGIF2, e ATF1, RUNX3, STAT1, STAT6, EGR3, GAS7, SMAD1 e EDF1</i>	qRT-PCR
Catherino WH e Malik M. Fertil Steril, 87, 1388-1398, 2007	Cultura de células de LU e MM	<i>ADH1A, ALDH1A1, CRABP1, RARA, RARB, RXRA</i> <i>CYP26A1</i>	RT-PCR IHQ

Dimitrova IK et al. Fertil Steril, 91, 2650-2663, 2009	LU (3) x MM (3) uma paciente pós-menopáusia	<i>FBLN1, DSG2, PRKCB, MMP11 e MMP14, FBLN2, FBLN3, EFEMP1, FBLN5, MST4, MMP7</i>	RT-PCR Western blot IHQ
Hodge JC et al. Am J Pathol, 172, 1403-1410, 2008	LU plexiformes (6) Análise cariotípica prévia FISH para o <i>HMG2</i>	<i>UBC, RDBP, CYC1, COX5B, CKS1B, DSS1, TMSB10, PIPPIN, ENPP1, GRIA2, CRMP1, TGFB3, NBL1, LAMA3, MORF4L2, KANK, CTNND1, SETDB1, TACR2, GSTM4, APOD, TIMP3, DPP6, DVL3, VCL, TPM4, GTF2I, THRSP, ID2, STAT4 e MMP2</i>	qRT-PCR IHQ
Crabtree JS et al. Cancer Res, 69, 6171-6178, 2009	LU (15) X MM (16) de ratos LU (23) X MM (23) humanos	Genes da via mTOR; <i>GRIN2B, CAPN6, MMP11, e MMP13</i>	Western blot, IHQ
Borsari R et al. Fertil Steril, 94, 1072-1077, 2010	30 pacientes nulíparas com LU Dois grupos - tratadas x não tratadas	<i>CYR61, SULF 2, HMG1, and EGR1 e WIF 1</i>	qRT-PCR
Jiang Y et al. Syst Biol Reprod Med, 56, 318-326, 2010	LU (42) x MM (42)	<i>CCN1, CCN2, AXL, PTK6, NRP1, VEGFA, JAK3, AVEN, BAG1, FGF9, CCNE2, LYN, EPHA7, FLT1, MET, FES, FGF23, IGFBP5, IGF2, EPHA8, PDGFC, EPHB4, EFNA4, BOK, EFNB2, HBEGF, BCL2L1, PGF, BCL2L14, CCND3, MDK, BCL2L12, APAF1, BCL2L11, PRAD1, FGFR1, EFNB3, PTK7,</i>	qRT-PCR

*Alguns estudos não informaram a fase do ciclo das pacientes. LU = leiomiomas uterinos; LMS = leiomiosarcomas; MM = miométrio adjacente; IHQ = imunohistoquímica; RT-PCR = PCR qualitativa; qRT-PCR = PCR quantitativa em tempo real; Verde = expressão diminuída; Vermelho = expressão aumentada.

Anexo 2. Dados clínicos e histopatológicos das pacientes com Leiomiomas Uterinos utilizadas neste estudo.

Caso	Amostras*	Idade	Número de tumores	Endométrio	Raça	Idade Menarca	IMC	Idade 1ª gestação	Gestações	História familiar ca	História pessoal ca
1	046T_A 046T_B	60	LUM	outros	B	18	42,2	ND	9	irmã gêmea LU e 2 irmãs ca endométrio	ca pele
2	084T	48	LUM	outros	B	13	28,3	23	3	tio materno ca hepático/tia paterna ca endometrial/irmã ca mama/ irmã LU	nega
3	085T	37	LUM	proliferativo	B	12	27,5	26	2	nega	nevus intradérmico/fibroma mole
4	091T	48	LUM	proliferativo	B	14	29,9	ND	0	nega	nega
5	092T	47	LU	proliferativo	B	10	23,2	23	4	nega	nega
6	147T_A 147T_B 147T_C	41	LUM	proliferativo	B	13	26,0	ND	3	ca mama (irmã)	nega
7	300T_A 300T_B	36	LUM	secretor	B	13	26,2	ND	3	nega	nega
8	301T_A 301T_B 301T_C	49	LUM	proliferativo	NB	11	36,1	15	7	pai ca próstata	nega
9	303T	47	LUM	secretor	B	13	29,3	23	1	pai ca?	nega
10	304T	44	LUM	proliferativo	B	13	26,6	ND	10	avó ca intestino	nega
11	307T_A 307T_B	48	LUM	proliferativo	B	12	34,9	26	4	nega	nega
12	317T_A	38	LUM	secretor	NB	12	24,8	20	5	nega	nega

	317T_B										
	317T_C										
13	614T	44	LU	proliferativo	B	12	29,9	17	10	nega	nega
14	615T	43	LUM	proliferativo	B	ND	22,9	0	0	nega	nega
15	625T	47	LUM	secretor	B	12	24,5	24	3	tia ca mama/tia ca estômago	nega
16	629T_A 629T_B	49	LUM	secretor	B	15	21,2	26	3	irmã LU	nega
17	630T_A 630T_B	47	LUM	proliferativo	B	13	26,3	18	2	nega	nega
18	631T_A 631T_B 631T_C	43	LUM	proliferativo	NB	9	24,1	ND	0	nega	nega
19	633T	46	LU	outros	NB	ND	26,2	ND	3	mãe LU	nega
20	634T	44	LU	secretor	B	15	23,6	23	3	ND	ND
21	642T_A 642T_B	46	LUM	secretor	B	9	29,9	21	1	mãe ca útero	nega
22	643T_A 643T_B 643T_C	39	LUM	outros	B	14	25,4	31	1	irmão leucemia	nega
23	644T	39	LU	proliferativo	B	12	24,3	ND	2	nega	nega
24	670T	38	LU	secretor	B	15	26,6	18	3	nega	nega
25	681T	46	LUM	proliferativo	B	15	21,5	21	4	tia materna ca útero	pólipo endometrial
26	683T_A 683T_B	49	LUM	secretor	B	13	28,7	23	3	irmã LU	nega
27	684T	49	LU	secretor	B	13	25,6	ND	6	nega	nega

28	708T	45	LU	secretor	B	13	33,9	22	3	tia paterna ca mama	nega
29	711T	43	LUM	outros	B	12	30,0	18	2	nega	exérese cisto no pescoço
30	719T	50	LU	secretor	B	13	24,2	0	0	pai óbito ca metastático?	nega
31	722T_A 722T_B 722T_C	50	LUM	secretor	B	11	34,9	ND	4	mãe ca rim	nega
32	728T	50	LUM	outros	B	16	27,3	18	6	nega	nega
33	729T_A 729T_B 729T_C	46	LUM	secretor	B	12	24,3	25	1	nega	nega
34	733T_A 733T_B	48	LUM	outros	B	11	25,8	25	2	nega	nega
35	759T	51	LUM	secretor	B	15	21,2	ND	3	nega	nega
36	761T	45	LU	secretor	B	13	18,7	ND	3	ca intestino (mãe)	nega
37	769T	47	LU	proliferativo	B	14	23,1	ND	2	tio ca?	nega
38	790T	52	LUM	outros	B	12	32,0	26	3	irmã ca mama	nega
39	842T	37	LUM	outros	B	11	29,5	20	1	mãe e irmã LU/avó, avô e tio ca?	nega
40	844T_A 844T_B	50	LUM	proliferativo	B	12	36,8	ND	4	nega	LU
41	845T_A 845T_B	54	LUM	outros	NB	14	22,4	28	6	pai ca intestino	nega
42	853T	35	LUM	secretor	B	11	26,8	21	2	nega	nega
43	854T	47	LUM	outros	B	12	34,4	33	1	pai ca próstata	nega
44	855T	37	LU	secretor	B	13	24,9	23	2	tio ca estômago	nega
45	857T	41	LUM	secretor	B	13	23,7	20	3	3 irmãs LU	nega

46	867T	46	LUM	outros	B	14	33,5	17	4	irmã ca intestino/sobrinha ca ósseo	nega
47	901T	51	LUM	proliferativo	B	9	34,1	ND	3	nega	nega
48	942T	48	LUM	outros	B	13	27,5	25	3	nega	nega
49	947T	47	LUM	proliferativo	B	ND	27,6	ND	ND	nega	nega
50	948T	46	LUM	proliferativo	B	13	26,6	29	2	tia paterna ca mama	nega
51	954T	45	LUM	secretor	B	12	33,7	18	3	nega	nega
52	1001T	46	LUM	outros	B	12	32,9	ND	10	avô ca boca	nega
53	1002T	46	LUM	secretor	B	13	27,5	18	2	nega	nega
54	1003T	36	LUM	outros	NB	13	21,6	0	0	avó materna ca estômago	nega
55	1005T	40	LUM	proliferativo	NB	12	34,2	ND	2	mãe ca cólon	nega
56	1015T	43	LU	secretor	NB	13	24,2	ND	2	nega	nega

Em negrito, amostras selecionadas para expressão gênica em larga escala. *Número de amostras avaliadas por caso; Outros=endométrio basal ou atrófico; ND=não definido; LUM=leiomiomas uterinos múltiplos; LU=leiomioma uterino único; B=branca; NB=não branca; IMC=índice de massa corpórea; ca=câncer; ?=dúvida na resposta.

Anexo 3. Regiões envolvidas em alto nível de ganhos e perdas homozigotas distribuídas entre as amostras de Leiomiomas Uterinos.

Amostras	Região	Posição	Evento	Tamanho	Símbolo do gene	miRNAs
642T_A	1p32.3	53,522,717-53,822,562	ANG	299846	<i>DMRTB1, FLJ40434, GLIS1, LRP8</i>	
759T	1p35.3	28,177,545-29,102,257	PH	924713	>10	
1005T	1p36.11	24,343,549-24,628,588	ANG	285040	<i>C1orf201, GRHL3, IL28RA, LOC284632, NIPAL3</i>	
1015T	1p36.12	20,825,832-21,667,911	ANG	842080	<i>DDOST, ECE1, EIF4G3, HP1BP3, KIF17, NBPF3, PINK1, SH2D5</i>	hsa-mir-1256
1015T	1p36.33	1,126,488-1,312,474	ANG	185987	>10	
733T_B		764,603-1,464,685	ANG	700083	>10	hsa-mir-429, hsa-mir-200b, hsa-mir-200a
147T_A	1q23.3	160,035,864-160,908,351	ANG	872488	<i>ATF6, C1orf111, C1orf226, DDR2, MIR556, NOS1AP, OLFML2B, SH2D1B, UAP1, UHMK1</i>	hsa-mir-556
634T	1q43	239,547,628-239,873,763	PH	326136	<i>CHML, FH, KMO, OPN3, RGS7</i>	
1015T	2q14.3	127,761,044-128,533,454	ANG	772411	>10	
615T, 769T	2q31.2	179,112,462-179,414,753	ANG	302292	<i>CCDC141, MIR548N, TTN</i>	
629T_B	2q33.1	198,610,861-199,908,639	ANG	1297779	<i>PLCL1, SATB2</i>	
733T_B	2q37.3	240,662,537-241,467,459	ANG	804923	>10	hsa-mir-149
790T		240,688,770-241,180,357	ANG	491588	<i>ANKMY1, CAPN10, DUSP28, GPC1, MIR149, MYEOV2, OTOS, PP14571, RNPEPL1</i>	hsa-mir-149
1015T	3q23	143,131,301-143,792,940	ANG	661640	<i>ATR, GK5, TFDP2, XRN1</i>	
901T	3q29	195,830,910-196,432,478	ANG	601569	<i>C3orf21, FAM43A, LSG1, TMEM44</i>	
1015T	4p16.3	221,876-1,773,213	ANG	1551338	>10	hsa-mir-571
1015T	4q13.3	73,330,498-74,347,744	ANG	1017247	<i>ADAMTS3, ANKRD17, COX18</i>	
630T_B	5p15.1	16,949,171-17,555,490	ANG	606320	<i>BASP1, LOC285696, MYO10</i>	
1001T	6q13	73,302,931-75,123,875	ANG	1820945	>10	

1015T	6q13-q14.1	75,806,978-76,314,627	ANG	507650	<i>COL12A1, COX7A2, FILIP1, TMEM30A</i>	
1015T	6q14.3	86,030,676-86,398,242	ANG	367567	<i>NT5E, SNX14, SYNCRIP</i>	
1015T	6q21	111,712,978-112,697,554	ANG	984577	<i>C6orf225, FYN, LAMA4, REV3L, TRAF3IP2, TUBE1, WISP3</i>	
769T	6q23.2	132,994,431-133,778,091	ANG	783661	>10	
1015T	7p22.1	6,410,776-6,924,151	ANG	513376	>10	
769T	7q22.1	100,323,922-103,058,551	PH	2734630	>10	hsa-mir-548o
759T	8p23.2	2,895,708-5,484,889	ANG	2589182	<i>CSMD1</i>	
1015T	8q11.21	48,621,770-49,021,474	ANG	399705	<i>CEBPD, KIAA0146, PRKDC</i>	
1015T	8q24.12	120,574,919-120,854,617	ANG	279699	<i>ENPP2, TAF2</i>	
901T	9p12-p11.2	42,014,069-42,702,421	ANG	688353	<i>ANKRD20A2, ANKRD20A3, FAM95B1</i>	
670T, 1001T 722T_B	9p13.3	33,871,385-34,349,039	ANG	477655	<i>DCAF12, KIF24, NUDT2, SNORD121A, SNORD121B, UBAP1, UBAP2, UBE2R2</i>	
1015T	9p22.3-p22.2	16,353,348-16,917,523	ANG	564176	<i>BNC2</i>	
1015T	9p23-p22.3	13,045,258-14,821,430	ANG	1776173	<i>CER1, FREM1, MPDZ, NFIB, ZDHHC21</i>	
625T	9q34.3	138,470,455-138,845,556	ANG	375102	>10	hsa-mir-126
1015T	10p11.22	32,304,680-32,823,769	ANG	519090	<i>CCDC7, EPC1, KIF5B</i>	
630T_A	10p15.3	0-1,064,364	ANG	1064365	<i>C10orf108, C10orf110, DIP2C, GTPBP4, IDI2, LARP4B, TUBB8, ZMYND11</i>	
1015T	10q21.2-q21.3	64,585,376-64,870,931	ANG	285556	<i>JMJD1C, MIR1296</i>	hsa-mir-1296
1015T	10q22.3	78,167,059-79,452,226	ANG	1285168	<i>DLG5, KCNMA1, LOC100128292, POLR3A</i>	
1015T	10q26.12-q26.13	123,015,655-123,682,321	ANG	666667	<i>ATE1, FGFR2</i>	
301T_C	10q26.3	134,828,307-134,960,478	ANG	132172	<i>ADAM8, KNDC1, MIR202, TUBGCP2, UTF1, VENTX</i>	hsa-mir-202
317T_A		134,834,654-135,048,859	ANG	214206	>10	hsa-mir-202
733T_B		134,341,466-135,035,217	ANG	693752	>10	hsa-mir-202
708T	11p12	41,208,731-43,395,213	ANG	2186483	<i>API5, TTC17</i>	

1005T	11p15.5	774,043-838,551	ANG	64509	<i>CD151, CEND1, EFCAB4A, LRDD, PNPLA2, POLR2L, RPLP2, SLC25A22, SNORA52, TSPAN4</i>	
1015T		767,539-838,551	ANG	71013	<i>CD151, CEND1, EFCAB4A, LRDD, PNPLA2, POLR2L, RPLP2, SLC25A22, SNORA52, TSPAN4</i>	
642T_A		396,940-551,034	ANG	154095	<i>ANO9, C11orf35, HRAS, LRRC56, PTDSS2, RASSF7, RNH1, SIGIRR</i>	
733T_B		209,012-2,143,485	ANG	1934474	>10	hsa-mir-483, hsa-mir-675, hsa-mir-210
1015T	11q13.1	65,775,949-65,880,765	ANG	104817	<i>B3GNT1, BRMS1, CD248, CNIH2, KLC2, RAB1B, RIN1, TMEM151A, YIF1A</i>	
954T	11q13.2	67,570,565-68,221,718	ANG	651154	<i>C11orf24, CHKA, GAL, LRP5, SAPS3, SUV420H1, TCIRG1</i>	
644T, 670T	11q14.1	82,275,057-83,407,676	ANG	1132620	<i>ANKRD42, C11orf82, CCDC90B, DLG2, PCF11, PRCP, RAB30, SNORA70E</i>	
681T		82,275,057-83,407,676	ANG	1132620	<i>ANKRD42, C11orf82, CCDC90B, DLG2, PCF11, PRCP, RAB30, SNORA70E</i>	
1015T	12p11.22-p11.21	30,581,792-31,656,430	ANG	1074639	<i>CAPRIN2, DDX11, DENND5B, FAM60A, FLJ13224, IPO8, TSPAN11</i>	
1001T	12p13.31	7,530,089-9,042,700	ANG	1512612	>10	
733T_B	12q13.3	55,799,155-55,951,830	ANG	152676	<i>LRP1, MIR1228, NDUFA4L2, NXPH4, R3HDM2, SHMT2, STAC3</i>	hsa-mir-1228
842T		55,805,133-55,923,685	ANG	118553	<i>LRP1, MIR1228, NDUFA4L2, NXPH4, SHMT2, STAC3</i>	hsa-mir-1228
300T_A	12q24.31	119,131,582-120,835,736	ANG	1704155	>10	
1015T	12q24.33	131,666,252-132,065,341	ANG	399090	<i>ANKLE2, CHFR, FBRSL1, GOLGA3, P2RX2, PGAM5, POLE, PXMP2, ZNF605</i>	
1015T	13q14.11	39,854,438-40,523,146	ANG	668709	<i>ELF1, FOXO1, LOC646982, MIR320D1, MIR621,</i>	hsa-mir-621, hsa-mir-

					<i>MRPS31, SLC25A15, SUGT1L1</i>	320d-1
630T_A	13q21.31-q21.32	61,500,686-65,823,692	ANG	4323007	<i>OR7E156P, PCDH9</i>	
1015T	13q32.1	95,190,823-95,510,926	ANG	320104	<i>DNAJC3, UGGT2</i>	
630T_A	14q21.3-q22.1	46,294,392-48,524,187	ANG	2229796	<i>MDGA2</i>	
1015T	14q22.1	49,863,385-50,450,539	ANG	587155	<i>ABHD12B, AT1L1, CDKL1, MAP4K5, NIN, PYGL, SAV1</i>	
954T	14q23.1	57,617,329-58,746,271	ANG	1128943	<i>ACTR10, ARID4A, C14orf37, DAAM1, DACT1, FLJ31306, KIAA0586, PSMA3, TIMM9, TOMM20L</i>	
683T_B	15q26.1	89,659,483-90,416,121	ANG	756639	<i>SLCO3A1</i>	
1015T	16q22.1	65,652,632-65,820,436	ANG	167805	>10	hsa-mir-328
630T_A	16q22.3	72,033,508-72,894,835	ANG	861328	<i>PSMD7</i>	
631T_B	16q24.3	88,060,983-88,283,196	ANG	222214	<i>ANKRD11, C16orf55, CDK10, CHMP1A, CPNE7, DPEP1, RPL13, SNORD68, SPG7</i>	
1015T	17q12	32,026,668-32,784,291	ANG	757624	<i>AATF, ACACA, DHRS11, LHX1, MRM1</i>	
1015T	17q23.3-q24.1	59,409,954-60,286,002	ANG	876049	>10	
1015T	17q25.3	74,300,696-74,491,524	ANG	190829	<i>LGALS3BP, TIMP2, USP36</i>	
733T_B		76,691,820-77,619,424	ANG	927605	>10	hsa-mir-1250, hsa-mir-338, hsa-mir-657
630T_A	18q12.2	33,417,167-35,366,305	ANG	1949139	<i>LOC647946</i>	
733T_B	19q13.42	60,234,352-60,585,968	ANG	351617	>10	
629T_B	21p11.1	10,013,263-10,117,957	PH	104695	<i>BAGE, BAGE2, BAGE3, BAGE4, BAGE5</i>	
630T_A, 1003T		10,013,263-10,117,957	ANG	104695	<i>BAGE, BAGE2, BAGE3, BAGE4, BAGE5</i>	
733T_B	21q22.11	33,097,167-33,677,532	ANG	580366	<i>C21orf54, C21orf62, IFNAR1, IFNAR2, IL10RB, OLIG1, OLIG2</i>	
728T	21q22.12	35,234,989-36,070,842	ANG	835854	<i>C21orf96, MIR802, RUNX1</i>	hsa-mir-802
733T_B	22q13.33	48,616,848-49,691,432	ANG	1074585	>10	

ANG = alto nível de ganhos; PH = perda homozigota; >10 = mais que 10 genes descritos na região.

Anexo 4. Genes e miRNAs mapeados nas regiões significativas de ganhos e perdas genômicos nas amostras de Leiomiomas Uterinos.

Região	Start	End	Evento	Símbolo do gene	miRNAs
1p36.33-p36.32	1,921,978	3,562,051	Perda	>10 (<i>TP73, ARHGEF16</i>)	hsa-mir-551a
1p31.1	82,163,689	84,655,179	Perda	<i>DNASE2B, LPHN2, PRKACB, SAMD13, TTLL7, UOX</i>	
1q21.3	149,836,310	150,010,690	Ganho	<i>SNX27, TNRC4, RIIAD1, CELF3, MRPL9, OAZ3</i>	
1q41	212,893,807	214,014,200	Perda	<i>CENPF, KCNK2, KCTD3, MIR548F3, USH2A</i>	
2q35	217,253,192	219,351,693	Ganho	>10 (<i>IGFBP5, CXCR1, CXCR2</i>)	hsa-mir-26b
4p14	40,185,743	40,457,450	Perda	<i>RBM47, NSUN7</i>	
4q13.1-q13.2	60,129,993	68,031,967	Perda	<i>CENPC1, EPHA5, LPHN3, TECRL</i>	hsa-mir-1269
4q21.21	80,461,995	82,110,583	Perda	<i>GDEP, PRDM8, ANTXR2, FGF5, ARD1B, NAA11, GK2</i>	
4q28.3	132,220,774	135,650,424	Perda	<i>PABPC4L, PCDH10</i>	
5q31.1-q31.2	134,274,907	135,585,827	Perda	>10 (<i>TGFBI, CXCL14</i>)	hsa-mir-886
6p21.32	31,962,130	32,132,073	Ganho	>10	hsa-mir-1236
6q13	73,782,170	74,568,779	Ganho	>10	
6q16.3	102,041,584	103,869,362	Ganho	<i>GRIK2</i>	
7p12.2	50,138,493	50,702,500	Perda	<i>ZNFN1A1, FIGNL1, IKZF1, DDC, GRB10</i>	
7q22.1	98,769,492	102,068,500	Ganho	>10	hsa-mir-25, hsa-mir-93, hsa-mir-106b, hsa-mir- 548o
7q22.1-q22.2	104,284,386	105,139,642	Perda	<i>ATXN7L1, EFCAB10, LHFPL3, LOC100216545, LOC723809, MLL5, PUS7, RINT1, SRPK2</i>	
7q31.33	124,268,700	125,951,548	Perda	<i>GRM8, POT1</i>	
8p12-p11.23	37,606,006	38,685,823	Ganho	>10 (<i>STAR, FGFR1</i>)	

8q22.3-q23.1	105,831,966	107,742,335	Ganho	<i>OXR1, ZFPM2</i>	
8q24.3	143,989,891	146,027,963	Ganho	>10 (<i>CYP11B2, C8orf51, PYCRL, FAM83H, CYC1, BOP1, NFKBIL2, MFSD3, RECQL4, LRRC14</i>)	hsa-mir-937, hsa-mir-661, hsa-mir-939, hsa-mir-1234
9p13.3	33,871,385	34,349,039	Ganho	<i>DCAF12, KIF24, NUDT2, SNORD121A, SNORD121B, UBAP1, UBAP2, UBE2R2</i>	
9q34.3	137,569,888	138,507,449	Perda	>10	
10p15.3	0	1,064,364	Ganho	<i>C10orf108, C10orf110, DIP2C, GTPBP4, IDI2, LARP4B, TUBB8, ZMYND11</i>	
10q21.3	68,975,747	70,562,567	Ganho	>10	hsa-mir-1254
11p15.5	0	1,034,144	Ganho	>10 (<i>HRAS, IFTM1</i>)	hsa-mir-210
11p14.1	29,114,213	29,511,311	Perda	0	0
11q12.3-q13.2	62,133,411	67,979,861	Ganho	>10 (<i>ESRRA, CD248, PPP2R5B</i>)	hsa-mir-1237, hsa-mir-192, hsa-mir-194-2
12q13.2-q13.3	54,386,743	56,279,267	Ganho	>10 (<i>APOF, CDK2, STAT2, STAT6, MYL6, MYL6B, MYO1A, HSD17B6, ITGA7, MMP19</i>)	hsa-mir-1228, hsa-mir-616
12q24.11-q24.12	108,897,746	110,392,895	Ganho	>10	
12q24.31	119,131,582	119,853,673	Ganho	>10	
13q31.1	79,882,431	85,782,478	Perda	<i>MIR548F1, SLITRK1, SLITRK6</i>	
14q32.2	96,180,194	99,038,571	Perda	<i>BCL11B, C14orf177, C14orf64, CCNK, SETD3, VRK1</i>	
15q25.3	85,542,221	86,595,151	Perda	<i>NCRNA00052, NTRK3</i>	
16p13.3	2,756,969	3,081,334	Ganho	>10	
16p11.2	28,184,420	31,429,424	Ganho	>10	
16q22.1	65,732,841	66,551,504	Ganho	>10	hsa-mir-328
16q24.3	88,134,943	88,526,617	Ganho	>10	

17p13.1	7,431,136	7,505,810	Ganho	<i>MPDU1, SOX15, FXR2, SHBG, SAT2, ATP1B2</i>	
17q21.31	38,105,718	39,165,518	Ganho	>10	
18q22.1-q22.2	60,070,073	65,021,204	Perda	<i>CCDC102B, CDH19, CDH7, DSEL, TMX3</i>	
19p13.12-p13.11	15,989,861	16,541,429	Ganho	>10	
19q13.32-q13.33	50,012,054	55,753,381	Ganho	>10	hsa-mir-330, hsa-mir-642, hsa-mir-769, hsa-mir-220c, hsa-mir-150
20q11.23-q12	37,049,227	38,416,306	Perda	>10	
20q13.33	61,881,839	62,198,449	Ganho	>10	hsa-mir-941-1, hsa-mir- 941-2, hsa-mir-941-3, hsa- mir-1914, hsa-mir-647
22q13.33	49,041,686	49,691,432	Ganho	>10	

>10 = mais que 10 genes descritos na região. Os genes representados entre parênteses foram selecionados como genes candidatos e apresentados na discussão dos achados.

Anexo 5. Genes identificados na análise integrada. Os genes estão separados de acordo com o padrão de ganhos e perdas genômicas e expressão aumentada ou diminuída entre as amostras.

CGH array	Microarray	Símbolo Gene	Nº amostras expressão diminuída	Nº amostras expressão aumentada	Nº amostras perdas genômicas	Nº amostras ganhos genômicos
+	+	<i>HSPB7</i>	1	50	12	39
-	-	<i>CENPF</i>	51	0	31	20
+	+	<i>CALCRL</i>	0	51	25	26
+	+	<i>COL3A1</i>	1	50	44	7
+	+	<i>IGFBP5</i>	1	50	20	31
+	+	<i>TNS1</i>	0	51	36	15
+	+	<i>CTDSP1</i>	0	51	37	14
+	-	<i>VIL1</i>	51	0	23	28
+	-	<i>RQCD1</i>	47	4	18	33
-	-	<i>RHOH</i>	51	0	19	32
+	-	<i>KIF20A</i>	51	0	31	20
+	-	<i>CDC25C</i>	51	0	33	18
+	-	<i>PRELID1</i>	51	0	21	30
+	-	<i>F12</i>	50	1	37	14
+	+	<i>DBN1</i>	1	50	23	28
+	-	<i>CPSF4</i>	51	0	34	17
+	-	<i>ATP5J2</i>	47	4	30	21
+	-	<i>ZNF655</i>	50	1	33	18
+	-	<i>AZGP1</i>	50	1	24	27
+	-	<i>MCM7</i>	49	2	27	24
+	-	<i>STAG3</i>	50	1	39	12
+	-	<i>C8orf51</i>	51	0	33	18
+	-	<i>PYCRL</i>	50	1	34	17
+	-	<i>CYC1</i>	49	2	17	34
+	+	<i>VPS28</i>	0	51	39	12
+	-	<i>TONSL</i>	48	3	24	27
+	-	<i>RECQL4</i>	51	0	13	38
+	-	<i>LRRC14</i>	50	1	20	31
+	+	<i>GPR124</i>	1	50	27	24
+	-	<i>EIF4EBP1</i>	47	4	29	22

+	+	<i>FGFR1</i>	1	50	20	31
+	-	<i>IDI1</i>	49	2	27	24
+	+	<i>DIP2C</i>	1	50	14	37
+	-	<i>MYPN</i>	50	1	17	34
+	-	<i>DDX21</i>	49	2	31	20
+	-	<i>HKDC1</i>	50	1	37	14
+	+	<i>IFITM1</i>	0	51	22	29
+	-	<i>PKP3</i>	50	1	34	17
+	-	<i>HRAS</i>	51	0	23	28
+	-	<i>RASSF7</i>	50	1	28	23
+	-	<i>TALDO1</i>	50	1	51	0
+	-	<i>SLC25A22</i>	49	2	17	34
+	-	<i>CHKA</i>	50	1	11	40
-	-	<i>BICD1</i>	49	2	38	13
+	-	<i>AICDA</i>	50	1	42	9
+	+	<i>MFAP5</i>	1	50	29	22
+	-	<i>BAZ1A</i>	51	0	28	23
+	+	<i>NUPR1</i>	0	51	22	29
+	-	<i>CD19</i>	49	2	23	28
+	+	<i>PRRT2</i>	0	51	33	18
+	+	<i>MVP</i>	0	51	33	18
+	-	<i>GDPD3</i>	50	1	23	28
+	-	<i>CORO1A</i>	50	1	30	21
+	+	<i>ZNF688</i>	0	51	26	25
+	+	<i>EXOC3L1</i>	1	50	35	16
+	-	<i>ELMO3</i>	50	1	16	35
+	+	<i>C16orf7</i>	0	51	25	26
+	-	<i>FANCA</i>	51	0	16	35
+	-	<i>TUBB3</i>	50	1	34	17
+	-	<i>BRCA1</i>	51	0	22	29
+	+	<i>AOC3</i>	1	50	23	28
+	+	<i>VASP</i>	1	50	8	43
+	+	<i>GPR4</i>	0	51	27	24
+	-	<i>SNRPD2</i>	50	1	29	22
+	+	<i>DMPK</i>	1	50	29	22
+	+	<i>NOVA2</i>	0	51	19	32

+	+	<i>CCDC8</i>	0	51	37	14
+	-	<i>CALM3</i>	51	0	24	27
+	+	<i>DACT3</i>	0	51	31	20
+	-	<i>SLC1A5</i>	51	0	13	38
+	-	<i>AP2S1</i>	50	1	35	16
+	-	<i>SPIB</i>	51	0	20	31
-	+	<i>CD93</i>	0	51	25	26
+	-	<i>TRIB3</i>	50	1	36	15
+	-	<i>TBC1D20</i>	51	0	22	29

*Os sinais positivos (+) e negativos (-) representam o *status* dos genes em relação ao padrão de ganhos (+) e perdas (-) e ao aumento (+) e diminuição (-) de expressão gênica. N°=número de amostras classificadas de acordo com o padrão de ganhos e perdas genômicas e expressão aumentada ou diminuída.

An integrative genomic and transcriptomic analysis reveals potential targets associated with cell proliferation in Uterine Leiomyomas

Priscila Daniele Ramos Cirilo, M.Sc.^{a,b}, Fábio Albuquerque Marchi^{b,c}, Maria Aparecida Custódio Domingues, M.D., Ph.D.^d, Anaglória Pontes, M.D., Ph.D.^e, Silvia Regina Rogatto, Ph.D.^{b,f*}

^aDepartment of Genetics, Institute of Biosciences, UNESP - Sao Paulo State University, Botucatu, SP, Brazil.

^bCIPE - NeoGene Laboratory, AC Camargo Hospital, Fundação Antonio Prudente, São Paulo, SP, Brazil.

^cInstitute of Mathematics and Statistics, USP – São Paulo University, São Paulo, SP, Brazil.

^dDepartment of Pathology, School of Medicine, UNESP - Sao Paulo State University, Botucatu, SP, Brazil.

^eDepartment of Gynecology and Obstetrics, School of Medicine, UNESP - Sao Paulo State University, Botucatu, SP, Brazil.

^fDepartment of Urology, School of Medicine, UNESP - Sao Paulo State University, Botucatu, SP, Brazil.

***Correspondence to:**

Silvia Regina Rogatto, PhD

NeoGene Laboratory - AC Camargo Hospital, Fundação Antonio Prudente.

Rua Taguá, 440, Liberdade

São Paulo, SP - Brazil

CEP: 01508-010

Telephone: 55-11-21895152 Fax: 55-11-21895152

E-mail: rogatto@fmb.unesp.br and silvia.rogatto@gmail.com

Key-words: Uterine leiomyomas, array CGH, gene expression microarrays, integrative analysis, molecular biomarkers, biological pathways

Submitted: PLoS ONE (Impact Factor 4.351).

ABSTRACT

Background: Uterine leiomyomas (UL) are the most common benign tumors affecting between 25-30% of women in reproductive age. Although little is known about its etiology, these tumors represent a major problem in public health being the main indication for hysterectomy. About 40-50% have nonrandom cytogenetic abnormalities, thus half of these tumors may have submicroscopic alterations, including copy number alterations (CNAs). Global gene expression studies have shown genes that act in proliferation and cell cycle process, retinoic acid, TGF-beta and IGF-1 signaling in response to estrogen and progesterone. However, few genes mapped at CNVs regions were directly associated with the UL development. In this study, we have used an integrative genomic and transcriptomic profiling to elucidate mechanisms and candidate genes associated with ULs.

Methodology/Principal Findings: CGH array and gene expression microarrays were applied in 51 ULs obtained from 34 patients. JISTIC was used to classify genes mapped at gains and losses genomic regions. Differentially expressed genes were identified by SAM statistical test. The integrative analysis, using CONEXIC algorithm, revealed 75 modulators. *In silico* functional analysis (Ingenuity Pathways Analysis) of top 30 modulators (*P* value <0.01 and highest scores) revealed that most of them were associated with cell cycle and cell proliferation. *CALCRL*, *COL3A1*, *FGFR1*, *HSPB7*, *GPR4*, *IGFBP5*, *RHOH*, and *TNSI* genes were significantly associated with canonical pathways as drug targets. In addition, *COL3A1*, *FGFR1* and *IGFBP5* genes could be potential therapeutic targets.

Conclusions/Significance: The findings of this study allowed to describe new CNAs regions, and also confirmed the involvement of others CNVs already described in ULs. The strategy of integrative genomic and transcriptomic data analysis revealed potential molecular markers that could be used on treatment of ULs, a common and poorly understood disease.

INTRODUCTION

Uterine Leiomyomas (ULs) are the most common benign tumors of reproductive age affecting 25-30% of women [1]. The ULs are smooth muscle tumors, often multiple tumors are found in the same uterus [2]. Although they are extremely common and represent an important public health problem, the biology of these tumors still remains unexplained. The majority of tumors are asymptomatic which only one fourth of individuals have clinical symptoms such as pelvic pain, abnormal bleeding, infertility and pregnancy complications [3]. Estrogen and progesterone are the most critical regulators of fibroid growth [4, 5]. In addition, growth factors [6], deregulation of microRNAs (miRNAs) [7], shortening of telomeres [8], excessive production of disorganized extracellular matrix [6, 9], genomic regions with loss of heterozygosity [10] and recurrent chromosomal aberrations [for review 11] have also been suggested to contribute to the growth of fibroids.

Around 50% of ULs harbor chromosomal rearrangements involving a small number of nonrandom chromosomal regions [for review 12]. Gene expression profile studies revealed mainly genes involved with cell proliferation, cell cycle, differentiation and extracellular matrix production [for review 13]. However, few studies have associated copy number alterations (CNAs) with deregulated gene expression in ULs [14, 15, 16]; these studies having used indirect correlation analysis.

Identifying genes as drivers (modulators) of tumorigenesis is a crucial challenge. DNA copy number alteration is the one of several events that can regulate gene expression [17]. Recently, studies using integrative analysis of genomic and transcriptomic data of cancer have revealed genes mapped at CNAs regions with an aberrant gene expression [18, 19]. These events could be underlying mechanisms of disease evolution and reveal new potential candidates for therapeutic intervention [20].

Thus, these findings prompt us to investigate the genomic and transcriptomic profiling in ULs using CONEXIC, an algorithm recently described [17]. In addition, functional analysis of networks and canonical pathways of modulators was applied to uncover molecular pathways involving in ULs pathogenesis that can be useful to select molecular markers and define target therapies.

MATERIAL AND METHODS

Tissue sample collection. Fifty-one fresh ULs samples were collected from 34 patients and stored at -80°C . These samples were collected from patients that undergone to hysterectomy procedure at the Department of Gynecology and Obstetrics, School of Medicine, Sao Paulo State University, São Paulo, Brazil, between October 1995 and February 2004. All the patients were advised of the procedures and thus provided informed consent. This study was approved by the Institutional Ethics Committee 146/2007-CEP.

Clinical and histopathological parameters. Nine patients (n=9) had one UL tumor. Twenty-five patients had multiple tumors: 13 of them had one sample evaluated (n=13), seven had two samples evaluated (n=14) and five had three samples evaluated (n=15). All women were premenopausal, showing regular menstrual cycles prior to diagnosis of UL, and they no received exogenous hormones or hormone suppression therapy at least three months before surgery. At surgery, 13 patients were in proliferative menstrual cycle phase and 21 in secretory period [21]. The medical records from patients were examined in 2011 to retrieve clinical and pathological data (Table S1). The age ranged from 35 to 51 years with mean age 45 years. All of the tumors were histopathologically diagnosed as usual ULs.

Array CGH. Genomic DNA from 51 ULs were prepared by standard sodium dodecyl sulfate/proteinase K digestion followed by phenol and chloroform extraction and ethanol precipitation and were stored at -20°C . The samples were treated with RNaseA $20\mu\text{g/ml}$

(Sigma-Aldrich, St. Louis, MO). High-quality genomic DNA (500ng) from each sample were hybridized on Agilent Human 4x44K CGH Microarrays (Agilent Technologies, Santa Clara, CA) using a standard direct method as described in the Agilent Oligonucleotide Array-Based CGH for Genomic DNA Analysis Enzymatic Labeling for Blood, Cells or Tissues (www.chem.agilent.com) using a male commercial genomic DNA (Promega, USA) as reference. After slides scanning (Agilent scanner at 5 μm resolution), the array data was extracted using the default CGH settings of the Feature Extraction v.10.1.1.1 (Agilent Technologies, Santa Clara, CA).

Identification of significant CNAs by using JISTIC. Copy number analysis was performed using segmented genomic dataset (DNAcopy, <http://www.bioconductor.org/packages/2.3/bioc/html/DNAcopy.html>) and JISTIC algorithm [22]. JISTIC gives G_{DEL} and G_{AMP} scores for each altered region (loss and gains, respectively), multiplied by the increase average in the ratio in \log_2 amplified sample. Scoring (G score) is based on the amplitude mean (a) of the aberration type (G_{DEL} or G_{AMP}) and frequency (f) on the data set. The meaning of each score is determined by comparison with similar results obtained after permutation within each sample. All regions with a q-value below a threshold (0.25) are deemed significant. For large aberrations, the sub-region with a minimal q-value is identified as a peak driver region.

Transcriptional profiling. Total RNA was extracted from frozen tissues using the RNeasy Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's instructions. Microarray experiments were performed using Two-Color Human GE 4x44K Microarrays. Isolated RNA (500-1000ng) was converted to cDNA with reverse transcriptase and an oligo(dT) primer bearing a T7 promotor followed by *in vitro* transcription with T7 RNA polymerase to create amplified antisense RNA. RNA reference [23] was amplified as above. Microarray image analysis was done using Feature Extraction v.10.1.1.1 (Agilent

Technologies, Santa Clara, CA). Statistical analysis was performed using background-corrected mean signal intensities from each dye channel. Microarray data were normalized using intensity-dependent global normalization (LOWESS) using Feature Extraction v.10.1.1.1 software (Agilent Technologies, Santa Clara, CA).

Significant genes based on expression analysis. The raw data from array scans were normalized by median-centering genes for each array, and \log_2 transformed. Additionally, genes with 30 or more ‘absent’ scores were filtered out. A total of 16354 sequences were analyzed. In order to identify significant genes, the Significance Analysis of Microarrays (SAM) method was applied [24]. The False Discovery Ratio (FDR) <0.05 was used to determine the significance threshold for genes and to limit the likelihood of type I error taking into account the fact that thousands of genes are simultaneously being tested [25, 26]. To select significant genes, it was used threshold values of $\log_2\text{ratio} \geq 1.0$ and ≤ -1.0 fold-change to classify up- or down-regulated genes, respectively. Hierarchical clustering analysis was performed using Complete Linkage Hierarchical algorithm with Pearson correlation (TMeV v.4.5).

Integrative Analysis. CONEXIC (COpy Number and EXpression In Cancer) was used to integrate genomic and transcriptomic data, which is an algorithm that matched CNAs and gene expression data from tumor samples to identify gene drivers and the processes that they influence [17]. Higher score values (top 30 modulators) indicate a greater possibility of the gene have some adaptive advantage on the tumor phenotype.

miRNA target prediction. The mechanisms of inverse association (genomic gains/up-regulations or genomic losses/down-regulations) identified among modulators could be explained by miRNA regulation. The miRNA target prediction analysis was done using the *in silico* tools TargetScan (<http://www.targetscan.org/>) and PicTar (<http://pictar.mdc-berlin.de/cgi-bin/PicTar Vertebrate.cgi>).

Gene Set Enrichment Analysis. Gene Set Enrichment Analysis (GSEA) [28] was performed for all selected modulators which sharing common biological function, chromosomal location, or regulation (modules).

Functional Analysis. Ingenuity Pathways Analysis (IPA) 8.0 (Ingenuity® Systems, <http://www.ingenuity.com>) was used in the set of modulators. Networks were generated based on Ingenuity Pathways Knowledge Database. Fischer's exact test was applied to identify the significant functions, networks and pathways represented within the respective gene sets. As result, it is displayed a score $-\log(\text{P-value})$ representing the probability of finding genes in networks and pathways relative to other genes, assembled them into specific network/pathways based on random chance.

RESULTS

DNA copy number alterations analysis. According JISTIC analysis, the overall pattern of CNAs displayed heterogeneity among all cases showing 142 chromosomal regions involved in gains (1032 genes) and 18 in losses (160 genes) totalizing 1192 genes identified. The most frequent genomic imbalances were gains on chromosomes 16 and 19, and losses on chromosomes 4 and 16 (Table S2).

Gene expression analysis. Unsupervised hierarchical clustering analysis was unable to identify subgroups of tumors, according to the gene expression profile and clinical features (data not shown). Gene expression analysis revealed 3325 significant genes; 1138 were up-regulated and 2187 down-regulated. About 40% (39.7%) of up-regulated and 38% of down-regulated genes were associated with cellular processes. Significant genes were preferentially associated with cell cycle (<http://www.funnet.info/>) (Figure S1).

Integrative analysis. The integrative analysis by CONEXIC using 1192 up-regulated genes and 3325 down-regulated genes with significant G scores, ranked 75 modulators

mapped at 1p36.13, 1q41, 2q32.1, 2q32.2, 2q35, 4p14, 5q31.2, 5q35.3, 7q22.1, 8p12, 8q24.3, 10p15.3, 10q21.3, 11p15.5, 11q13.2, 12p11.21, 12p13.31, 14q13.2, 16p11.2, 16q24.3, 17q21.31, 19q13.32, 19q13.33, 20p11.2 and 20p13 (Table 1). The modulators had scores ranged from the 0 (*TBC1D20* - TBC1 domain family, member 20) to 21227.32 (*KIF20A* - kinesin family member 20A).

miRNA target prediction - Among the 75 modulators, 26 showed positive association (genomic gains/up-regulated) and three negative association (genomic losses/down-regulated). An inverse correlation was found involving 45 genes and one ORF sequence (*C8orf51*); 30 of them could be explained by miRNA regulation (Table 1) (Table S3). The miRNA target prediction analysis was unable to predict the inverse correlation found in 15 genes.

Modulators characterization. Unsupervised hierarchical clustering analysis was done in order to verify association between genomic and transcriptomic data of 75 modulators considering the number of tumors evaluated, menstrual cycle phase and diagnosis of multiple or solitary tumors. Both genomic and transcriptomic profile were not statistically associated with the clustering of samples according to these clinical data (Figure 1A). In addition, the distribution of the 75 modulators along the chromosomes revealed that 44 of them (58.66%) were mapped in regions not usually described as breakpoints targets (1p36.13, 1q41, 2q32.1, 2q32.2, 2q35, 4p14, 5q31.2, 7q22.1, 11q13.2, 12p13.31, 14q13.2, 16q22.1, 17q21.31, 19q13.32, 19q13.33), and 23 and 5 genes (30.67% and 6.67%, respectively) were mapped in telomeric or pericentromeric regions, which are frequently targets of chromosomal instability (Table 1) (Figure 1B).

Gene Set Enrichment Analysis. The results revealed that 32 out of 75 genes were enriched between 26 modules (Table S4). Among these modules, *CORO1A*, *FGFR1*, *DDX21* and *DBN1* genes were the most frequently identified whereas *DBN1* and *FGFR1* showed

positive association. Values of enrichment for the 26 modules were significantly associated with the 32 genes. Module 8 had the highest number of genes in overlapping (9 genes) including *CORO1A*, *MCM7*, *SLC1A5*, *DDX21*, *NUPR1*, *DBN1*, *EIF4EBP1*, *VIL1*, and *CHKA* ($P < 10^{-3}$) (Table 2).

Selection of central modulators. The top 30 modulators were selected based on higher CONEXIC scores (Table 1). Positive association was found involving 12 genes (*TNS1*, *HSPB7*, *DBN1*, *CALCRL*, *COL3A1*, *IGFBP5*, *MFAP5*, *NUPR1*, *GPR4*, *DIP2C*, *CTDSP1*, and *FGFR1*), while negative association was detected for *RHOH* and *CENPF*. Genes with positive association were mapped at 1p36.13, 2q32.1, 2q32.2, 2q35, 5q35.3, 8p12, 10p15.3, 12p13.31, 16p11.2, and 19q13.32 while genes showing negative association were located at 1q41 and 4p14 genomic regions (Figure 1C). Moreover, the GSEA values for these genes revealed that *DBN1* and *FGFR1* were in all eight modules of cancer genes, while *DIP2C* and *MFAP5* genes had only one associated module. The *TNS1*, *HSPB7*, *IGFBP5*, *GPR4* and *CTDSP1* were not observed in any module (Table 2).

Functional analysis of modulators. The top 30 modulators were subjected to functional *in silico* analysis using the Ingenuity Pathway Analysis (IPA). The networks generated from the input yields networks based on the known functions and interconnectivity of the affected genes. The modulators were present at eight gene interactions networks with scores values ranging from 2 to 38 (Table S5). Selected gene networks analysis showed *CALCRL*, *CENPF*, *COL3A1*, *FGFR*, *IGFBP5*, *GPR4*, *NUPR1*, *RHOH*, and *TNS1* genes with the remaining pathway molecules incorporated by IPA on Network 1, associated with cellular movement, skeletal and muscular system development and function, cell morphology (Figure 2A). Similarly, Network 2 (*HSPB7*, *DBN1* - developmental disorder, neurological disease, DNA replication, recombination, and repair) (Figure 2B) showed gene interactions between selected genes and IPA database. In addition, cancer, reproductive system disease, and

genetic disorder showed the higher number of associated molecules ($P < 10^{-4}$). Cell cycle, cellular assembly and organization, as well as cellular growth and proliferation were the major functions associated with dataset ($P < 10^{-3}$) (Figure 3B).

Sixty-one canonical pathways were found to be significantly associated with modulators, even though only 37 pathways were associated with the selected genes (Figure 3A). Eight out of 14 genes were associated with different canonical pathways including the intrinsic prothrombin activation pathway (*COL3A1*), FGF signaling (*FGFR1*), ERK/MAPK signaling (*HSPB7*), VDR/RXR activation (*IGFBP5*), mTOR signaling (*RHOH*), FAK signaling (*TNSI*), and G-Protein coupled receptor signaling (*CALCRL*, *GPR4*) (Figure 4A). The canonical pathway hepatic fibrosis/hepatic stellate cell activation revealed the involvement of *COL3A1*, *FGFR1*, and *IGFBP5* genes ($P < 0.01$), showing similarity to biological pathways already described in ULs (Figure 3C).

DISCUSSION

Based on *in silico* approaches integrating genomic and transcriptomic profiles, the present study has revealed candidate genes and molecular pathways associated with ULs pathogenesis.

Array CGH and gene expression microarray data of 51 ULs samples from 34 patients revealed 1192 genomic imbalances and 3325 up- or down-regulated transcripts, respectively. The integrative analysis using CONEXIC [17] identified 75 modulators that explain the 3325 differentially expressed genes. The modulators were distributed on 25 chromosomal regions, seven of them were previously described by HR-CGH or array CGH studies in ULs, including 1p36.13, 4p14, 7q22.1, 11p15.5, 14q13.2, 19q13.32 and 19q13.33 [30, 31, 15, 11, 32]. Furthermore, it was also verified new genomic imbalances involved in ULs mapped on 1q41, 2q32.1, 2q32.2, 2q35, 5q31.2, 5q35.3, 8p12, 8q24.3, 10p15.3, 10q21.3, 11q13.2,

12p11.21, 12p13.31, 16p11.2, 16q24.3, 17q21.31, 20p11.2 and 20p13. The 75 modulators play an important role in ULs pathogenesis regardless of tumor multiplicity and menstrual cycle phase.

The integrative analysis allowed the identification of 75 modulators based on gene expression data and then generates the association of these findings with genomic imbalances. Forty-five out of 75 modulators have shown an inverse association between genomic alterations (gains or losses) and gene expression pattern (up- or down-regulation). Several mechanisms, besides CNAs, are associated with gene expression regulation. Epigenetic mechanisms and miRNAs, for example, can regulate transcriptional events. Gene expression regulation by miRNAs has been a major focus in studies of tumors, including ULs [33, 34]. In the present study, 30 out of 45 genes that showed an inverse association between genomic/transcriptomic data could be explained by miRNA regulation. The additional 15 genes could be regulated by other mechanisms. In addition, 44 out of 75 modulators are mapped on chromosomal regions that generally are not frequent targets of breakpoints. Therefore, this study suggests that deregulated expression genes in ULs frequently are associated with genomic alterations and regulation by miRNAs. Altogether, CNAs and miRNA could explain the main mechanisms of regulation of gene expression of ULs samples.

Based on Akavia et al. [17], it was selected 30 modulators showing the highest scores values with 14 of them showing positive association (*CALCRL*, *COL3A1*, *CTDSP1*, *DBN1*, *DIP2C*, *FGFR1*, *GPR4*, *HSPB7*, *IGFBP5*, *MFAP5*, *NUPR1*, and *TNS1*) or negative association (*CENPF* and *RHOH*). The *TNS1*, *HSPB7*, *IGFBP5*, *GPR4* and *CTDSP1* modulators were not detected in modules, indicating that they could be classified into a new class of genes specifically associated with ULs. These selected modulators genes could be putative candidates to uterine leiomyomas development and are discussed below.

The modulators were mainly associated with cancer involving cell cycle ($P < 10^{-4}$) and cellular growth and proliferation functions ($P < 10^{-3}$). Damage of cell cycle has been described in the pathogenesis of several tumors [35, 36], including ULs [37, 38, 39]. In addition, cellular proliferation stimulated by growth factors and/or steroid hormones is one of the mechanisms accountable for volume increase observed in ULs tumors [5].

Menorrhagia, characterized by excessive uterine bleeding, is one of the most frequently symptoms associated with ULs and may have implications for fertility and contraception. The canonical pathway intrinsic mechanism for prothrombin activation begins with trauma to the blood vessel or exposure of blood to collagen in a traumatized vessel wall. Prothrombin overactivation could be associated with excessive bleeding observed in affected patients [40]. In the present study, the gene associated with this pathway was *COL3A1*, which showed positive association. The up-regulation of this gene has been previously associated with increase of collagen deposition in ULs [41]. The *in silico* functional analysis revealed that *COL3A1* molecule was associated with response to collagenase *Clostridium histolyticum*, an treatment recently approved for progressive Dupuytren contractures disease (DD) [42]. DD is a fibroproliferative disorder of unknown etiology that often results in shortening and thickening of the palmar fascia, leading to permanent and irreversible flexion contracture of the digits [43]. Therefore, *COL3A1* is as candidate gene for further studies aiming to evaluate its correlation with menorrhagia and fibroid formation in UL patients.

Fibroblast growth factor receptor (*FGFR*), showing positive association in the present study, belongs to the family of receptor tyrosine kinases (RTKs). Activated RTKs play an important role in the enhanced proliferation described in ULs [for review 44]. In addition, up-regulation of *FGF1* was associated with menorrhagia in patients with ULs [45]. *FGFR1* has been reported as a potential therapeutic target in breast cancers [46]. The *in silico* functional analysis showed an association with *FGF1* molecule and pazopanib, a tyrosine kinase

inhibitor recently approved to renal cell carcinoma (RCC) treatment [47]. Pazopanib acts in RCC through its antiangiogenic properties via inhibition of the intracellular tyrosine kinases. Thus, *FGFR1* amplification translated by up-regulation is a target of studies with drugs that inhibit cellular growth in ULs menorrhagia-associated.

The ERK (extracellular-regulated kinase)/MAPK (mitogen activated protein kinase) signaling is a key pathway that transduces cellular information on meiosis/mitosis, growth, differentiation and carcinogenesis. The *HSPB7* encodes one heat shock 27kDa protein family, member 7 that has been associated with cardiomyopathy and belongs to ERK/MAPK signaling [48]. To the best of our knowledge, this is the first report showing the involvement of *HSPB7* in ULs.

Vitamin D3-bound VDR-RXR, along with other co-activator proteins, mediates the transcriptional regulation of a number of genes, including *IGFBP5*. This gene encodes insulin-like growth factor binding protein 5 that stimulates gene targets transcription leading to cell proliferation. Up-regulation of *IGFBP5* has been associated with cervical intraepithelial neoplasia and breast cancer [49, 50]. Therefore, increased activity of *IGFBP5* could be associated with estrogen-dependent tumors, as well as ULs.

A principal pathway that signals through mTOR is the PI3K/AKT that is involved in cell survival and proliferation. Crabtree et al. [39] reported the mammalian target of rapamycin (mTOR) pathway as one of the most highly up-regulated pathways in both human and rat ULs. The *RHOH*, associated with mTOR signaling, encodes a small G-like protein. Down-regulation of *RHOH* was associated with acute myeloid leukemia [51]). The present findings suggest that *RHOH* genomic losses and transcript down-regulation could be associated with deregulation of cell cycle and mTOR signaling pathway in ULs.

FAK signaling regulates a number of key cellular processes including growth factor-induced mitogenic signals, cell survival, cell proliferation and migration, cell locomotion and

regulation of cell cycle. Chegini and Kornberg [52] demonstrated that GnRHa therapy resulted in a noticeable decrease in FAK in ULs cells. These data suggested that ULs regression induced by GnRHa is mediated in part through a mechanism involving growth factors and adhesion molecules. The *TNSI* gene associated on FAK signaling acts in extracellular matrix production. Therefore, the current study showing that up-regulation of *TNSI* associated with genomic amplifications gives an additional support for treatment with GnRH analogs in ULs.

Heterotrimeric G proteins are the key players in transmembrane signaling by coupling a multitude of receptors to enzymes, channel proteins and other effector molecules. The *CALCRL* gene encodes the receptor-like calcitonin that forms complexes with protein-G. The *CALCRL* mediates the effects of adrenomedullin (AM), which is an angiogenic factor induced by hypoxia. Despite its benign nature, ULs have aberrant angiogenesis, essential step for the growth and proliferation [for review 4]. Nikitenko et al. [53] reported an increased activity of *CALCRL* in the vascular endothelium of ULs suggesting that *CALCRL* could be a potential therapeutic target for angiogenesis inhibitors in ULs. In agreement with the findings, *CALCRL* gains and up-regulation could explain the angiogenesis phenotype detected in ULs.

The hepatic fibrosis/hepatic stellate cell activation pathways revealed an association of *COL3A1*, *FGFR1*, and *IGFBP5*. Hepatic fibrosis is a chronic liver disease associated with the extracellular matrix accumulation, very similar to pathological conditions of ULs. The inhibitor of tyrosine kinase activity of *FGFR1* (NP603) recently described, inhibits the proliferation of myofibroblasts associated with liver fibrosis in rats [54]. In ULs, the increased tyrosine kinase activity of *FGFR1* in association with *IGFBP5* in response to FGF1, IGF-1 or TGF-beta might trigger cell proliferation, while increased activity of *COL3A1*, in response to ET-1 could lead to extracellular matrix accumulation and collagen deposition. Additional

studies using *in vivo* models are needed to assess the rate of response to drugs that inhibit cell growth and proliferation in ULs.

In conclusion, the integrative analysis of genomic and transcriptomic data provided a comprehensive and biologically meaningful insight into tumorigenesis of ULs revealing genomic amplifications translated by up-regulation of modulators. To our knowledge, this is the first study using a large series of ULs evaluated by integrative genomic and transcriptomic analysis. The findings indicate genes that could be targets for development of specific therapies related to extracellular matrix and cell proliferation related to Uterine Leiomyomas.

REFERENCES

- [1] Baird DD, Dunson DB (2003) Why is parity protective for uterine fibroids? *Epidemiology* 14: 247-250.
- [2] Mantovani MS, Barbieri Neto J, Philbert PMP (1999) Multiple uterine leiomyomas: Cytogenetic analysis. *Gynecol Oncol* 72: 71–75.
- [3] Cook H, Ezzati M, Segars JH et al. (2010) The impact of uterine leiomyomas on reproductive outcomes. *Minerva Ginecol* 62: 225-236.
- [4] Flake GP, Andersen J, Dixon D (2003) Etiology and pathogenesis of uterine leiomyomas: a review. *Environ Health Perspect* 111: 1037-1054.
- [5] Ishikawa H, Ishi K, Serna VA et al. (2010) Progesterone is essential for maintenance and growth of uterine leiomyoma. *Endocrinology* 151:2433-2442.
- [6] Sozen I and Arici A (2002) Interactions of cytokines, growth factors, and the extracellular matrix in the cellular biology of uterine leiomyomata. *Fertil Steril* 78: 1-12.
- [7] Marsh EE, Lin Z, Yin P, Milad M et al. (2008) Differential expression of microRNA species in human uterine leiomyoma versus normal myometrium. *Fertil Steril* 89: 1771-1776.
- [8] Bonatz G, Frahm SO, Andreas S et al. (1998) Telomere shortening in uterine leiomyomas. *Am J Obstet Gynecol* 179: 591-596.
- [9] Malik M, Norian J, McCarthy-Keith D et al. (2010) Why leiomyomas are called fibroids: the central role of extracellular matrix in symptomatic women. *Semin Reprod Med* 28: 169–179.
- [10] Canevari RA, Pontes A, Rogatto SR (2005) Microallelotyping Defines Novel Regions of Loss of Heterozygosity in Uterine Leiomyomas. *Mol Carcinog* 42: 177–182.
- [11] Canevari, RA and Rogatto SR (2007) Uterine Leiomyoma: Updates in Cytogenetics and Molecular Analysis. *Asian Journal of Cancer* 6: 15-28.
- [12] El-Gharib MN and Elsobky ES (2010) Cytogenetic aberrations and the development of uterine leiomyomata. *J Obstet Gynaecol Res* 36: 101–107.
- [13] Arslan AA, Gold LI, Mittal K et al. (2005) Gene expression studies provide clues to the pathogenesis of uterine leiomyoma: new evidence and a systematic review. *Human Reproduction* 20: 852-863.
- [14] Quade BJ, Wang TY, Sornberger K et al. (2004) Molecular pathogenesis of uterine smooth muscle tumors from transcriptional profiling. *Genes Chrom Cancer* 40: 97–108.
- [15] Vanharanta S, Wortham NC, Laiho P et al. (2005) 7q deletion mapping and expression profiling in uterine fibroids. *Oncogene* 24: 6545-6554.
- [16] Hodge JC, Park PJ, Dreyfuss JM et al. (2009) Identifying the molecular signature of the interstitial deletion 7q subgroup of uterine leiomyomata using a paired analysis. *Genes Chrom Cancer* 48: 865–885.
- [17] Akavia UD, Litvin O, Kim J et al. (2010) An integrated approach to uncover drivers of cancer. *Cell* 143: 1005–1017.
- [18] Staaf J, Jönsson G, Ringnér M et al. (2010) High-resolution genomic and expression analyses of copy number alterations in HER2-amplified breast cancer. *Breast Cancer Research*, 12: R25.

- [19] Green MR, Aya-Bonilla C, Gandhi MK et al. (2011) Integrative Genomic Profiling Reveals Conserved Genetic Mechanisms for Tumorigenesis in Common Entities of Non-Hodgkin's Lymphoma. *Genes Chrom Cancer* 50: 313–326.
- [20] Jönsson G, Staaf J, Vallon-Christersson J et al. (2010) Genomic subtypes of breast cancer identified by array-comparative genomic hybridization display distinct molecular and clinical characteristics. *Breast Cancer Research* 12: R42.
- [21] Noyes RW, Hertig AT, Rock J (1950) Dating the endometrial biopsy. *Fertil Steril* 1: 3-25.
- [22] Sanchez-Garcia F, Akavia UD, Mozes E et al. (2010) Identification of Significant Targets in Cancer. *BMC Bioinfo* 11: 189-198.
- [23] Gomes LI, Esteves GH, Carvalho AF, et al. (2005) Expression profile of malignant and nonmalignant lesions of esophagus and stomach: differential activity of functional modules related to inflammation and lipid metabolism. *Cancer Res* 65: 7127-7136.
- [24] Tusher VG, Tibshirani R and Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *PNAS* 98: 5116–5121.
- [25] Benjamini Y and Hochberg Y (1995) Controlling the false discovery rate — a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol* 57: 289–300.
- [26] Storey JD and Tibshirani R (2003) Statistical significance for genome-wide studies. *Proc Natl Acad Sci U S A* 100: 9440–9445.
- [27] Segal E, Shapira M, Regev A et al. (2003) Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genetics* 34: 166-176.
- [28] Subramanian A, Tamayo P, Mootha VK et al. (2005) Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *PNAS* 102: 15545-15550.
- [29] Reich M, Liefeld T, Gould J et al. (2006) GenePattern 2.0. *Nature Genetics* 38: 500-501.
- [30] Packenham JP, Du Manoir S, Schrock E et al. (1997) Analysis of genetic alterations in uterine leiomyomas and leiomyosarcomas by comparative genomic hybridization. *Mol Carcinog* 19: 273-279.
- [31] Levy B, Mukherjee T, Hirschhorn K (2000) Molecular cytogenetic analysis of uterine leiomyomata and leiomyosarcoma by comparative genomic hybridization. *Cancer Genet Cytogenet* 121: 1-8.
- [32] Vanharanta S, Wortham NC, Langford C et al. (2007) Definition of a minimal region of deletion of chromosome 7 in uterine leiomyomas by tiling-path microCGH array and mutation analysis of known genes in this region. *Genes Chrom Cancer* 46: 451–458.
- [33] Ryan BM, Robles AI and Harris CC (2010) Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer* 10: 389-402.
- [34] Zavadil J, Ye H, Liu Z et al. (2010) Profiling and Functional Analyses of MicroRNAs and Their Target Gene Products in Human Uterine Leiomyomas. *PLoS ONE* 5: e12362.
- [35] Hoeijmakers JHJ (2001) Genome maintenance mechanisms for preventing cancer. *Nature* 411: 366-374.
- [36] Holthausen JT, Wyman C, Kanaar R (2010) Source. Regulation of DNA strand exchange in homologous recombination. *DNA Repair (Amst)* 9: 1264-1272.

- [37] Luo X, Ding L, Xu J et al. (2005) Gene expression profiling of leiomyoma and myometrial smooth muscle cells in response to transforming growth factor-beta. *Endocrinol* 146: 1097–1118.
- [38] Dimitrova IK, Richer JK, Rudolph MC et al. (2009) Gene expression profiling of multiple leiomyomata uteri and matched normal tissue from a single patient. *Fertil Steril* 91: 2650-2663.
- [39] Crabtree JS, Jelinsky SA, Harris HA et al. (2009) Comparison of human and rat uterine leiomyomata: identification of a dysregulated mammalian target of rapamycin pathway. *Cancer Res* 69: 6171-6178.
- [40] Shah AA, Grainger DA (1996) Contemporary Concepts in Managing Menorrhagia. *Medscape Womens Health* 12: 8.
- [41] Behera MA, Feng L, Yonish B et al. (2007) Thrombospondin-1 and Thrombospondin-2 mRNA and TSP-1 and TSP-2 protein expression in uterine fibroids and correlation to the genes COL1A1 and COL3A1 and to the collagen cross-link hydroxyproline. *Reprod Sci* 14: 63-76.
- [42] Desai SS, Hentz VR (2011) The treatment of Dupuytren disease. *J Hand Surg Am* 36: 936-942.
- [43] Shih B, Bayat A (2010) Scientific understanding and clinical management of Dupuytren disease. *Nat Rev Rheumatol* 6: 715-726.
- [44] Yu L, Saile K, Swartz CD et al. (2008) Differential expression of receptor tyrosine kinases (RTKs) and IGF-I pathway activation in human uterine leiomyomas. *Mol Med* 14:264-275.
- [45] Anania CA, Stewart EA, Quade BJ et al. (1997) Expression of the fibroblast growth factor receptor in women with leiomyomas and abnormal uterine bleeding. *Mol Hum Reprod* 3: 685-691.
- [46] Shiang CY, Qi Y, Wang B et al. (2010) Amplification of fibroblast growth factor receptor-1 in breast cancer and the effects of brivanib alaninate. *Breast Cancer Res Treat* 123: 747-755.
- [47] Keisner SV, Shah SR (2011) Pazopanib: the newest tyrosine kinase inhibitor for the treatment of advanced or metastatic renal cell carcinoma. *Drugs* 71:443-454.
- [48] Villard E, Perret C, Gary F et al. (2011) A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J* 32:1065-1076.
- [49] Hou XJ, Zhang YZ, Liu X et al. (2009) Expressions of IGFBP-5, cFLIP in cervical intraepithelial neoplasia, cervical carcinoma and their clinical significances: a molecular pathology. *J Exp Clin Cancer Res* 28: 1-16.
- [50] Taylor KJ, Sims AH, Liang L et al. (2010) Dynamic changes in gene expression in vivo predict prognosis of tamoxifen-treated patients with breast cancer. *Breast Cancer Res* 12: R39.
- [51] Galiègue-Zouitina S, Delestré L, Dupont C et al. (2008) Underexpression of RhoH in Hairy Cell Leukemia. *Cancer Res*, 68:4531-4540.
- [52] Chegini N, Kornberg L (2003) Gonadotropin releasing hormone analogue therapy alters signal transduction pathways involving mitogen-activated protein and focal adhesion kinases in leiomyoma. *J Soc Gynecol Investig* 10:21-26.

- [53] Nikitenko LL, Cross T, Campo L et al. (2006) Expression of terminally glycosylated calcitonin receptor-like receptor in uterine leiomyoma: endothelial phenotype and association with microvascular density. *Clin Cancer Res* 12: 5648-5658.
- [54] Lin N, Chen S, Pan W et al. (2011) NP603, a novel and potent inhibitor of FGFR1 tyrosine kinase, inhibits hepatic stellate cell proliferation and ameliorates hepatic fibrosis in rats. *Am J Physiol Cell Physiol*, in press.

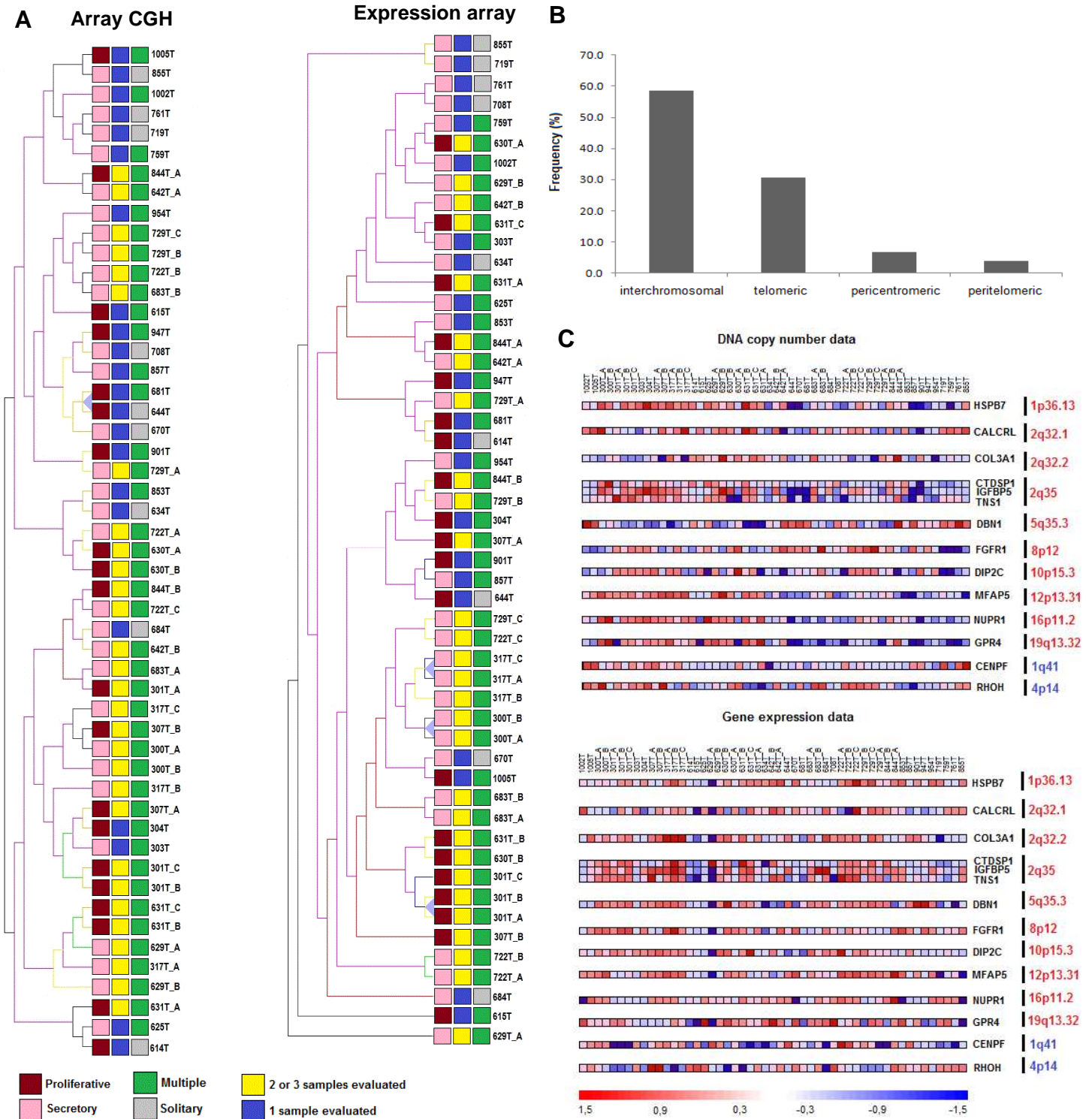


Figure 1. Characterization of modulators **A.** Hierarchical clustering of samples according to menstrual cycle phase (proliferative and secretory), number of samples evaluated (2 or 3 samples or 1 sample) and diagnosis of multiple or solitary tumors (TMeV v.4.5). **B.** Chromosomal mapping of 75 modulators which most of them (~60%) were mapped at interchromosomal regions and telomeric regions (30%). Around 6% were mapped in regions close to telomeres. **C.** Selection of genes with positive (red) and negative (blue) association among top 30 gene modulators, totalizing 14 genes as potential drivers [29].

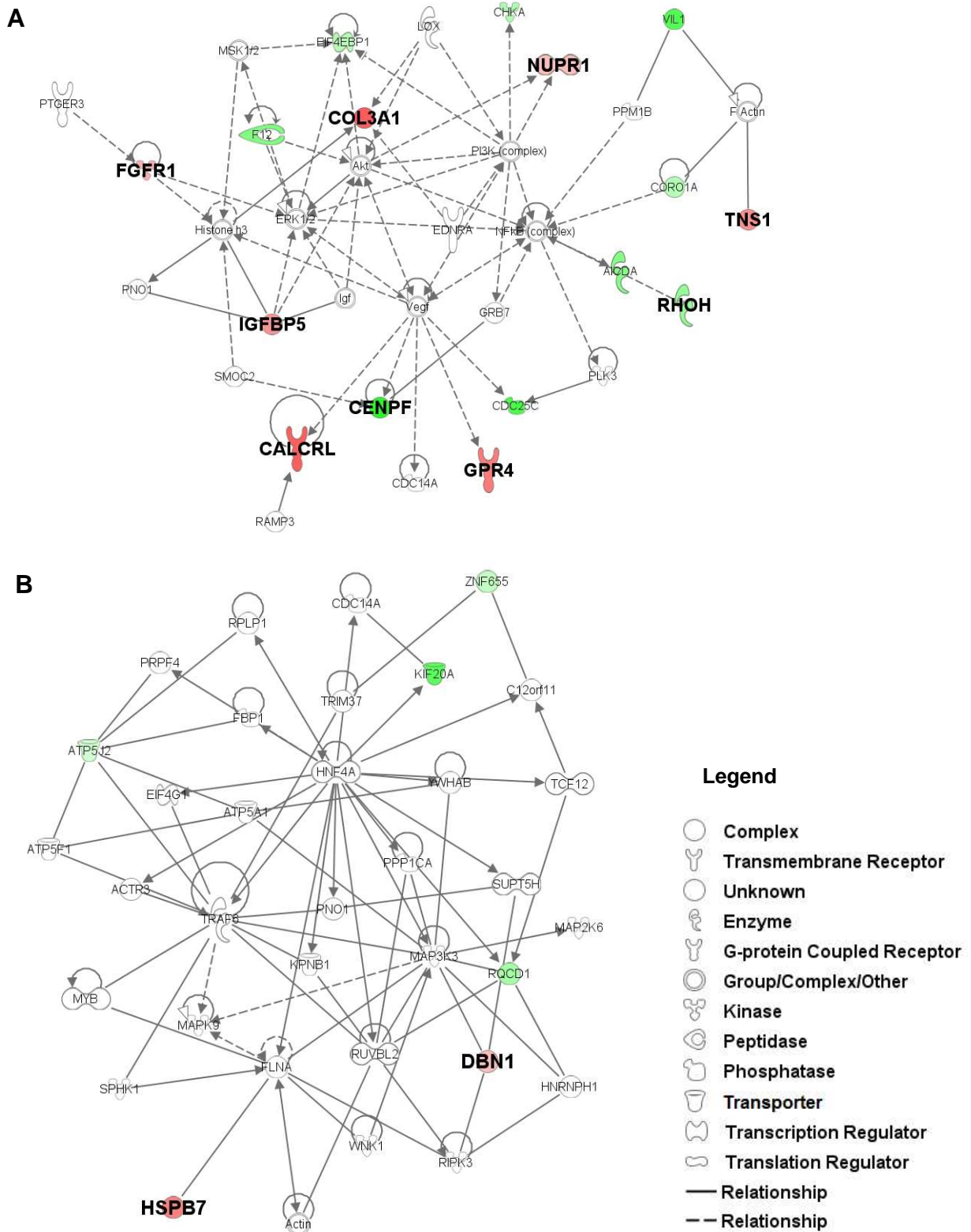
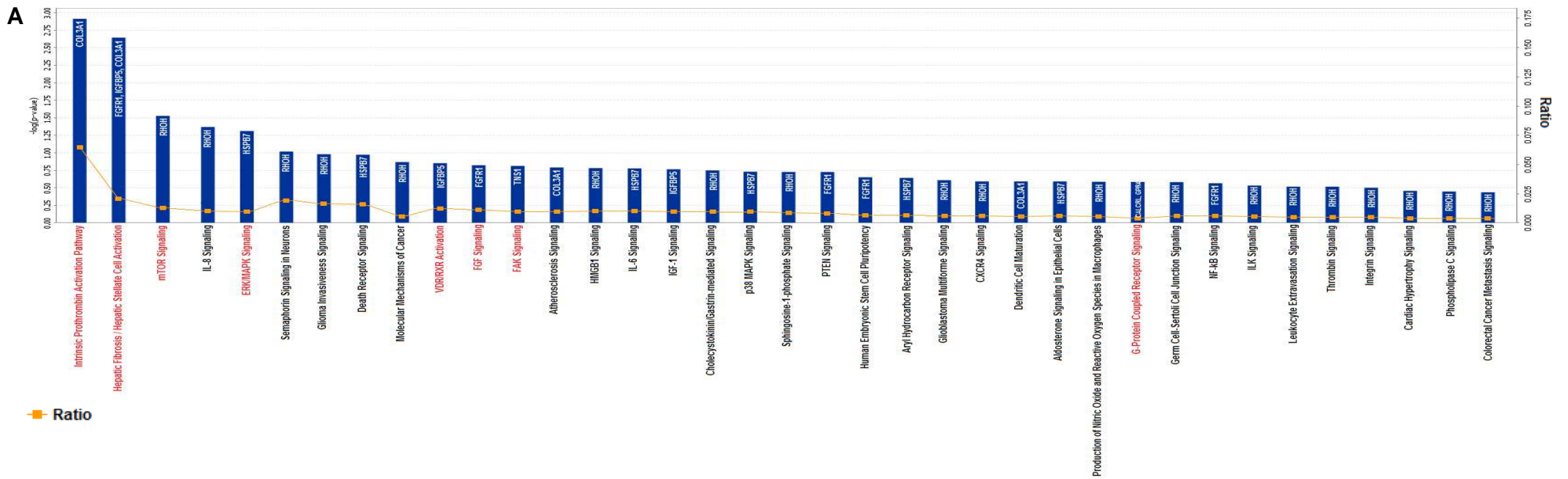


Figure 2. Networks generated from top 30 modulators and IPA database **A.** Network 1 showing interactions among 9/30 modulators. Genes were located on network periphery interconnected with Akt, ERK, PI3K, NFKB, and VEGF central complex. **B.** Network 2 with *DBN1* and *HSPB7* gene interactions with FLNA, MAP3K3, and HNF4A central molecules (IPA). Genes in the network with positive (red) and negative (green) association.



B

Diseases and Disorders	P value	Molecules
Cancer	1.88E-04 - 4.87E-02	15
Reproductive System Disease	3.67E-04 - 4.87E-02	8
Genetic Disorder	6.68E-04 - 4.16E-02	9
Metabolic Disease	6.68E-04 - 1.53E-02	3
Connective Tissue Disorders	1.93E-03 - 3.61E-02	1
Molecular and Cellular Functions	P value	Molecules
Cell Cycle	1.28E-04 - 4.72E-02	7
Amino Acid Metabolism	1.93E-03 - 1.93E-03	1
Cellular Assembly and Organization	1.93E-03 - 4.35E-02	7
Cellular Growth and Proliferation	1.93E-03 - 4.53E-02	10
Molecular Transport	1.93E-03 - 3.61E-02	2

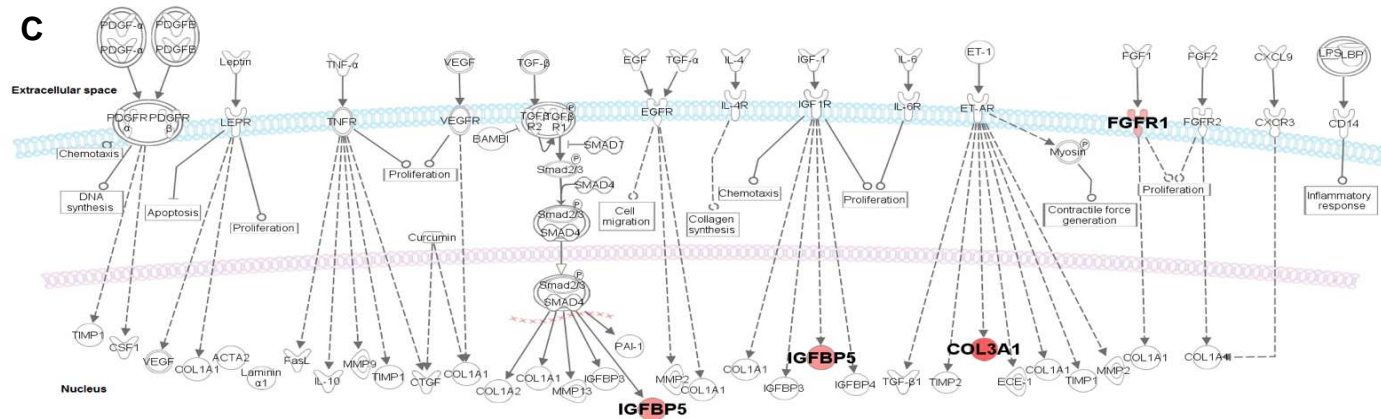


Figure 3. Canonical pathways associated with top 30 modulators **A.** Sixty-one canonical pathways were associated with modulators. In red, top pathways associated with 8/14 selected genes, which have been described as drug targets. **B.** Diseases and biological functions statistically associated with top 30 modulators such as cancer and cell cycle, respectively. **C.** Pathway showing cell proliferation and extracellular matrix deposition process with high similarity observed in ULs. Genes *COL3A1*, *FGFR1*, and *IGFBP5* were selected as potential molecular markers to ULs treatment.

Table 1. Seventy-five modulators obtained from integrative analysis.

JISTIC q-value	CONEXIC score	Array CGH	Expression array	Region	Position Hg18 (Start-End)	Gene symbol
<0.01	9832.22	+	+	1p36.13*	16,213,110-16,217,872	<i>HSPB7</i>
<0.01	2280.2	-	-	1q41*	212,843,155-212,904,537	<i>CENPF</i>
<0.01	3214.2	+	+	2q32.1*	187,916,094-188,021,266	<i>CALCRL</i>
<0.01	3211.1	+	+	2q32.2*	189,547,344-189,585,717	<i>COL3A1</i>
<0.01	2576.68	+	+	2q35*	217,245,073-217,268,517	<i>IGFBP5</i>
<0.01	15242.9	+	+		218,372,757-218,517,041	<i>TNSI</i>
<0.01	1774	+	+		218,972,722-218,978,908	<i>CTDSP1</i>
<0.01	3222	+	-		218,996,724-219,022,487	<i>VILI</i>
<0.01	8213.98	+	-		219,141,922-219,167,243	<i>RQCD1</i>
<0.01	3009.88	-	-	4p14*	39,874,922-39,922,676	<i>RHOH</i>
<10 ⁻⁹	21227.32	¥+	-	5q31.2*	137,543,248-137,551,261	<i>KIF20A</i>
<0.01	2487.12	¥+	-		137,648,858-137,695,415	<i>CDC25C</i>
<0.01	2290.81	¥+	-	5q35.3	176,663,441-176,666,556	<i>PRELID1</i>
<0.01	2001.89	¥+	-		176,761,745-176,769,183	<i>FI2</i>
<0.01	3678.31	+	+		176,816,220-176,833,300	<i>DBN1</i>
<0.01	3222.8	¥+	-	7q22.1*	98,874,499-98,892,932	<i>CPSF4</i>
<0.01	4222.7	¥+	-		98,893,720-98,901,744	<i>ATP5J2</i>
<0.01	1908.18	¥+	-		98,994,384-99,012,013	<i>ZNF655</i>
<0.01	1232.98	+	-		99,402,286-99,411,623	<i>AZGP1</i>
<0.01	442.98	¥+	-		99,528,340-99,537,363	<i>MCM7</i>
0.05	1132.32	+	-		99,613,474-99,649,946	<i>STAG3</i>
<0.01	1883.3	¥+	-	8q24.3	144,520,168-144,522,180	<i>C8orf51</i>
<0.01	998.3	+	-		144,757,532-144,762,887	<i>PYCR1</i>
<0.01	774.1	+	-		145,221,948-145,224,416	<i>CYC1</i>
<0.01	523.2	+	+		145,619,808-145,624,735	<i>VPS28</i>
<0.01	1124.2	+	+		145,624,971-145,640,620	<i>TONSL</i>
<0.01	523.65	+	-		145,707,479-145,714,008	<i>RECQL4</i>
<0.01	235.65	¥+	-		145,714,199-145,721,365	<i>LRRC14</i>
<0.01	998.4	+	+	8p12	37,773,932-37,820,650	<i>GPR124</i>
<0.01	1534.34	¥+	-		38,007,177-38,037,040	<i>EIF4EBP1</i>
<0.01	1552.55	+	+		38,387,813-38,445,509	<i>FGFR1</i>
<0.01	1523.52	¥+	-	10p15.3	1,075,964-1,085,061	<i>ID11</i>
<0.01	1774.98	+	+		311,432-725,606	<i>DIP2C</i>
<0.01	1883.3	¥+	-	10q21.3	69,539,196-69,641,779	<i>MYPN</i>
0.01	998.3	¥+	-		70,385,898-70,414,285	<i>DDX21</i>
<0.01	774.1	¥+	-		70,650,065-70,697,321	<i>HKDC1</i>
<0.01	523.2	+	+	11p15.5	303,991-305,272	<i>IFITM1</i>
<0.01	1124.2	+	-		384,217-394,908	<i>PKP3</i>
0.02	523.65	¥+	-		522,242-525,550	<i>HRAS</i>
<0.01	235.65	+	-		551,450-554,018	<i>RASSF7</i>
<0.01	998.4	+	-		737,432-755,024	<i>TALDO1</i>
<0.01	1534.34	¥+	-		780,475-786,221	<i>SLC25A22</i>
<0.01	1552.55	¥+	-	11q13.2*	67,576,902-67,645,434	<i>CHKA</i>
<0.01	1523.52	-	-	12p11.21	32,151,452-32,422,408	<i>BICD1</i>
<0.01	1774.98	¥+	-	12p13.31*	8,646,029-8,656,706	<i>AICDA</i>
<0.01	2001.32	+	+		8,689,807-8,706,700	<i>MFAP5</i>
<0.01	1232	¥+	-	14q13.2*	34,291,688-34,414,604	<i>BAZ1A</i>
<0.01	1822.11	+	+	16p11.2	28,456,163-28,457,996	<i>NUPR1</i>
<0.01	909.81	¥+	-		28,850,761-28,858,164	<i>CD19</i>
<0.01	987.05	+	+		29,730,910-29,734,703	<i>PRRT2</i>
<0.01	1034.55	+	+		29,739,288-29,766,842	<i>MVP</i>
<0.01	1993.31	+	-		30,023,632-30,032,379	<i>GDPD3</i>
<0.01	1738.2	+	-		30,102,427-30,107,898	<i>CORO1A</i>
<0.01	783.5	+	+		30,488,524-30,491,229	<i>ZNF688</i>
<0.01	984.2	+	+	16q22.1*	65,775,770-65,781,608	<i>EXOC3L1</i>
0.05	578.87	+	-		65,790,529-65,795,428	<i>ELMO3</i>
0.01	1378.55	+	+	16q24.3	88,301,042-88,314,895	<i>C16orf7</i>
<0.01	566.98	¥+	-		88,331,460-88,410,566	<i>FANCA</i>
<0.01	665.35	¥+	-		88,517,246-88,530,006	<i>TUBB3</i>
<0.01	533.8	¥+	-	17q21.31*	38,449,840-38,530,994	<i>BRCA1</i>
<0.01	968.66	+	+		38256727-38263666	<i>AOC3</i>
<0.01	1372.76	+	+	19q13.32*	50,702,528-50,722,080	<i>VASP</i>

<0.01	1799.91	+	+	50,784,865-50,787,557	<i>GPR4</i>
0.01	879.54	+	-	50,882,558-50,887,282	<i>SNRPD2</i>
<0.01	876.1	+	+	50,964,816-50,977,655	<i>DMPK</i>
<0.01	866.98	+	+	51,134,631-51,168,497	<i>NOVA2</i>
<0.01	655.5	+	+	51,605,929-51,608,681	<i>CCDC8</i>
<0.01	675.01	¥+	-	51,796,352-51,805,879	<i>CALM3</i>
<0.01	989.32	+	+	51,842,709-51,856,235	<i>DACT3</i>
<0.01	377.3	¥+	-	51,969,980-51,983,653	<i>SLC1A5</i>
<0.01	656.33	¥+	-	52,033,263-52,046,043	<i>AP2S1</i>
<0.01	886.44	¥+	-	19q13.33* 55,614,028-55,624,060	<i>SPIB</i>
<0.01	1252.32	¥-	+	20p11.2 23,007,993-23,014,977	<i>CD93</i>
0.17	0	¥+	-	20p13 309,308-326,203	<i>TRIB3</i>
0.23	0	¥+	-	364,124-391,187	<i>TBC1D20</i>

In bold, top 30 modulators based on CONEXIC scores; Hg18: Human genome version 18 (Mar 2006 NCBI36); ¥ miRNA target prediction; Positive (+) and negative (-) signs indicate gene status in relation to genomic gains and losses and up- or down-regulated gene expression, respectively. *Regions not usually involved in chromosomal breakpoints.

Table 2. Genes identified on modules of cancer genes obtained from gene set enrichment analysis – GSEA.

Gene Set Name	Genes in Gene Set (K)	Genes in Overlap (k)	k/K	P value
533	42	<i>CENPF, CORO1A, AOC3</i>	0.0714	<10 ⁻³
8	406	<i>CORO1A, MCM7, SLC1A5, DDX21, NUPRI, DBN1, EIF4EBP1, VILI, CHKA</i>	0.0222	<10 ⁻³
509	14	<i>AP2S1, ID1</i>	0.1429	<10 ⁻³
315	15	<i>CENPF, TUBB3</i>	0.1333	<10 ⁻³
5	423	<i>CORO1A, AOC3, DDX21, NUPRI, FGFR1, CALCRL, COL3A1, MVP</i>	0.0189	<10 ⁻²
292	133	<i>IFITM1, SPIB, CD19, STAG3</i>	0.0301	<10 ⁻²
3	375	<i>MCM7, SLC1A5, DDX21, DBN1, EIF4EBP1, FGFR1, CALCRL</i>	0.0187	<10 ⁻²
118	395	<i>CENPF, CORO1A, DDX21, DBN1, EIF4EBP1, FGFR1, RECQL4</i>	0.0177	<10 ⁻²
86	43	<i>HRAS, RHOH</i>	0.0465	<10 ⁻²
137	531	<i>CORO1A, NUPRI, DBN1, EIF4EBP1, FGFR1, HRAS, NOVA2, BICD1</i>	0.0151	<10 ⁻²
12	351	<i>DDX21, NUPRI, DBN1, FGFR1, COL3A1, DIP2C</i>	0.0171	<10 ⁻²
9	114	<i>EIF4EBP1, VILI, HRAS</i>	0.0263	<10 ⁻²
430	52	<i>CORO1A, FANCA</i>	0.0385	<10 ⁻²
1	361	<i>AOC3, DDX21, NUPRI, DBN1, FGFR1, COL3A1</i>	0.0166	<10 ⁻²
105	194	<i>CHKA, ID1, FGFR1, COL3A1</i>	0.0206	<10 ⁻²
57	54	<i>MCM7, CDC25C</i>	0.0370	<10 ⁻²
345	122	<i>IFITM1, SPIB, CD19</i>	0.0246	<10 ⁻²
254	58	<i>NUPRI, IFITM1</i>	0.0345	<10 ⁻²
337	60	<i>CENPF, DDX21</i>	0.0333	<10 ⁻²
53	391	<i>CENPF, CORO1A, DDX21, DBN1, EIF4EBP1, RECQL4</i>	0.0153	<10 ⁻²
129	217	<i>DBN1, FGFR1, HRAS, NOVA2</i>	0.0184	<10 ⁻²
119	138	<i>SPIB, CD19, STAG3</i>	0.0217	<10 ⁻²
122	138	<i>AOC3, COL3A1, MFAP5</i>	0.0217	<10 ⁻²
308	67	<i>CORO1A, CDC25C</i>	0.0299	<10 ⁻²
438	67	<i>CENPF, CORO1A</i>	0.0299	<10 ⁻²
32	232	<i>SNRPD2, EIF4EBP1, TUBB3, RHOH</i>	0.0172	<10 ⁻¹

GSEA - <http://www.broadinstitute.org/gsea/index.jsp>; K=overlapping modulators associated with modules; k=differentially expressed genes associated with modulators; k/K=ratio between genes on modules.

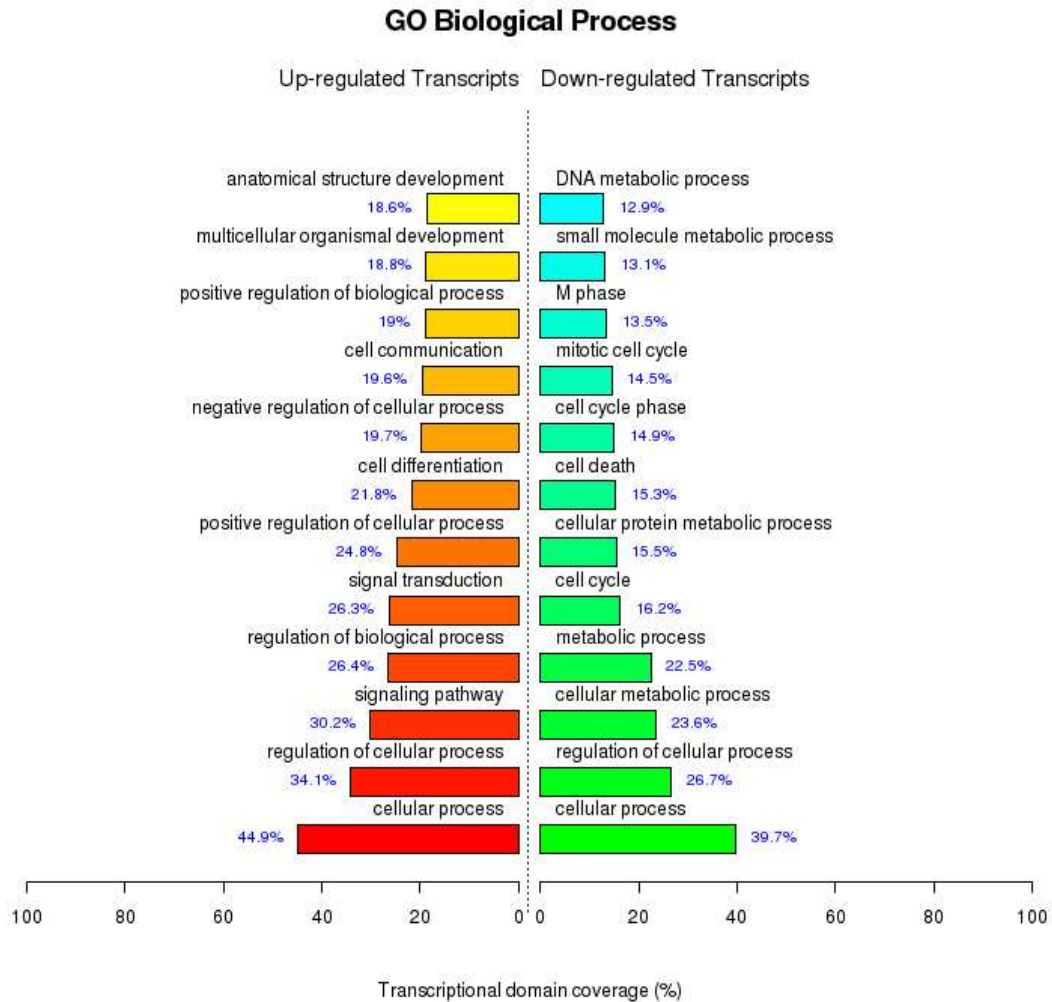


Figure S1. Biological functions of 3325 significant genes identified by gene expression microarrays. Differentially expressed transcripts (\log_2 ratio) were classified based on fold-change ≤ -1.0 (down-regulated) and ≥ 1.0 (up-regulated). Among up-regulated genes (red), 44.9% had biological functions associated with cellular process. Approximately 40% (39.7%) of down-regulated genes (green) showed the same functional category. (FUNNET - <http://www.funnet.info/>).

Table S1. Clinical parameters and histopathological data from patients with Uterine Leiomyomas.

ID	Samples*	Age	N° tumors	Menstrual cycle phase	Race	Menarche age	BMI	Age at 1 ^a pregnancy	N° pregnancy	Familial history of cancer	Personal history of cancer
1	300T_A 300T_B	36	M	secretory	W	13	26,2	ND	3	no	no
2	301T_A 301T_B 301T_C	49	M	proliferative	B	11	36,1	15	7	father prostate ca	no
3	303T	47	M	secretory	W	13	29,3	23	1	father ca?	no
4	304T	44	M	proliferative	W	13	26,6	ND	10	grandmother bowel ca	no
5	307T_A 307T_B	48	M	proliferative	W	12	34,9	26	4	no	no
6	317T_A 317T_B 317T_C	38	M	secretory	B	12	24,8	20	5	no	no
7	614T	44	S	proliferative	W	12	29,9	17	10	no	no
8	615T	43	M	proliferative	W	ND	22,9	0	0	no	no
9	625T	47	M	secretory	W	12	24,5	24	3	aunt breast ca / aunt stomach ca	no
10	629T_A 629T_B	49	M	secretory	W	15	21,2	26	3	sister UL	no
11	630T_A 630T_B	47	M	proliferative	W	13	26,3	18	2	no	no
12	631T_A 631T_B 631T_C	43	M	proliferative	B	9	24,1	ND	0	no	no

13	634T	44	S	secretory	W	15	23,6	23	3	ND	ND
14	642T_A 642T_B	46	M	secretory	W	9	29,9	21	1	mother uterine ca	no
15	644T	39	S	proliferative	W	12	24,3	ND	2	no	no
16	670T	38	S	secretory	W	15	26,6	18	3	no	no
17	681T	46	M	proliferative	W	15	21,5	21	4	aunt uterine ca	endometrial polyp
18	683T_A 683T_B	49	M	secretory	W	13	28,7	23	3	sister UL	no
19	684T	49	S	secretory	W	13	25,6	ND	6	no	no
20	708T	45	S	secretory	W	13	33,9	22	3	aint breast ca	no
21	719T	50	S	secretory	W	13	24,2	0	0	father died metastatic ca	no
22	722T_A 722T_B 722T_C 729T_A	50	M	secretory	W	11	34,9	ND	4	mother renal ca	no
23	729T_B 729T_C	46	M	secretory	W	12	24,3	25	1	no	no
24	759T	51	M	secretory	W	15	21,2	ND	3	no	no
25	761T	45	S	secretory	W	13	18,7	ND	3	mother bowel ca	no
26	844T_A 844T_B	50	M	proliferative	W	12	36,8	ND	4	no	UL
27	853T	35	M	secretory	W	11	26,8	21	2	no	no
28	855T	37	S	secretory	W	13	24,9	23	2	uncle stomach ca	no
29	857T	41	M	secretory	W	13	23,7	20	3	three sisters UL	no
30	901T	51	M	proliferative	W	9	34,1	ND	3	no	no
31	947T	47	M	proliferative	W	ND	27,6	ND	ND	no	no

32	954T	45	M	secretory	W	12	33,7	18	3	no	no
33	1002T	46	M	secretory	W	13	27,5	18	2	no	no
34	1005T	40	M	proliferative	B	12	34,2	ND	2	mother colon ca	no

*Number of samples evaluated by case; N^o=number; 1^a=first; ND=undefined; M= multiples; S=solitary; B=black; W=white; BMI=body index mass; ca=cancer; ?= doubt in response.

Table S2. Recurrent copy number alterations identified by JISTIC among 51 Uterine Leiomyomas samples.

Chromosome	Position	Size (bp)	Cytoband	Event
1	16210136-16404475	194339	p36.13	Gain
1	193404142-194937836	1533694	q31.3	Loss
1	212893807-214486070	1592263	q41	Loss
2	187336585-189552794	2216209	q32.1 - q32.2	Gain
2	217253192-219351693	2098501	q35	Gain
4	39121392-40671431	1550039	p14	Loss
4	63649592-66171072	2521480	q13.1	Loss
4	132475881-135650424	3174543	q28.3	Loss
5	32031565-32459137	427572	p13.3	Gain
5	134274907-135585827	1310920	q31.1 - q31.2	Loss
5	137509066-137894706	385640	q31.2	Gain
5	176628647-176902434	273787	q35.3	Gain
7	98882914-102068500	3185586	q22.1	Gain
7	124268700-124352498	83798	q31.33	Loss
7	150157293-150323899	166606	q36.1	Gain
8	37606006-38685823	1079817	p12 - p11.23	Gain
8	105119923-108955123	3835200	q22.3 - q23.1	Gain
8	143934407-144604977	670570	q24.3	Gain
8	144723137-146027963	1304826	q24.3	Gain
9	33871385-34349039	477654	p13.3	Gain
9	34380035-34602962	222927	p13.3	Gain
9	42014069-42702421	688352	p12 - p11.2	Gain
9	137569888-138038551	468663	q34.3	Loss
9	138467512-138845556	378044	q34.3	Gain
10	0-1167989	1167989	p15.3	Gain
10	69249439-70694950	1445511	q21.3	Gain
11	209012-838551	629539	p15.5	Gain
11	66956644-67906757	950113	q13.1 - q13.2	Gain
11	82287494-82844323	556829	q14.1	Gain
12	7620638-8938808	1318170	p13.31	Gain
12	31233552-32196110	962558	p11.21	Loss
12	109668284-109732691	64407	q24.11	Gain
13	79882431-84123782	4241351	q31.1	Loss
14	34245057-34584738	339681	q13.2	Gain
14	104210820-104323543	112723	q32.33	Gain
16	154832-649034	494202	p13.3	Gain
16	14674167-15423013	748846	p13.12 - p13.11	Loss

16	28184420-30915100	2730680	p11.2	Gain
16	34059589-34739434	679845	p11.2 - p11.1	Loss
16	65732841-65799873	67032	q22.1	Gain
16	65977069-66551504	574435	q22.1	Gain
16	74665577-75131336	465759	q23.1	Loss
16	88147752-88526617	378865	q24.3	Gain
17	10029176-12411350	2382174	p13.1 - p12	Gain
17	36917868-36931811	13943	q21.2	Gain
17	38117300-39051741	934441	q21.31	Gain
17	76857866-77619424	761558	q25.3	Gain
18	61548752-64944845	3396093	q22.1 - q22.2	Loss
19	1173033-1426319	253286	p13.3	Gain
19	11723538-12106305	382767	p13.2	Loss
19	50340704-52381739	2041035	q13.32	Gain
19	53504306-54293549	789243	q13.32 - q13.33	Gain
19	54838449-55064096	225647	q13.33	Gain
19	55617858-55704434	86576	q13.33	Gain
20	0-512804	512804	p13	Gain
20	22971225-23206367	235142	p11.21	Loss
20	37240929-38699013	1458084	q12	Loss
21	10013263-10117957	104694	p11.1	Loss
21	10013263-10117957	104694	p11.1	Gain
22	49041686-49401541	359855	q13.33	Gain

bp=base pairs.

Table S3. miRNA target prediction analysis to genes identified on integrative analysis that showed an inverse association between genomic and transcriptomic data.

Gene symbol	TargetScan	PicTar
<i>AICDA</i>	<i>hsa-miR-155</i>	<i>hsa-miR-155</i>
<i>AP2S1</i>	-	<i>hsa-miR-34b, hsa-miR-34c, hsa-miR-34a</i>
<i>ATP5J2</i>	-	-
<i>AZGP1</i>	-	-
<i>BAZ1A</i>	<i>hsa-miR-137</i>	-
<i>BRCA1</i>	<i>hsa-miR-218</i>	<i>hsa-miR-197, hsa-miR-143, hsa-miR-205, hsa-miR-132, hsa-miR-370, hsa-miR-30a-3p, hsa-miR-140, hsa-miR-185, hsa-miR-154</i>
<i>C8orf51</i>	ND	ND
<i>CALM3</i>	<i>hsa-miR-22</i>	<i>hsa-miR-122a, hsa-miR-22, hsa-miR-27b, hsa-miR-27a, hsa-miR-196b, hsa-miR-196a, hsa-miR-29c, hsa-miR-29b, hsa-miR-29a, hsa-miR-320,</i>
<i>CD19</i>	-	-
<i>CDC25C</i>	<i>hsa-miR-767-3p</i>	-
<i>CD93</i>	<i>hsa-miR-216a</i>	-
<i>CHKA</i>	<i>hsa-miR-30e, hsa-miR-30a, hsa-miR-30d, hsa-miR-30b, hsa-miR-30c</i>	-
<i>CORO1A</i>	-	-
<i>CPSF4</i>	<i>hsa-miR-214</i>	<i>hsa-miR-214, hsa-miR-23a, hsa-miR-23b, hsa-let-7c, hsa-let-7g, hsa-let-7b, hsa-let-7f, hsa-let-7i, hsa-let-7a, hsa-miR-98, hsa-let-7e, hsa-let-7d</i>
<i>CYC1</i>	-	-
<i>DDX21</i>	<i>hsa-miR-607</i>	-
<i>EIF4EBP1</i>	<i>hsa-miR-125b, hsa-miR-125a-5p</i>	<i>hsa-miR-125b, hsa-miR-125a</i>
<i>ELMO3</i>	-	-
<i>F12</i>	<i>hsa-miR-330-3p</i>	-
<i>FANCA</i>	-	<i>hsa-miR-26a, hsa-miR-26b</i>
<i>GDPD3</i>	-	-
<i>HKDC1</i>	<i>hsa-miR-876-5p, hsa-miR-1243</i>	-
<i>HRAS</i>	<i>hsa-miR-892a</i>	-
<i>ID11</i>	<i>hsa-miR-570</i>	-
<i>KIF20A</i>	<i>hsa-miR-153</i>	<i>hsa-miR-369, hsa-miR-374</i>
<i>LRRC14</i>	-	<i>hsa-miR-370, hsa-miR-28, hsa-miR-122a</i>
<i>MCM7</i>	<i>hsa-miR-519b-3p, hsa-miR-519a, hsa-miR-519c-3p, hsa-miR-548p</i>	-

<i>MYPN</i>	<i>hsa-miR-214</i>	-
<i>NFKBIL2</i>	-	-
<i>PKP3</i>	-	-
<i>PRELID1</i>	<i>hsa-miR-22</i>	-
<i>PYCRL</i>	-	-
<i>RASSF7</i>	-	-
<i>RECQL4</i>	-	-
<i>RQCD1</i>	ND	ND
<i>SLC1A5</i>	<i>hsa-miR-125a-5p, hsa-miR-125b</i>	<i>hsa-miR-122a, hsa-miR-137</i>
<i>SLC25A22</i>	<i>hsa-miR-613, hsa-miR-1, hsa-miR-206</i>	<i>hsa-miR-296, hsa-miR-1, hsa-miR-206, hsa-miR-337</i>
<i>SNRPD2</i>	-	-
<i>SPIB</i>	<i>hsa-miR-1299, hsa-miR-520a-5p, hsa-miR-525-5p</i>	<i>hsa-miR-146, hsa-miR-218, hsa-miR-328</i>
<i>STAG3</i>	-	-
<i>TALDO1</i>	-	-
<i>TBC1D20</i>	<i>hsa-miR-150</i>	-
<i>TRIB3</i>	<i>hsa-miR-24</i>	-
<i>TUBB3</i>	<i>hsa-miR-429, hsa-miR-200c, hsa-miR-200b</i>	-
<i>VIL1</i>	-	-
<i>ZNF655</i>	<i>hsa-miR-181c, hsa-miR-181a, hsa-miR-181b, hsa-miR-181d</i>	-

TargetScan (<http://www.targetscan.org/>); *PicTar* (http://pictar.mdc-berlin.de/cgi-bin/PicTar_vertebrate.cgi); ND=genes not identified on databases; (-) no miRNA prediction.

Table S4. Genes identified in gene set enrichment analysis.

Gene symbol	N° of enrichment modules	Array CGH	Expression array	Region
<i>CORO1A</i>	8	+	-	16p11.2
<i>FGFR1</i>	8	+	+	8p12
<i>DDX21</i>	8	+	-	10q21.3
<i>DBN1</i>	8	+	+	5q35.3
<i>CENPF</i>	7	-	-	1q41
<i>EIF4EBP1</i>	7	+	-	8p12
<i>NUPR1</i>	6	+	+	16p11.2
<i>COL3A1</i>	5	+	+	2q32.2
<i>AOC3</i>	4	+	+	17q21.31
<i>HRAS</i>	4	+	-	11p15.5
<i>MCM7</i>	3	+	-	7q22.1
<i>IFITM1</i>	3	+	+	11p15.5
<i>SPIB</i>	3	+	-	19q13.33
<i>CD19</i>	3	+	-	16p11.2
<i>SLC1A5</i>	2	+	-	19q13.32
<i>VILI</i>	2	+	-	2q35
<i>CHKA</i>	2	+	-	11q13.2
<i>ID1</i>	2	+	-	10p15.3
<i>TUBB3</i>	2	+	-	16q24.3
<i>CALCRL</i>	2	+	+	2q32.1
<i>STAG3</i>	2	+	-	7q22.1
<i>RECQL4</i>	2	+	-	8q24.3
<i>RHOH</i>	2	-	-	4p14
<i>NOVA2</i>	2	+	+	19q13.32
<i>CDC25C</i>	2	+	-	5q31.2
<i>AP2S1</i>	1	+	-	19q13.32
<i>SNRPD2</i>	1	+	-	19q13.32
<i>MVP</i>	1	+	+	16p11.2
<i>MFAP5</i>	1	+	+	12p13.31
<i>BICD1</i>	1	-	-	12p11.21
<i>DIP2C</i>	1	+	+	10p15.3
<i>FANCA</i>	1	+	-	16q24.3

In bold, genes that showed positive or negative association. Positive (+) and negative (-) signs indicate gene status in relation to genomic gains and losses and up- or down-regulated gene expression, respectively.

Table S5. Ingenuity Pathways Analysis (IPA) networks from top 30 modulators.

Network	Molecules in Network	Score	Focus Molecules	Top Functions
1	↓ AICDA , Akt, ↑ CALCRL , CDC14A, ↓ CDC25C , ↓ CENPF , ↓ CHKA , ↑ COL3A1 , ↓ CORO1A , EDNRA, ↓ EIF4EBP1 , ERK1/2, ↓ F12 , F Actin, ↑ FGFR1 , ↑ GPR4 , GRB7, Histone h3, Igf, ↑ TIGFBP5 , LOX, MSK1/2, NFκB (complex), ↑ NUPR1 , PI3K (complex), PLK3, PNO1, PPM1B, PTGER3, RAMP3, ↓ RHOH , SMOC2, ↑ TNS1 , Vegf, ↓ VIL1	38	16	Cellular Movement, Skeletal and Muscular System Development and Function, Cell Morphology
2	Actin, ACTR3, ATP5A1, ATP5F1, ↓ ATP5J2 , C12orf11, CDC14A, ↑ DBN1 , EIF4G1, FBP1, FLNA, HNF4A, HNRNP1, ↑ HSPB7 , ↓ KIF20A , KPNB1, MAP2K6, MAP3K3, MAPK9, MYB, PNO1, PPP1CA, PRPF4, RIPK3, RPLP1, ↓ RQCD1 , RUVBL2, SPHK1, SUPT5H, TCF12, TRAF6, TRIM37, WNK1, YWHAB, ↓ ZNF655	12	6	Developmental Disorder, Neurological Disease, DNA Replication, Recombination, and Repair
3	↑ DIP2C , PCBD1	3	1	Genetic Disorder, Metabolic Disease, Organ Development
4	↓ PRELID1 , STAT6	3	1	Cell-mediated Immune Response, Cellular Development, Cellular Function and Maintenance
5	Cpsf, ↓ CPSF4 , FIP1L1, MARK3	2	1	RNA Post-Transcriptional Modification, Cancer, Infection Mechanism
6	DLL1, JAG1, JAG2, ↑ MFAP5 , NOTCH1	2	1	Nervous System Development and Function, Tissue Morphology, Gene Expression
7	CDCA3, ↑ CTDSP1 , mir-124, POLR2A, SMAD1, SNAI1	2	1	Embryonic Development, Tissue Development, Cellular Growth and Proliferation
8	ACTN2, ACTN3, ANKRD1, ANKRD23, ↓ MYPN , NEB, NEBL, PPP1CA, TTN	2	1	Cell Morphology, Cellular Assembly and Organization, Cellular Function and Maintenance

In bold, molecules on focus from dataset; arrows above indicate genes with positive association; arrows below indicate genes with negative association.

