

**ASPECTOS MOLECULARES ENVOLVIDOS NO  
SURGIMENTO DO TUMOR TRIPLO-NEGATIVO DE  
MAMA EM PACIENTES PORTADORES OU NÃO DE  
MUTAÇÃO GERMINATIVA EM *BRCA1***

**KIVVI DUARTE DE MELLO NAKAMURA**

**Tese de Doutorado apresentada à Fundação  
Antônio Prudente para obtenção do título de Doutor  
em Ciências**

**Área de concentração: Oncologia**

**Orientador: Dra. Dirce Maria Carraro**

**Co-Orientadora: Dra. Elisa Napolitano e Ferreira**

**São Paulo**

**2019**

## FICHA CATALOGRÁFICA

Preparada pela Biblioteca da Fundação Antônio Prudente

Nakamura, Kivvi Duarte de Mello

**Aspectos moleculares envolvidos no surgimento do tumor triplo-negativo de mama em pacientes portadoras ou não de mutação germinativa em *BRCA1* /**

Kivvi Duarte de Mello Nakamura - São Paulo, 2019.

140p.

Tese (Doutorado)-Fundação Antônio Prudente.

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

Orientadora: Dirce Maria Carraro

Descritores: 1. Neoplasias de Mama Triplo Negativas/genética/Triple Negative Breast Neoplasms/genetics. 2. Mutação em Linhagem Germinativa/Germ-Line Mutation. 3. Genes *BRCA1*/Genes, *BRCA1*. 4. Análise de Sequência de RNA/Sequence Analysis, RNA. 5. Neoplasias da Mama/Breast Neoplasms. 6. Sequenciamento de Nucleotídeos em Larga Escala/High-Throughput Nucleotide Sequencing.

*“Pois o Senhor é quem dá sabedoria;  
de Sua boca procedem o conhecimento e o discernimento”.*

*Provérbios 2:6*

## **DEDICATÓRIA**

Dedico este trabalho aos meus pais, Ribamar e Lílian, que sempre estiveram ao meu lado, me apoiando a trilhar o árduo caminho longe de casa e por me darem toda a força e coragem para enfrentar os grandes desafios da carreira científica.

Dedico também ao meu esposo, Eliézer, e ao meu filho, Eliel, os quais estiveram presentes, me dando todo o suporte e apoio necessários durante todos esses últimos anos, nunca me deixando desistir e sempre me incentivando a buscar meus objetivos.

## AGRADECIMENTOS

Agradeço aos meus pais, Ribamar e Lílian, e à minha irmã, Kimmi, por toda a educação investida, por estarem sempre ao meu lado, por todo o incentivo e força durante esses anos que estive longe, por todo amor e carinho, pelo apoio incondicional, por acreditarem no meu sonho e me ajudarem a torná-lo realidade;

Ao meu amado marido, Eliézer, por todo apoio e força, por me incentivar a nunca desistir, por estar me lembrando dos sonhos e metas, por me ajudar com sua atenção, cuidado, companheirismo, amizade; pelos longos dias e longas noites que parou tudo para cuidar do nosso filho, me permitindo estudar e trabalhar para a conclusão desta etapa;

Ao meu filho Eliel, que embora seja um pequeno bebê, tanto me ajudou com seu jeitinho quieto de ser durante os momentos mais difíceis e desafiadores da minha maternidade; sem você em meu ventre e em meus braços, hoje, seria muito mais difícil a conclusão dessa fase da minha vida;

Aos meus queridos sogros Lia e Orlando, por todo apoio, cuidado e carinho que tiveram comigo aqui em São Paulo; por toda ajuda nos cuidados com Eliel que me permitiu trabalhar na conclusão da escrita tese; serei eternamente grata a vocês;

À minha orientadora, Dra. Dirce Maria Carraro, por toda ajuda científica, pelo apoio profissional; por me ter aberto as portas para a realização deste trabalho e pela confiança e dedicação a esta pesquisa; pelos momentos de discussões científicas que tanto me impulsionaram a crescer profissionalmente;

À minha co-orientadora, Dra. Elisa Napolitano e Ferreira, por toda ajuda científica nos primeiros 2 anos do projeto; pelas discussões e ensinamentos na bancada que também tanto me ajudaram a crescer cientificamente;

À toda equipe do Laboratório de Genômica e Biologia Molecular (LGBM), pelo carinho e amizade; pelos ensinamentos da genômica e por toda ajuda de que precisei durante esses anos; em especial, quero agradecer à minha amiga Isabella Tanus pela amizade e por toda ajuda no laboratório; pela atenção e carinho nos momentos mais difíceis desses anos de doutorado e por ter me ajudado a deixar esta jornada mais leve;

À querida amiga Melissa Pizzi, por toda amizade, força, atenção profissional e emocional; por toda parceria e por ter tornando momentos difíceis em momentos de muitas risadas;

À equipe do Biobanco de Macromoléculas, por todo auxílio e processamento das amostras utilizadas neste estudo;

Aos funcionários da pós-graduação, do CIPE e da biblioteca, em especial a Suely, por toda ajuda na formação da tese;

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), por meio do processo nº 2013/23277-8, pela concessão da bolsa de doutorado.

## RESUMO

Nakamura KDM. **Aspectos moleculares envolvidos no surgimento do Tumor Triplo-Negativo de Mama em pacientes portadoras ou não de mutação germinativa em *BRCA1***. São Paulo; 2019. [Tese de Doutorado-Fundação Antônio Prudente]

O Câncer de Mama Triplo-Negativo (CMTN), caracterizado pela perda de expressão dos receptores de estrógeno (RE), de progesterona (RP) e pela não super-expressão/amplificação do receptor do fator de crescimento epidermal humano do tipo 2 (HER-2), é considerado um subtipo bastante agressivo e heterogêneo molecularmente. Clinicamente, o CMTN apresenta prognóstico ruim, altas taxas de recorrência e menor sobrevida global em relação aos outros subtipos de câncer de mama. Ele corresponde a aproximadamente 15% dos casos de câncer de mama e não apresenta nenhuma terapia alvo efetiva. Mutação patogênica germinativa nos genes *BRCA1* e *BRCA2* leva a um aumento de risco para o desenvolvimento de Câncer de Mama e Ovário, sendo que mutação germinativa em *BRCA1* está associada ao desenvolvimento de CMTN, especialmente em pacientes diagnosticadas antes dos 50 anos. A deficiência de *BRCA1* leva ao mecanismo ineficiente do reparo do DNA e ao desenvolvimento do tumor. Recentemente, nosso grupo classificou as pacientes diagnosticadas com CMTN em grupo Hereditário (com mutação patogênica germinativa em *BRCA1*) e grupo Esporádico (sem mutação germinativa em *BRCA1*). Estes foram ainda classificados em tumores *BRCA1*-deficiente (com hipermetilação no promotor de *BRCA1*) e em tumores *BRCA1*-proficiente (sem hipermetilação no promotor de *BRCA1* e sem mutação germinativa em *BRCA1*). As diferenças moleculares entre esses grupos de CMTN são de grande interesse clínico e biológico, embora ainda não sejam conhecidas. Dentro desse contexto, nosso objetivo foi investigar o perfil transcricional de CMTN sob diferentes aspectos de deficiência de *BRCA1*. Para isso, foi avaliado o sequenciamento do RNA (RNA-Seq) de 37 casos de CMTN em idade jovem ( $\leq 50$  anos), sendo 9 casos Hereditário (com mutação patogênica germinativa em *BRCA1*) e 28 casos Esporádico. Dos casos Esporádicos, 9

foram *BRCA1*-deficiente (com o promotor de *BRCA1* hipermetilado) e 19 *BRCA1*-proficiente (com o promotor de *BRCA1* não hipermetilado). Utilizamos o RNA total a partir das 37 amostras tumorais e 25 amostras normais pareadas adjacentes ao tumor. As bibliotecas de cDNA foram construídas a partir do RNA total através do método de depleção do rRNA e sequenciadas na plataforma *NextSeq* (Illumina). Para a normalização dos dados de expressão de cada gene, foi aplicada a medida FPKM. Os critérios para a obtenção dos genes diferencialmente expressos (GDEs), nas amostras tumorais em relação às normais, foram *fold change*  $\geq 4,0$  e  $\leq -4,0$  e p valor ajustado  $\leq 0,05$ ). Para classificação molecular dos tumores TN, foi utilizada a ferramenta online *TNBCtype*. As curvas de sobrevida global de acordo com essa classificação molecular foram calculadas através do método de Kaplan-Meier. Para verificar os processos biológicos envolvidos com a tumorigênese do CMTN, os GDEs foram submetidos a uma análise funcional *in silico* através do programa *Ingenuity Pathway Analysis* (IPA). Em média, 48.627.204 milhões de sequências foram geradas por amostra, das quais 79,5% foram mapeadas no genoma humano referência, revelando, em média, 15.071 genes expressos com pelo menos 10 sequências única mapeada por amostra. O perfil transcricional das amostras CMTN permitiu a classificação nos 7 subtipos moleculares de CMTN, sendo a maioria das amostras classificadas no subtipo imunomodulador (IM) e mesenquimal (M). No grupo das amostras *BRCA1*-deficiente, observamos um predomínio do subtipo IM enquanto que, nas amostras *BRCA1*-proficiente, houve um maior número de amostras classificadas em M. Não observamos diferença nas curvas de sobrevida global entre os dois grupos de amostras. Porém, os resultados mostraram que o subtipo IM parece ter uma tendência de melhor sobrevida global comparado com os outros subtipos, independente do status de *BRCA1*. A expressão diferencial das amostras tumorais em relação às amostras normais no grupo de CMTN Hereditário revelou 1965 GDEs, sendo 589 mais expressos e 1376 menos expressos no tumor; e, no grupo de CMTN Esporádico, a análise revelou 1837 GDEs, sendo 645 mais expressos e 1192 menos expressos no tumor. A partir dos GDEs de cada grupo, a análise do IPA mostrou que os dois grupos apresentaram vias canônicas enriquecidas comuns significativamente ativadas e envolvidas, de forma geral, com a Regulação do ciclo celular. Entretanto, as vias canônicas significativamente inibidas mostraram-se



exclusivas em cada grupo: a via de Sinalização do receptor do glutamato foi a mais significativa (z-score = -2,12) nas amostras de CMTN Hereditário; e a via de Ativação de LXR/RXR foi a mais significativa (z-score = -2,98) nas amostras CMTN Esporádico. Além disso, observamos reguladores transcricionais significativamente ativados e inibidos exclusivamente em cada grupo. Os genes *TNF* e *E2F1* foram os mais significativamente ativados apenas nas amostras de CMTN Hereditário (z-score = 4,06) e de CMTN Esporádico (z-score = 2,22), respectivamente. Por outro lado, o microRNA mir-21 e o gene *PPARG* foram os reguladores inibidos exclusivamente nas amostras CMTN Hereditário (-score = -4,68) e nas amostras CMTN Esporádico (z-score = -4,24), respectivamente. Esse estudo, de forma geral, revelou potenciais vias e genes reguladores de cascatas de sinalização envolvidos na tumorigênese do CMTN no contexto da deficiência de *BRCA1*, contribuindo na elucidação da complexidade funcional de tumores TN.

**Descritores:** Neoplasias de Mama Triplo Negativas/genética. Mutação em Linhagem Germinativa. Genes BRCA1. Análise de Sequência de RNA. Neoplasias da Mama. Sequenciamento de Nucleotídeos em Larga Escala.

## SUMMARY

Nakamura KDM. [Molecular aspects involved in development of triple-negative breast cancer in patients with or without germline mutation in *BRCA1*]. São Paulo; 2019. [Tese de Doutorado-Fundação Antônio Prudente]

Triple-negative breast cancer (TNBC), characterized by lack of expression of the estrogen receptor (ER), progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2), results in aggressive biology, early peak of recurrence, and shorter overall survival than other subtypes. It comprises approximately 15% of breast cancer cases and yet there is no effective therapy. Mutations in the *BRCA1* and *BRCA2* genes are associated with increased risk of breast and ovarian cancers since germline mutation in *BRCA1* is associated with the development of TNBC, especially in patients diagnosed before age 50. *BRCA* deficiency leads to impaired DNA repair and tumor development. Recently, our group classified TNBC patients into hereditary *BRCA1*-mutated and sporadic *BRCA1*-proficient. These were further classified into *BRCA1*-deficient tumors (with *BRCA1* promoter hypermethylation) and *BRCA1*-proficient tumors (no *BRCA1* promoter hypermethylation neither *BRCA1* germline mutation). Molecular differences between hereditary and sporadic TNBC groups are clinically and biologically interesting although it remains unclear. In this context, we aimed to investigate the transcriptional profile of TNBC-associated or not with *BRCA1* deficiency. For that, RNA sequencing (RNA-seq) from 37 early-onset TNBC ( $\leq 50$  years old) was evaluated, comprising 9 Hereditary (with *BRCA1* germline pathogenic mutation) and 28 Sporadic cases, of which 9 *BRCA1*-deficient and 19 *BRCA1*-proficient. Total RNA from 37 tumors samples and 25 adjacent normal samples of paired cases were used to constructed RNAseq libraries by depleting ribosomal RNA and sequenced on Illumina NextSeq 500 platform. Expression values were normalized by FPKM. Differentially expressed genes (DEGs) between tumor and normal samples were obtained using fold-change  $\geq |4|$  and p value adjusted  $\leq 0.05$  as statistical criteria. The TNBC samples were subtyping using the web-based prediction tool TNBCType. Overall survival curves

were calculated using the Kaplan.Meir method. IPA software was used to detect activated/inactivated canonical pathways and relevant upstream regulators, considering a z-score  $\leq -2.0$  and  $\geq 2.0$ , respectively. On average, 49 million reads were generated per sample, of which 79.5% were mapped to the human reference genome revealing 15,071 expressed genes with at least 10 reads per sample. From the transcriptional profile of TNBC samples we classified into seven TNBC subtypes being the majority of tumors classified as immunomodulatory (IM) and mesenchymal (M) subtype. We detected no difference in overall survival for both groups. However, trends towards better overall were observed for TNBC samples classified as IM compared with other subtypes, without associations with *BRCA1* status. Differential gene expression analysis between tumor and normal samples in the hereditary group revealed 1,965 DEGs, being 589 upregulated and 1,376 downregulated; and in the sporadic group, the analysis revealed 1,837 DEGs, being 645 upregulated and 1,192 downregulated. Using the DEGs of each group, the IPA analysis revealed that Cell Cycle Regulation signaling was activated in both groups. Regarding inactivated pathways, we detected the Glutamate Receptor signaling (z-score = -2,12) in hereditary TNBC and the LXR/RXR activation in sporadic TNBC (z-score = -2,98). Also, the IPA analysis revealed relevant specific transcription regulators of each group. The *TNF* and *E2F1* were the most significantly activated genes in hereditary- and sporadic-TNBC, respectively. On the other hand, the mir-21 and *PPARG* were the most significantly unique inhibited regulators in hereditary and sporadic, respectively. In general, this study unveiled potential pathways and regulatory genes for signaling cascades involved in TNBC tumorigenesis considering the deficiency of *BRCA1*, contributing to the elucidation of the functional complexity of the tumorigenic process of TNBC patients.

**Key-words:** Triple Negative Breast Neoplasms/genetics. Germ-Line Mutation. Genes, BRCA1. Sequence Analysis, RNA. Breast Neoplasms. High-Throughput Nucleotide Sequencing.

## LISTA DE FIGURAS

<b>Figura 1</b>	Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma.....	2
<b>Figura 2</b>	Representação esquemática da função das proteínas envolvidas com o reparo de quebra de dupla fita do DNA através da recombinação homóloga.....	28
<b>Figura 3</b>	Esquema representativo da biblioteca final de cDNA .....	44
<b>Figura 4</b>	Representação esquemática das etapas do sequenciamento na plataforma Illumina.....	47
<b>Figura 5</b>	Fluxograma das etapas do pipeline desenvolvido pela equipe do Laboratório de Biologia Computacional do CIPE .....	49
<b>Figura 6</b>	Perfil da amostra de RNA tumoral ID_M256T analisado pelo <i>Agilent Bioanalyzer 2100</i> .....	55
<b>Figura 7</b>	Perfil da amostra de RNA total ID_M284N analisado pelo <i>Agilent Bioanalyzer 2100</i> .....	56
<b>Figura 8</b>	Perfil da distribuição dos fragmentos de 9 bibliotecas finais de cDNA analisado pelo <i>Agilent Bioanalyzer 2100</i> .....	60

<b>Figura 9</b>	Dados do sequenciamento das 62 bibliotecas finais de cDNA gerados pela plataforma NextSeq 500.....	65
<b>Figura 10</b>	Distribuição dos subtipos moleculares nos quais as 35 amostras de CMTN foram classificadas. ....	71
<b>Figura 11</b>	Curvas de sobrevida global das amostras de CMTN que foram classificadas nos subtipos moleculares .....	72
<b>Figura 12</b>	Curvas de sobrevida global das amostras de CMTN <i>BRCA1</i> -deficiente e CMTN <i>BRCA1</i> -proficiente.....	73
<b>Figura 13</b>	Distribuição da porcentagem de todos os Genes Diferencialmente Expressos (GDEs) no tumor e dos GDEs classificados de acordo com o <i>RefSeq</i> .....	75
<b>Figura 14</b>	Proporção da porcentagem de todos os Genes Diferencialmente Expressos (GDEs) no tumor para cada grupo das amostras de CMTN Hereditário e CMTN Esporádico .....	78
<b>Figura 15</b>	Porcentagem de concordância dos GDEs encontrados em todas as amostras CMTN do estudo (CMTN_Estudo) com os GDEs das amostras CMTN do TCGA.....	80
<b>Figura 16</b>	Porcentagem de concordância dos GDEs encontrados em todas as amostras de CMTN do estudo (CMTN_Estudo) com os GDEs nas amostras de CMTN do TCGA (CMTN_TCGA) por grupo de idade .	81

<b>Figura 17</b>	Vias canônicas enriquecidas em ordem decrescente da significância estatística em todas as amostras CMTN.....	83
<b>Figura 18</b>	Rede de interação biológica 1 identificada a partir dos GDEs em todas as amostras CMTN .....	84
<b>Figura 19</b>	Enriquecimento das vias canônicas distribuídas em ordem decrescente de significância estatística nas amostras de CMTN Hereditário .....	88
<b>Figura 20</b>	Enriquecimento das vias canônicas distribuídas em ordem decrescente de significância estatística nas amostras de CMTN Esporádico.....	88
<b>Figura 21</b>	Rede de interação biológica 1 cujos 35 genes focais das amostras CMTN Hereditário.....	90
<b>Figura 22</b>	Rede de interação biológica 1 cujos 34 genes focais das amostras CMTN Esporádico .....	91

## LISTA DE QUADROS E TABELAS

<b>Quadro 1</b>	Características moleculares e alvos terapêuticos dos subtipos do CMTN .....	18
<b>Quadro 2</b>	Características moleculares dos casos de CMTN selecionados .....	34
<b>Quadro 3</b>	Dados da avaliação anátomo-patológica de todas as amostras tumorais e normais e da extração e quantificação das moléculas de RNA total a partir dos casos de CMTN utilizados no estudo .....	38
<b>Quadro 4</b>	Dados das bibliotecas finais de cDNA construídas a partir do RNA total de todas as amostras tumorais e normais .....	57
<b>Quadro 5</b>	Amostras de CMTN baseada na classificação de Lehmann pela ferramenta <i>TNBCtype</i> .....	69
<b>Quadro 6</b>	Amostras CMTN <i>BRCA1</i> -deficiente baseada na classificação de Lehmann pela ferramenta <i>TNBCtype</i> .....	70
<b>Quadro 7</b>	Amostras de CMTN <i>BRCA1</i> -proficiente baseada na classificação de Lehmann pela ferramenta <i>TNBCtype</i> .....	70
<b>Quadro 8</b>	Dados dos GDEs das amostras de CMTN do TCGA .....	79

<b>Quadro 9</b>	Vias canônicas enriquecidas em todas as amostras de CMTN .....	82
<b>Quadro 10</b>	Redes de interação biológica identificadas a partir dos GDEs em todas as amostras CMTN .....	84
<b>Quadro 11</b>	Reguladores <i>upstream</i> ativados ou inibidos identificados em todos os GDEs quando todas as amostras de CMTN foram comparadas com os seus respectivos normais .....	85
<b>Quadro 12</b>	Vias canônicas enriquecidas nas amostras CMTN Hereditário .....	87
<b>Quadro 13</b>	Vias canônicas enriquecidas nas amostras CMTN Esporádico .....	87
<b>Quadro 14</b>	Redes de interação biológica das amostras CMTN Hereditário .....	89
<b>Quadro 15</b>	Redes de interação biológica das amostras CMTN Esporádico.....	90
<b>Quadro 16</b>	Reguladores <i>upstream</i> ativados ou inibidos identificados nas amostras CMTN Hereditário.....	92
<b>Quadro 17</b>	Reguladores <i>upstream</i> ativados ou inibidos identificados nas amostras CMTN Esporádico.....	93
<b>Tabela 1</b>	Dados do sequenciamento das bibliotecas de cDNA gerado pela plataforma <i>NextSeq 500</i> .....	61



<b>Tabela 2</b>	Dados estatísticos dos 2039 GDEs nas amostras de CMTN pareadas	75
<b>Tabela 3</b>	Dados estatísticos dos 1965 GDEs nas amostras de CMTN Hereditário .....	77
<b>Tabela 4</b>	Dados estatísticos dos 1837 GDEs nas amostras de CMTN Esporádico.....	77

## LISTA DE SIGLAS E ABREVIATURAS

<b>AKT</b>	<i>v-akt murine thymoma viral oncogene homologs</i>
<b>ATM</b>	<i>Ataxia Telangiectasia Mutated serine/threonine kinase</i>
<b>ATR</b>	<i>Ataxia Telangiectasia and Rad3-Related</i>
<b>BARD1</b>	<i>BRCA1 associated RING domain 1</i>
<b>BL1</b>	<i>Basal-like 1</i>
<b>BL2</b>	<i>Basal-like 2</i>
<b>BLIS</b>	<i>Basal-like Imunosupressor</i>
<b>BLIA</b>	<i>Basal-like Imunoativador</i>
<b>BRCA1</b>	<i>Breast cancer 1, early onset</i>
<b>BRCA2</b>	<i>Breast cancer 2, early onset</i>
<b>BRIP1</b>	<i>BRCA1 Interacting Protein C-Terminal Helicase 1</i>
<b>BWA</b>	<i>Burrows Wheeler Aligner</i>
<b>CDH1</b>	<i>Cadherin 1, type 1, E-cadherin (epithelial)</i>
<b>CDKN2A</b>	<i>Cyclin-dependent kinase inhibitor 2A</i>
<b>CEP</b>	<i>Comitê de Ética em Pesquisa</i>
<b>CHEK2</b>	<i>Checkpoint kinase 2</i>
<b>CMTN</b>	<i>Câncer de Mama Triplo-Negativo</i>
<b>CTLA4</b>	<i>Cytotoxic T-Lymphocyte Associated Protein 4</i>
<b>DNA</b>	<i>Deoxyribonucleic acid</i>
<b>EGFR</b>	<i>Epidermal growth factor receptor 1</i>
<b>ER</b>	<i>Estrogen receptor</i>
<b>ERBB2</b>	<i>Avian Erythroblastic Leukemia Viral Oncogene Homolog 2</i>
<b>E2F1</b>	<i>E2F Transcription Factor 1</i>
<b>FDA</b>	<i>Food and Drug Administration</i>
<b>FOXM1</b>	<i>Forkhead Box M1</i>
<b>FPKM</b>	<i>Fragments Per Kilobase Million</i>
<b>GDEs</b>	<i>Genes Diferencialmente Expressos</i>
<b>HBOC</b>	<i>Hereditary Breast and Ovarian Cancer</i>
<b>HER2</b>	<i>Human epidermal growth factor receptor 2</i>
<b>H2AX</b>	<i>H2A Histone Family, Member X</i>
<b>IHQ</b>	<i>imuno-histoquímica</i>
<b>IM</b>	<i>imuno-modulador</i>

<b>INCa</b>	Instituto Nacional do Câncer
<b>INDEL</b>	Alteração de Inserção ou Deleção
<b>INS</b>	Instáveis
<b>IPA</b>	<i>Ingenuity Pathway Analysis</i>
<b>Ki67</b>	<i>Antigen KI-67</i>
<b>LOF</b>	<i>Loss of Function</i>
<b>LRA</b>	Luminal Receptor de Andrógeno
<b>LXR</b>	<i>The liver X receptor</i>
<b>M</b>	Mesenquimal
<b>Mb</b>	Megabase
<b>MES</b>	Mesenquimal
<b>METABRIC</b>	<i>Molecular Taxonomy of Breast Cancer International Consortium</i>
<b>miRNA</b>	MicroRNA
<b>mM</b>	Milimolar
<b>MLPA</b>	<i>Multiplex Ligation-Dependent Probe Amplification</i>
<b>mRNA</b>	RNA mensageiro
<b>MSL</b>	<i>mesenchymal stem-like</i>
<b>mTOR</b>	<i>Mammalian target of rapamycin</i>
<b>N</b>	Tecido Normal
<b>ncRNA</b>	RNA não codificador de proteína
<b>Ng</b>	nanograma
<b>NGS</b>	<i>Next Generation Sequencing</i>
<b>PALB2</b>	<i>Partner and localizer of BRCA2</i>
<b>PARP1</b>	<i>Poly (adenosine diphosphate-ribose) polymerase 1</i>
<b>Pb</b>	Par de base
<b>PCR</b>	<i>Polymerase chain reaction</i>
<b>PDGF</b>	<i>Platelet-derived growth factor</i>
<b>PD1</b>	<i>Programmed Cell Death Protein 1</i>
<b>PD-L1</b>	<i>Programmed Cell Death Ligant 1</i>
<b>PI3K</b>	<i>phosphatidylinositol 3-kinase</i>
<b>PI3KCA</b>	<i>Phosphatidylinositol-4, 5-bisphosphate 3-kinase catalytic subunit</i>
<i>alpha</i>	<i>isoform</i>
<b>pM</b>	Picomolar
<b>PPARG</b>	<i>Peroxisome Proliferator-Activated Receptor Gamma</i>
<b>PTEN</b>	<i>Phosphatase and tensin homolog</i>

<b>QDF</b>	Quebra de Dupla Fita
<b>pCR</b>	<i>Complete Pathological Response</i>
<b>qPCR</b>	<i>Quantitative Polymerase Chain Reaction</i>
<b>RAD50</b>	<i>RAD50 homolog (S. cerevisiae)</i>
<b>RAD51</b>	RAD51 recombinase
<b>RAD51B</b>	<i>RAD51 Paralog B</i>
<b>RB1</b>	Retinoblastoma 1
<b>R</b>	Receptor de Estrógeno
<b>RH</b>	Recombinação Homóloga
<b>RHs</b>	Receptores Hormonais
<b>RIN</b>	<i>RNA Integrity Number</i>
<b>RP</b>	Receptor de Progesterona
<b>RNA</b>	<i>Ribonucleic acid</i>
<b>RNASeq</b>	Sequenciamento massivo em paralelo do RNA
<b>rRNA</b>	RNA ribossomal
<b>RXR</b>	<i>The retinoid X receptor</i>
<b>STAT3</b>	<i>Signal Transducer and Activator of Transcription 3</i>
<b>STK1</b>	<i>Ser/Thr kinase 1</i>
<b>T</b>	Tecido Tumoral
<b>TAMs</b>	<i>Tumor Associated Macrophages</i>
<b>TCGA</b>	<i>The cancer genome Atlas</i>
<b>TEM</b>	Transição epitélio-mesênquima
<b>TGF-β</b>	<i>Transforming growth factor beta</i>
<b>TILs</b>	<i>Tumor-Infiltrating Lymphocytes</i>
<b>TN</b>	Triplo-negativo
<b>TNF</b>	<i>Tumor Necrosis Factor</i>
<b>TP53</b>	<i>Tumor Protein p53</i>
<b>uL</b>	Microlitro
<b>VEGF</b>	<i>Vascular endothelial growth factor</i>
<b>VTCN1</b>	<i>V-Set Domain Containing T Cell Activation Inhibitor 1</i>

# ÍNDICE

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>1</b>
1.1	Câncer de Mama .....	1
1.1.1	Incidência e mortalidade .....	1
1.1.2	Fatores de risco .....	2
1.1.3	Classificação molecular .....	5
1.2	Câncer de Mama Triplo-Negativo (CMTN).....	8
1.2.1	Aspectos clínicos.....	8
1.2.2	Caracterização Molecular .....	15
1.2.3	Caracterização Genética.....	23
1.3	Gene <i>BRCA1</i> e a Deficiência na Recombinação Homóloga.....	26
<b>2</b>	<b>OBJETIVOS</b> .....	<b>31</b>
2.1	Objetivo Geral.....	31
2.2	Objetivos Específicos.....	31
<b>3</b>	<b>MATERIAIS E MÉTODOS</b> .....	<b>32</b>
3.1	Aspectos éticos.....	32
3.2	Casuística .....	32
3.3	Extração de RNA .....	36
3.4	Preparo das bibliotecas de cDNA .....	41
3.5	Sequenciamento do transcriptoma completo (RNA-Seq).....	45
3.6	Análise de Bioinformática .....	47
3.7	Classificação molecular das amostras de CMTN Baseada na Classificação de Lehmann.....	50
3.8	Análise funcional <i>in silico</i> dos Genes Diferencialmente Expressos (GDEs)..	52
<b>4</b>	<b>RESULTADOS</b> .....	<b>54</b>
4.1	Seleção das amostras e extração de RNA .....	54
4.2	Sequenciamento do RNA.....	54

4.2.1	Construção das bibliotecas de cDNA.....	54
4.2.2	Sequenciamento das bibliotecas de cDNA na Plataforma <i>NextSeq</i> 500.....	60
4.3	Classificação molecular das amostras de CMTN .....	65
4.4	Análise dos Genes Diferencialmente Expressos.....	74
4.5	Análise funcional <i>in silico</i> através do IPA.....	82
<b>5</b>	<b>DISCUSSÃO.....</b>	<b>94</b>
<b>6</b>	<b>CONCLUSÃO.....</b>	<b>107</b>
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>109</b>

## **ANEXO**

**Anexo 1** Carta de aprovação do Comitê de Ética em Pesquisa-CEP

# **1 INTRODUÇÃO**

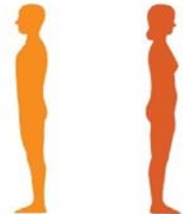
## **1.1 CÂNCER DE MAMA**

### **1.1.1 Incidência e mortalidade**

O câncer de mama é o segundo tipo de tumor mais frequentemente diagnosticado depois do câncer de pulmão, correspondendo a 11,6% de todos os novos casos do mundo. Ele está na quarta posição de causa de morte e representa 6,6% de todas as mortes por câncer no mundo. Entre as mulheres, o câncer de mama é o mais comumente diagnosticado e a principal causa de morte por câncer (BRAY et al. 2018). Enquanto que nos países mais desenvolvidos o câncer de mama é a segunda causa de morte, depois do câncer de pulmão, nos países menos desenvolvidos é a causa de morte mais frequente entre as mulheres (BRAY et al. 2018).

No Brasil, o câncer de mama é o tipo de tumor mais frequente entre as mulheres, depois do câncer de pele não melanoma, correspondendo por cerca de 29,5% dos novos casos de câncer a cada ano. De acordo com as estimativas do Instituto Nacional do Câncer - INCa, são estimados 59.700 casos novos de câncer de mama para o biênio 2018-2019, com um risco estimado de 56,33 casos a cada 100 mil mulheres (Ministério da Saúde 2018) (Figura 1). Em todas as regiões do Brasil, ele é o mais incidente, exceto na região Norte, onde o câncer do colo do útero está na primeira posição. E, apesar do avanço no diagnóstico precoce e no tratamento, as taxas de mortalidade e de sobrevida média para esta neoplasia continuam elevadas na

população feminina brasileira (Ministério da Saúde 2016). Porém, estudos mostram que o câncer de mama é uma neoplasia que apresenta um bom prognóstico quando analisada a sobrevida média das pacientes nos primeiros cinco anos pós-diagnóstico (PARKIN 2001; Ministério da Saúde 2018).

Localização Primária	Casos	%			Localização Primária	Casos	%
Próstata	68.220	31,7%	<b>Homens</b> 	<b>Mulheres</b>	Mama Feminina	59.700	29,5%
Traqueia, Brônquio e Pulmão	18.740	8,7%			Cólon e Reto	18.980	9,4%
Cólon e Reto	17.380	8,1%			Colo do Útero	16.370	8,1%
Estômago	13.540	6,3%			Traqueia, Brônquio e Pulmão	12.530	6,2%
Cavidade Oral	11.200	5,2%			Glândula Tireoide	8.040	4,0%
Esôfago	8.240	3,8%			Estômago	7.750	3,8%
Bexiga	6.690	3,1%			Corpo do Útero	6.600	3,3%
Laringe	6.390	3,0%			Ovário	6.150	3,0%
Leucemias	5.940	2,8%			Sistema Nervoso Central	5.510	2,7%
Sistema Nervoso Central	5.810	2,7%			Leucemias	4.860	2,4%

**Fonte:** Ministério da Saúde (2018).

**Figura 1** - Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma.

### 1.1.2 Fatores de risco

Os fatores de risco para câncer de mama, assim como para outros tipos de câncer, têm crescido nos últimos anos devido, em sua grande maioria, às mudanças no estilo de vida da população brasileira, tais como alimentação hipercalórica, consumo de álcool e tabaco, sedentarismo, obesidade e exposição à radiação ionizante. Diversos outros fatores também estão associados com o aumento de risco para desenvolver o câncer de mama, tais como: idade da primeira menstruação antes dos 12 anos; uso de alguns anticoncepcionais e da terapia de reposição hormonal na menopausa; mulheres que nunca engravidaram ou nunca tiveram filhos (nuliparidade); menopausa tardia, após os 55 anos; e predisposição genética e hereditária, onde o histórico familiar é de grande importância para o aumento de



risco (STEWART e WILD 2014; American Cancer Society-ACS 2017; Ministério da Saúde 2018).

Estima-se que cerca de 5-10% dos cânceres de mama são causados, principalmente, por alterações genéticas nas linhagens germinativas, as quais são transmitidas ao longo das gerações e que caracterizam as síndromes hereditárias (PHAROAH et al. 2002). A principal síndrome hereditária associada ao câncer de mama é a Síndrome de Câncer de Mama e Ovário Hereditário (HBOC, do inglês *Hereditary Breast and Ovarian Cancer*), a qual representa cerca de 10% dos casos de câncer de mama e ovário. Os principais genes para essa síndrome são os genes de alta penetrância *BRCA1* (do inglês *breast cancer 1, early onset*) e *BRCA2* (do inglês *breast cancer 2, early onset*) (MIKI et al. 1994; WOOSTER et al. 1995; ANTONIOU et al. 2003).

A mutação patogênica germinativa em *BRCA1* está associada a uma maior predisposição para o desenvolvimento de câncer de mama, conferindo um risco de 85%, enquanto que mutação patogênica germinativa em *BRCA2* aumenta o risco de desenvolvimento do câncer de mama de 45% em relação à população geral que não apresenta nenhuma alteração genética (ANTONIOU et al. 2003; CHEN e PARMIGIANI 2007). O mecanismo que leva as portadoras de mutação germinativa em heterozigose, com perda de função da proteína (do inglês *LOF - Loss of function mutation*), a terem um maior risco de desenvolver tumores na mama e no ovário ainda não está completamente esclarecido. Uma das hipóteses que explica essa alta associação é que a função de reparo da quebra de fita dupla por recombinação homóloga nos tecidos da mama e ovário é extremamente importante durante o estresse oxidativo gerado a cada ciclo menstrual. Assim, esses tecidos tornam-se

sensíveis ao dano causado a cada ciclo menstrual, estando propensos à carcinogênese já que o reparo do DNA está ineficiente devido à perda de função das proteínas BRCA1/BRCA2 (ROY et al. 2012; YU e SHAO 2012). Além disso, tumores associados com mutações nos genes *BRCA1/2* parecem estar restritos a malignidades epiteliais (PETRUCELLI et al. 2010).

Algumas diferenças são observadas nos tumores que carregam mutação germinativa em *BRCA1* daqueles com mutações em *BRCA2*. Tumores que apresentam mutação em *BRCA1* acometem mulheres a desenvolverem câncer de mama em idade mais jovem que os tumores com mutação em *BRCA2* (BREKELMANS et al. 2007). Uma outra característica importante dos tumores que carregam mutação germinativa em *BRCA1* é que, em sua grande maioria, são tumores com baixa ou ausente expressão protéica do receptor de estrógeno (RE) (LAKHANI et al. 2005; FACKENTHAL e OLOPADE 2007; MAVADDAT et al. 2013), estando, por isso, associado a um risco aumentado para o desenvolvimento de Câncer de Mama Triplo Negativo (CMTN) (60 - 80%), o qual não expressa RE e nem o receptor de progesterona (RP) (FACKENTHAL e OLOPADE 2007; MAVADDAT et al. 2013). Diferentemente, a maioria dos tumores associados com mutação germinativa em *BRCA2* apresentam expressão protéica positiva para o RE (SØRLIE et al. 2003; MAVADDAT et al. 2013).

Mutações germinativas em *BRCA1* também predispõem os indivíduos ao aumento de risco para desenvolver outros tipos de câncer, como o melanoma. Já mutações em *BRCA2* foram observadas por estarem associadas a uma maior predisposição de desenvolver câncer de pâncreas, próstata e de mama em homens (THOMPSON et al. 2002; GALLAGHER et al. 2010; MERSCH et al. 2015).

Além dos genes *BRCA1* e *BRCA2*, mutações em outros genes de alta penetrância, tais como *CDH1*, *PTEN*, *STK1* e *TP53*, também podem elevar o risco para o desenvolvimento de síndromes de susceptibilidade ao câncer de mama (MAXWELL e DOMCHECK 2013). Além disso, a suscetibilidade ao câncer de mama também está associada às mutações em genes de penetrância moderada como *ATM*, *CHEK2* e *PALB2* (ANTONIOU et al. 2014).

### **1.1.3 Classificação molecular**

O câncer de mama é uma doença que tem sido descrita por ser bastante heterogênea a nível molecular, apresentando diferentes subtipos caracterizados por distintos perfis de expressão gênica, os quais implicam diretamente no prognóstico e na sobrevida das pacientes (PEROU et al. 2000). Os tumores de mama são, portanto, classificados de acordo com o perfil de expressão imunohistoquímica de marcadores moleculares e de expressão gênica. Além disso, a taxa de proliferação das células tumorais, através da expressão dos marcadores de proliferação celular, como do Ki67, bem como variáveis clínico-patológicas, incluindo tamanho, tipo e grau do tumor, também são usadas para classificar os tumores de mama e auxiliar nas propostas terapêuticas (BERNHARDT et al. 2016). Estudos clássicos realizados por PEROU et al. (2000) e SØRLIE et al. (2001), com base na análise de expressão gênica, caracterizaram o câncer de mama em cinco principais padrões moleculares com diferentes implicações clínicas: Luminal A, Luminal B, HER-2 positivo, *Basal-like* e *Normal-like*. Entre cada um desses subtipos há diferenças significativas na sobrevida geral e livre de doença.

Os dois primeiros subtipos, como o nome sugere, tem maior expressão de genes característicos do tecido luminal além de apresentarem marcação positiva dos receptores hormonais de estrógeno (RE) e progesterona (RP). O luminal A é caracterizado pela alta expressão dos RE e RP e das citoqueratinas luminiais 8 e 18 assim como ausência de expressão do gene *ERBB2* que codifica o receptor do fator de crescimento epidermal humano tipo 2 (HER-2, do inglês *Human Epidermal growth factor Receptor-type 2*). É o subtipo mais comum, representando 50-60% de todos os casos de câncer de mama. Pacientes diagnosticadas com esse subtipo tem um bom prognóstico, apresentam maior sobrevida geral e livre de doença comparada com os outros subtipos. O tratamento é baseado principalmente em terapias hormonais com adição de quimioterapia, dependendo do estadio clínico. Já os tumores Luminal B são caracterizados pela baixa expressão dos RE e RP, alta positividade para o Ki67, e apresentam um alto grau histológico comparado com os tumores Luminal A. Eles compreendem 15-20% de todos os casos de câncer de mama e respondem bem à quimioterapia, alcançando resposta patológica completa (pCR - definida como doença invasiva não residual na mama e nos linfonodos) em 16% dos tumores comparado com 6% dos tumores Luminal A. Os tumores HER-2 positivo, *Basal-like* e *Normal-like*, apresentam baixa ou nenhuma expressão dos RE e RP. Os tumores HER-2 positivo apresentam aumento de expressão do HER2 e, assim, o tratamento é feito com o uso do anticorpo monoclonal contra o HER2, o trastuzumabe, junto com a quimioterapia. Eles representam 15-20% de todos os casos de câncer de mama e têm um bom prognóstico e boa sobrevida geral. Os tumores *Basal-like* apresentam alta expressão das citoqueratinas basais e contabilizam 5-10% de todos os casos de câncer de mama. Os *basal-like* não

apresentam positividade para RE e RP e são negativos para aumento de expressão ou amplificação de *ERBB2* (HER2), o que, por sua vez, são frequentemente associados com os tumores triplo-negativos (TN). Embora os termos *basal-like* e TN não sejam sinônimos, cerca de 80% dos tumores TN são *basal-like* assim como a maioria dos *basal-like* são TN (BERTUCCI et al. 2008; FOULKES et al. 2010). Além disso, os tumores *Basal-like* são associados com um comportamento clínico agressivo e maiores taxas de metástase. Pacientes diagnosticadas com tumores *Basal-like* têm um pior prognóstico e menor sobrevida geral e livre de doença comparado com os outros subtipos. E o subtipo *normal-like* apresenta elevada expressão de genes característicos de tecido adiposo e de outros tipos celulares não epiteliais (PEROU et al. 2000; SØRLIE et al. 2001, 2003).

Além da classificação baseada no perfil imunofenotípico dos marcadores RE, RP, HER-2 e Ki67, em 2011, o Consenso Internacional de Especialistas de Saint Gallen acrescentou a análise de amplificação do gene que codifica para HER-2, adotando-se as deliberações sobre as estratégias terapêuticas. Assim, os tumores de mama foram reclassificados em luminal A/B, hiperexpressão de HER2 e TN (GOLDHIRSCH et al. 2011). De maneira geral, são oferecidos tratamentos com hormonioterapia para os tumores luminais, anticorpo monoclonal anti-HER-2 (trastuzumab) para os tumores que expressavam HER-2 e quimioterapia para os triplo-negativo.

Além desses subtipos, HERSCHKOWITZ et al. (2007) propuseram um outro subtipo, *claudin-low*, caracterizado por uma baixa expressão de genes envolvidos com proteínas de adesão célula-célula, como as claudinas, ou até ausência da expressão de marcadores de diferenciação luminal e moderada expressão de genes de

proliferação. São tumores com alta expressão de marcadores da transição epitélio-mesenquima, marcadores endoteliais e linfocíticos e de células tronco tumorais. É um subtipo que apresenta um padrão imunofenotípico triplo-negativo não basalóide, mas que ainda não possui marcadores imunohistoquímicos protocolados para a sua caracterização e nenhuma terapia específica (HERSCHKOWITZ et al. 2007; PRAT et al. 2010).

## **1.2 CÂNCER DE MAMA TRIPLO-NEGATIVO (CMTN)**

### **1.2.1 Aspectos clínicos**

O CMTN é caracterizado pela ausência de expressão dos receptores hormonais RE e RP através da análise de imunohistoquímica, especificamente com < 1% de expressão nos núcleos das células tumorais, e por não apresentar aumento de expressão do receptor HER2, especificamente com positividade 1+ ou 2+ pela imuno-histoquímica, e/ou não amplificação do gene *ERBB2* pela Hibridação Fluorescente *In Situ* (FISH, do inglês *fluorescence in situ hybridization*) (CAREY et al. 2006; DENT et al. 2007; BIANCHINI et al. 2010). Ele compreende 10-20% de todos os casos de câncer de mama e são tumores maiores em tamanho, de alto grau histológico, biologicamente mais agressivos que os outros subtipos de câncer de mama e comumente acometem mulheres jovens com idade inferior a 50 anos e de origem afro-americana (BAUER et al. 2007; DENT et al. 2007; CAREY et al. 2007; STURTZ et al. 2014). Comparado com outros subtipos de câncer de mama, o CMTN está associado ao risco significativamente aumentado de recaída e morte e piores taxas de sobrevida global nos primeiros 5 anos de seguimento (LI et al. 2017).

Metástases à distância, preferencialmente para órgãos sólidos como pulmão e cérebro, também tendem a ocorrer dentro dos 3 a 5 primeiros anos após o diagnóstico (CAREY et al. 2006; BAUER et al. 2007; LIEDTKE et al. 2008). A média de sobrevida global das pacientes com CMTN metastático é bem curta, de 12 a 18 meses (DENT et al. 2009). Razões para esse desfavorável prognóstico incluem a natureza heterogênea e agressiva da doença bem como ausência de terapias alvo bem definidas, já que são tumores que não expressam os alvos terapêuticos convencionais, como os receptores hormonais (RE e RP) e HER-2, não sendo elegíveis para tratamento com terapias hormonais e anti-HER2, respectivamente, que são terapias bastante efetivas para os demais subtipos de câncer de mama. Assim, as opções de tratamento tornam-se, de forma geral, limitadas à quimioterapia convencional sistêmica na neoadjuvância, adjuvância e na doença metastática (BAUER et al. 2007; CAREY et al. 2007, 2012; LEHMANN et al. 2011).

Apesar do comportamento clínico bastante agressivo do CMTN, cerca de 30% a 40% dos pacientes atingem pCR sem evidência de doença no momento da cirurgia após quimioterapia neoadjuvante com altas taxas de sobrevida (LIEDTKE et al. 2008; GLÜCK et al. 2012). Entretanto, pacientes que apresentam doença residual após a neoadjuvância tem seis vezes mais chances de ter recorrência e doze vezes mais probabilidade de morte devido à doença metastática (LIEDTKE et al. 2008; ESSERMAN et al. 2012; MASUDA et al. 2013).

A quimioterapia baseada em antraciclina (doxorubicina, epirubicina), acrescida de taxanos (docetaxel, paclitaxel), foi evidenciada por conferir um maior benefício às pacientes com CMTN quando comparado com aquelas não TN, sendo, portanto, usada, rotineiramente, no contexto da quimioterapia neoadjuvante

(HAYES et al. 2007; OAKMAN et al. 2010; VON MINCKWITZ et al. 2012). Nesse sentido, os estudos mostram que as taxas de pCR, depois da quimioterapia neoadjuvante, são bem maiores nas pacientes com tumores TN com excelente sobrevida quando comparado com aquelas que não são TN. Contudo, mais da metade das pacientes TN não alcançam pCR e apresentam ruim sobrevida, principalmente nos primeiros 3 anos após o diagnóstico (LIEDTKE et al. 2008; DENT et al. 2009; VON MINCKWITZ et al. 2012; HATZIS et al. 2016). Com base nesses dados, houve um maior interesse das pesquisas para o desenvolvimento de efetivas terapias alvo para os tumores TN (SILVER et al. 2010; VON MINCKWITZ e MARTIN 2012; BASELGA et al. 2013).

O uso dos agentes de platinas, como carboplatina e cisplatina, que induzem dano ao DNA, tem se mostrado eficiente para diversos cânceres. Um estudo clínico randomizado de fase II evidenciou um aumento de 21% na pCR entre as pacientes TN que receberam a quimioterapia neoadjuvante baseada em paclitaxel e doxorubicina com a adição de carboplatina (VON MINCKWITZ et al. 2013). Recentemente, outro estudo clínico também demonstrou um aumento de 16% na taxa da pCR para as pacientes com TN quando submetidas ao tratamento quimioterápico baseado em doxorubicina, ciclofosfamida e paclitaxel acrescido com carboplatina (SIKOV et al. 2015).

O aumento na pCR causada pela carboplatina tem sido observado ainda maior em pacientes com mutação germinativa ou somática nos genes *BRCA1* ou *BRCA2* (BYRSKI et al. 2012, 2014; VON MINCKWITZ et al. 2014). Um recente estudo clínico de fase III mostrou um efeito positivo da carboplatina em pacientes de CMTN com mutação germinativa em *BRCA1*, porém o mesmo não foi observado em



pacientes com deficiência no reparo do DNA devido à metilação em *BRCA1* e ao baixo nível de expressão de *BRCA1* (TUTT et al. 2018). Células com mutações nos genes *BRCA1/2* apresentam mecanismo de reparo de quebra de fita dupla do DNA deficiente, o que, por sua vez, as tornam sensíveis aos agentes de platina (ROTTENBERG et al. 2007). Por exemplo, estudos têm mostrado que pacientes TN portadoras de mutação em *BRCA1* alcançam altas taxas de pCR após quimioterapia neoadjuvante combinada com cisplatina (BYRSKI et al. 2010). Esses resultados alertam para a importância de melhor investigar pacientes TN no contexto da deficiência do mecanismo de reparo do DNA para melhor manejo clínico dessas pacientes. Embora o uso dos agentes de platina tenha sido associado com qCR nas pacientes com CMTN no contexto da mutação em *BRCA1*, esses estudos mostram sua alta toxicidade hematológica e neurológica além de ser incerta a eficácia desse tratamento nesse subgrupo de pacientes com CMTN (BYRSKI et al. 2014; CLIFTON et al. 2018; CAMELO et al. 2019). Considerando isso, outros estudos têm focado na identificação de alvos moleculares para um tratamento mais individualizado, o que permitirá uma abordagem terapêutica personalizada e menos citotóxica.

A terapia combinada com anticorpos monoclonais está sendo objeto de grande interesse. O bevacizumab, o inibidor do fator de crescimento vascular endotelial (VEGF, do inglês *vascular endothelial growth factor*), está sendo estudado para o tratamento do CMTN, uma vez que o VEGF tem sido altamente correlacionado com a progressão tumoral, invasão e angiogênese nesses tumores (SU et al. 2016). Os estudos mostram que a neoadjuvância baseada em docetaxel e antraciclina, combinada com o bevacizumab, significativamente aumenta a pCR nas

pacientes com CMTN (BEAR et al. 2011; GERBER et al. 2013; SIKOV et al. 2015). Em contrapartida, VON MINCKWITZ et al. (2013) reportaram que esse aumento na taxa de pCR foi observado significativamente nas pacientes positivas para os receptores hormonais (RHs) e não para as pacientes TN. Logo, os resultados avaliando o papel dessa terapia combinada nos tumores TN ainda são incertos.

Outro alvo molecular, que tem sido foco de estudos clínicos para tratamento em CMTN, é o receptor do fator de crescimento epidérmico (EGFR, do inglês *epidermal growth factor receptor*) (NAKAI et al. 2016). O EGFR é amplificado em 2% dos cânceres de mama (The Cancer Genome Atlas Network 2012) e sua expressão aumentada é frequentemente observada no subtipo *basal-like* do câncer de mama (CAREY et al. 2012; BASELGA et al. 2013). Em CMTN, PARK et al. (2014) reportaram que, embora a mutação em EGFR seja um evento raro, alteração em número de cópia em EGFR foi encontrada como um fator prognóstico independente para uma ruim sobrevida livre de doença. Contudo, os resultados ainda não mostram significativa melhora das pacientes tratadas com o seu inibidor, o Cetuximab, em combinação com agentes de platina (CAREY et al. 2012; BASELGA et al. 2013).

Recentemente, várias outras terapias alvo têm sido emergentes e estão sendo abordadas como modalidade terapêutica para CMTN, uma vez que são alvos de cascatas de sinalização ativas nesse subgrupo de câncer de mama (PAPADIMITRIOU et al. 2017). Esses incluem agentes alvo das vias de sinalização de Notch (LOCATELLI e CURIGLIANO 2017), da sinalização de Wnt/ $\beta$ -catenina (POHL et al. 2017), da via Hedhehog (HABIB e O'SHAUGHNESSY 2016), da via de PI3K/Akt/mTOR (MASSIHANIA et al. 2016), da polimerase poli ADP-ribose 1 (PARP1, do inglês *poly adenosine diphosphate-ribose polymerase*)

(ANDREOPOULOU et al. 2017; SZÉKELY et al. 2017; GEENEN et al. 2018) e do receptor de andrógeno (RA) (AYCA e TRAINA 2010; RAMPURWALA et al. 2016).

Dentre esses alvos para modalidade terapêutica, o crescente número das pesquisas clínicas está sendo focado na enzima polimerase poli ADP-ribose 1 (PARP) para o tratamento do CMTN com mutação nos genes *BRCA1/2*, uma vez que essa enzima atua no reparo ao dano simples-fita do DNA. No contexto de tumores com mutação em *BRCA1* (80% dos casos de CMTN), que exibem deficiência no sistema de reparo dupla-fita do DNA, os inibidores de PARP têm a principal função de levar a letalidade sintética (FARMER et al. 2005; BRYANT e HELLEDAY 2006). A inibição de PARP leva ao acúmulo de quebras de fita simples do DNA, que, durante o processo de replicação do DNA, levam à formação de quebras de dupla fita do DNA, as quais, por sua vez, não podem ser reparadas com precisão nos tumores com recombinação homóloga deficiente devido à proteína *BRCA1* não funcional (ASHWORTH 2008). Portanto, nos tumores TN com mutação em *BRCA1/2* que perdem a capacidade do reparo dupla-fita do DNA por recombinação homóloga, os inibidores de PARP têm sido avaliados como uma promissora terapia alvo de baixa toxicidade (DEDES et al. 2011; PAPADIMITRIOU et al. 2017; GEENEN et al. 2018). Como observado por GELMON et al. (2011) e por PAHUJA et al. (2014), os inibidores de PARP mostraram taxas objetivas de resposta ao tumor entre 10% e 30% no CMTN metastático com mutação germinativa em *BRCA1*, com menor ou nenhuma atividade no CMTN sem mutação em *BRCA1* (GELMON et al. 2011; PAHUJA et al. 2014). Vários são os inibidores de PARP que estão sendo avaliados. Dentre eles, o Olaparibe tem apresentado maior benefício com taxas de

resposta ao tumor em 54% das pacientes com CMTN *BRCA1*-mutado (TUTT et al. 2009). Estudos clínicos de fase I observaram que o Olaparibe teve efeito antitumoral em pacientes com CMTN associado com mutação em *BRCA1* (FONG et al. 2009; TUTT et al. 2010). Esse inibidor também tem sido combinado com agentes de platina (cisplatina ou carboplatina) em CMTN com mutação em *BRCA1*, induzindo uma elevada taxa de resposta global de 88% (DENT et al. 2013; BALMAÑA et al. 2014; LEE et al. 2014). Dados *in vitro* já mostraram melhores respostas ao tratamento com agente de platina na presença de mutação em *BRCA1/2* (ROTTENBERG et al. 2007; TASSONE et al. 2009), o que é corroborado pelos estudos clínicos onde pacientes portadoras de mutação em *BRCA* exibem melhor sobrevida livre de progressão com o tratamento combinado de inibidor de PARP e componente de platina (HUANG et al. 2009; SILVER et al. 2010; HU et al. 2015; MYLAVARAPU et al. 2018). Com todos esses dados, recentemente, o Olaparibe foi aprovado pela agência regulatória norte-americana FDA e, no Brasil, pela Agência Nacional de Vigilância Sanitária (ANVISA) como tratamento para pacientes com câncer de ovário seroso de alto grau recidivado com mutação nos genes *BRCA1/2* e com doença sensível à platina, após publicação do estudo randomizado de fase II que mostrou que terapia com Olaparibe prolonga de forma significativa a sobrevida livre de progressão (LEDERMANN et al. 2014). Ele também foi aprovado, mais recentemente, para o tratamento de CMTN em pacientes com mutação germinativa em *BRCA1/2*, que já foram previamente submetidas à quimioterapia neoadjuvante, adjuvante ou no cenário metastático. Essa aprovação foi baseada no estudo de fase III que avaliou 302 pacientes com câncer de mama HER2-negativo, portadoras de mutação germinativa em *BRCA1/2*, tratadas com quimioterapia ou com Olaparibe.

Os autores demonstraram uma redução estatisticamente significativa no risco de progressão de doença ou morte, maior taxa de resposta, menores efeitos adversos e menores taxas de descontinuação do tratamento por toxicidade no grupo das pacientes que receberam o tratamento com Olaparibe em comparação às pacientes tratadas apenas com a quimioterapia (ROBSON et al. 2017).

Entretanto, algumas pacientes portadoras de mutação germinativa ou somática nos genes *BRCA1/2*, tratadas com o inibidor de PARP1, podem adquirir resistência ao tratamento pois os tumores podem reverter a mutação e apresentarem uma mutação secundária nos genes *BRCA1/2* capaz de restaurar suas funções e, conseqüentemente, gerar resistência a essa terapia (EDWARDS et al. 2008; DHILLON et al. 2011; BARBER et al. 2013). Essa resistência tem sido reportada em pacientes com câncer de ovário avançado que desenvolveram progressão da doença durante o tratamento com inibidor de PARP (SWISHER et al. 2008; SAKAI et al. 2009; NORQUIST et al. 2011).

De forma geral, a aprovação do Olaparibe, mesmo que em cenários mais restritivos, abre expectativas positivas para a utilização de inibidores de PARP em pacientes que apresentam resistência ao tratamento convencional.

### **1.2.2 Caracterização molecular**

O CMTN é um subgrupo de câncer de mama que apresenta uma grande heterogeneidade biológica e clínica. Vários grupos de pesquisas motivados pela necessidade de contribuir com a caracterização molecular do CMTN, dada a grande heterogeneidade molecular, usaram várias ferramentas como imunohistoquímica, análise global de expressão gênica e de metilação, entre outras.

Usando análises de expressão gênica, LEHMANN et al. (2011) estabeleceram assinaturas gênicas e ontologias a partir de 587 tumores TN e identificaram 6 distintos subtipos moleculares: dois subtipos *basal-like* (*BL1* e *BL2*), um imunomodulador (IM), um mesenquimal (M), um mesenquimal *stem-like* (*MSL*) e um luminal receptor de andrógeno (LRA). O subtipo *BL1* foi caracterizado pela alta expressão de genes envolvidos com a regulação do ciclo celular, proliferação celular e resposta a dano ao DNA. Essa característica explica como esses tumores apresentam uma alta pCR quando tratados com quimioterapia baseada em taxanos (BAUER et al. 2010; JUUL et al. 2010; LEHMANN et al. 2011). O subtipo *BL2* apresentou ontologia gênica enriquecida para vias de sinalização ativadas por fatores de crescimento, como a via do EGFR e da Wnt/ $\beta$ -catenina. Além disso, apresentou alta expressão de marcadores mioepiteliais e basais (LEHMANN et al. 2011). O subtipo IM foi caracterizado pelo enriquecimento de processos biológicos da resposta imune celular, tais como via de sinalização das células do sistema imune, vias de sinalização de citocinas e processamento e apresentação de antígenos. Já os subtipos M e *MSL* compartilhavam elevada expressão de genes envolvidos com motilidade e diferenciação celular e vias de sinalização ativadas por fatores de crescimento. Porém, foi observado que alguns fatores de crescimento estavam expressos apenas no *MSL*, como o *EGFR*, *PDGF*, dentre outros. Além disso, alta expressão de genes envolvidos com a angiogênese e de marcadores de células-tronco, assim como a baixa expressão de genes envolvidos com a proliferação celular, também foram observadas apenas no subtipo *MSL* (LEHMANN et al. 2011). O subtipo LRA foi o mais diferente de todos os subtipos. Apesar de serem tumores RE negativos, esse subtipo demonstrou enriquecimento de vias de sinalização

reguladas por hormônios, incluindo síntese de esteróides, metabolismo de porfirina e metabolismo de andrógeno/estrógeno através do receptor de andrógeno que foi encontrado estar 9 vezes aumentado nesse subtipo em relação aos outros. Além disso, esse subtipo apresentou um padrão aumentado de expressão de genes luminais (LEHMANN et al. 2011).

Além dos autores caracterizarem cada subtipo molecular do CMTN, eles identificaram alvos moleculares em modelos de linhagens celulares para todos os subtipos que podem fornecer plataformas pré-clínicas para o desenvolvimento de terapias efetivas no futuro (Quadro 1). Por exemplo, os autores mostraram que as linhagens celulares dos subtipos *BL1* e *BL2* tiveram boa resposta à cisplatina. E os subtipos *M* e *MSL* foram sensíveis ao NVP-BEZ235 (inibidor de PI3K/mTOR) e ao dasatinib (inibidor de Src/Abl) (LEHMANN et al. 2011; LEHMANN e PIETENPOL 2014). Todas essas características moleculares e os alvos potencialmente terapêuticos estão resumidas no quadro 1.

**Quadro 1** - Características moleculares e alvos terapêuticos dos subtipos do CMTN de acordo com a classificação molecular de LEHMANN et al. (2011) e do estudo LEHMANN e PIETENPOL (2014).

Subtipo molecular	Alterações moleculares	Potenciais alvos terapêuticos
BL1	Ciclo celular	Inibidores de PARP
	Reparo do DNA (ATR; BRCA)	Agentes genotóxicos
BL2	Proliferação celular Sinalização dos fatores de crescimento (EGFR, MET, NGF, Wnt/ $\beta$ -catenin, IGF-1R)	Inibidores de mTOR
	Glicólise, gliconeogênese	Inibidor dos fatores de crescimento
IM	Marcadores mioepiteliais Processos do sistema imune (Vias CTLA4, IL12, IL7, processamento e apresentação de antígenos) Assinatura gênica para o CM medular (raros CMTN com prognóstico favorável)	Inibidores PD1/PD-L1
M	Motilidade celular	Inibidores de mTOR
	Diferenciação celular	Tratamento alvo da TEM e CSC
	Sinalização de fator de crescimento Transição epitélio-mesênquima (TEM)	
MSL	Similar ao M	Inibidores de PI3K
	Baixa proliferação celular	Terapia antiangiogênica
	Angiogênese	Antagonista de Src
LRA	Receptor de andrógeno (RA)	Terapia antiandrogênica
	Padrão de expressão gênica Luminal Subtipo apócrino molecular	

Abreviações: BL, *basal-like*; CSC, *cancer stem cells*; EGFR, *epidermal growth factor receptor*; TEM, transição epitélio-mesênquima; IGF-1R, *insulin-like growth factor I receptor*; IL, *interleukin*; IM, imumodulador; LAR, *luminal androgen receptor*; M, *mesequimal*; MET, *hepatocyte growth factor*; mTOR, *mammalian target of rapamycin*; MSL, *mesenquimal stem-like*; NGF, *nerve growth factor*; PARP, *poly ADP ribose polymerase*; PD1, *programmed cell death 1*; PD-L1, *programmed death-ligand 1*; PI3K, *phosphatidylinositol 3-kinase*; CMTN, câncer de mama triplo-negativo.

MASUDA et al. (2013) também investigaram a aplicabilidade clínica da classificação molecular de LEHMANN et al. (2011), tendo corroborado com os dados deste estudo, mostrando que o subtipo BL1 apresenta maior pCR (51%) em comparação ao BL2 (0%) e LRA (10%) após quimioterapia neoadjuvante baseada em antraciclina e ciclofosfamida seguida de taxanos.

Mais recentemente, LEHMANN et al. (2016) realizaram uma análise de validação mais acurada dos subtipos previamente estabelecidos e observaram que o



perfil de expressão dos subtipos IM e MSL foi influenciado pela alta presença dos linfócitos infiltrantes do tumor, TILs (do inglês, *tumor-infiltrating lymphocytes*) presentes nos tumores IM, e das células estromais associadas ao tumor e vistas nos tumores MSL. Assim, eles propuseram uma nova classificação do CMTN, sendo as mesmas resumidas em *basal-like 1 (BL1)*, *basal-like 2 (BL2)*, mesenquimal (M) e luminal receptor de andrógeno (LRA). Esses resultados sugerem, portanto, que o microambiente desses tumores, como observado pelos marcadores do sistema imune e pelo estroma, pode influenciar na classificação molecular do CMTN.

Nesse contexto, temos observado que, nos últimos anos, a contribuição do sistema imune no CMTN tem sido alvo de muitas pesquisas pois são tumores que apresentam uma alta imunogenicidade devido à alta infiltração de células do sistema imune, como os TILs e os macrófagos associados ao tumor (TAM, do inglês *tumor associated macrophages*). Além disso, é conhecido que a presença dos TILs e dos TAMs nos tumores pode levar ao bom prognóstico e altas taxas de pCR, como observado por muitos estudos (WEST et al. 2011; LOI et al. 2014; ADAMS et al. 2014; TUNG e WINER 2015; KARN et al. 2015; SIKANDAR et al. 2017). Assim, diferentes estratégias estão sendo desenvolvidas para a modulação da resposta imune para melhor resposta terapêutica no CMTN (STANTON et al. 2016). Os marcadores PD-1 (do inglês, *Programmed cell death protein 1*), expressos nas células T ativas, e PD-L1 (do inglês, *Programmed cell death ligand 1*), expressos em células no sistema imune, como as células apresentadoras de antígenos, e também em células tumorais, representam promissores inibidores do checkpoint do sistema imune para diversos tipos de câncer, principalmente para o CMTN cuja expressão de PD-L1 é frequentemente encontrada (MITTENDORF et al. 2014; HERSEY e

GOWRISHANKAR 2015; LEE et al. 2016; FELD e HORN 2016; NANDA et al. 2016; SU et al. 2016). Assim, o CMTN caracterizado por um perfil de expressão imunogênica pode representar tumores que melhor responderão à imunoterapia (BONSANG-KITZIS et al. 2016; BROCKHOFF et al. 2017), embora apenas 8 a 20% dos tumores TN possam se beneficiar com imunoterapias anti-PD-1 e anti-PD-L1 (SOLINAS et al. 2017).

Para melhor elucidação das características moleculares do CMTN, temos visto um crescente número de pesquisas que avaliam o perfil global de expressão gênica através da tecnologia de sequenciamento de nova geração (NGS, do inglês *Next Generation Sequencing*) como uma importante abordagem para investigar a complexidade molecular desse subgrupo de câncer de mama (WUSHOU et al. 2015; JIANG et al. 2016). Com o avanço do NGS, o perfil de expressão global através do sequenciamento do RNA (RNA-Seq) tem sido uma medida quantitativa de grande abrangência e sensibilidade que permite o acesso da alta diversidade das espécies de RNAs e, portanto, do transcriptoma completo de uma célula (WILHELM e LANDRY 2009; KUKURBA e MONTGOMERY 2016). Com suas vantagens em termos de um maior alcance dinâmico e da capacidade de descobrir novos RNAs, dentre outras aplicações, junto ao seu crescente avanço, a metodologia do RNA-Seq tem esclarecido a complexidade do genoma humano em diversas doenças, incluindo do CMTN (ILOTT e PONTING 2013; DING et al. 2014).

Recentemente, BURSTEIN et al. (2015) analisaram o perfil genômico e do transcriptoma completo por RNA-Seq de 198 tumores TN, classificando o CMTN em 4 subtipos: luminal receptor de andrógeno (LRA), mesenquimal (MES), *basal-like* imunossupressor (BLIS) e *basal-like* imunoativador (BLIA). Dentre estes, o BLIS

apresentou pior prognóstico e o *BLIA* foi o que mostrou melhor sobrevida livre de doença. Os autores também identificaram alvos específicos de cada subtipo: para o LRA, os alvos foram o receptor de andrógeno (RA) e a mucina de superfície celular (MUC1); para o MES: o receptor de fator de crescimento derivado das plaquetas (PDGF) e c-Kit; *BLIS*: uma molécula imunossupressora (VTCN1); e para *BLIA*: moléculas de transição de sinal como STAT3 e citocinas (BURSTEIN et al. 2015). Outros trabalhos já mostraram que o perfil de expressão global dos tumores *BLIA* é caracterizado pelo enriquecimento das funções das células do sistema imune, como das células B, células T e das natural-killer (LOI et al. 2014; PARK et al. 2017). Em outras pesquisas também foi observado, através do perfil de expressão global de 165 pacientes com CMTN, que os tumores TN podem ser classificados em 4 subtipos: imunomodulador (IM); luminal receptor de andrógeno (LRA); mesenquimal (M) e *basal-like* imunossupressor (*BLIS*), cada um apresentando seu único padrão de expressão. Dentre esses subtipos, o *BLIS* foi o que apresentou pior sobrevida livre de recaída quando comparado com os outros subtipos (LIU et al. 2016). Nesse último estudo, eles observaram que esses 4 subtipos tiveram associação significativa com os subtipos da classificação molecular de LEHMANN et al. (2011). JIANG et al. (2016) também definiram uma assinatura gênica a partir do perfil de expressão global como uma ferramenta de valor preditivo e prognóstico para pacientes com CMTN. De acordo com cada assinatura gênica, as pacientes foram classificadas em baixo e alto risco. Estas últimas tiveram maior probabilidade de recidiva e alcançaram menor benefício da quimioterapia baseada em taxanos, o que demonstra a importância do conhecimento do padrão de expressão global, ou seja, do transcriptoma completo desses tumores para melhor manejo clínico das pacientes com CMTN (JIANG et al.

2016). Assim, elucidar o perfil do transcriptoma completo do CMTN é de grande contribuição para uma maior compreensão da heterogeneidade dessa neoplasia, a qual suporta a necessidade de investigar genes potencialmente alvos terapêuticos ou vias alteradas dentro de cada subtipo molecular.

Apesar do crescente número de pesquisas envolvendo o perfil transcricional por RNA-Seq do CMTN, ainda são poucos aqueles que investigam os processos biológicos envolvidos com a tumorigênese da doença (CRAIG et al. 2013; KODURU et al. 2017; LIU et al. 2017). A pesquisa desenvolvida por CRAIG et al. (2013) revelou que o CMTN é enriquecido com genes que controlam o ciclo celular e mitose. Entre os genes mais diferencialmente expressos nos tumores TN está o *FOXMI* e, entre as ontologias gênicas, estão o enriquecimento de vias relacionadas com o sistema imune, metabolismo do RE, do RA e metabolismo de pentose e glicorunato.

Recentemente, KODURU et al. (2017) desenvolveram um estudo com o objetivo de avaliar a diferença do perfil transcricional dos tumores TN em relação ao tecido normal da mama adjacente ao tumor através do RNA-Seq. Nesse estudo, eles identificaram genes diferencialmente expressos (GDEs) no CMTN em relação às amostras normais que podem servir de novos biomarcadores para a precoce detecção da doença. E, dentre as vias de sinalização desreguladas pelos GDEs, a do TGF-beta foi a mais significativamente afetada no CMTN. Isso ressalta a importância do desenvolvimento de trabalhos de NGS que sejam promissores para o conhecimento do surgimento e progressão do CMTN.

### 1.2.3 Caracterização genética

Recentes avanços na análise do genoma completo por NGS têm revelado com precisão uma grande variedade de alterações genômicas, permitindo o conhecimento da grande complexidade do perfil mutacional do CMTN (DING et al. 2010; CURTIS et al. 2012; SHAH et al. 2012; The Cancer Genome Atlas Network 2012; NIK-ZAINAL et al. 2016).

Os genes mais mutados somaticamente no câncer de mama, *PI3K* e *TP53*, são também predominantemente mutados no CMTN. Mutações nos genes *NRAS*, *EGFR*, *USH2A*, *MYO3A*, *MLL2*, *BRCA1*, *RBI*, *PTEN* também têm sido reportadas como frequentes em CMTN (The Cancer Genome Atlas Network 2012; SHAH et al. 2012; KOBOLDT et al. 2012; JUDES et al. 2016; WEISMAN et al. 2016). Mutações em genes envolvidos com a forma e motilidade celular e genes importantes da matrix extracelular também foram reportadas no CMTN (SHAH et al. 2012) assim como amplificação e aumento de expressão de *MYC* (CURTIS et al. 2012).

Embora mais de 80% das pacientes com CMTN carreguem mutações deletérias no gene *TP53* e um grupo menor de pacientes carreguem mutações pontuais de ganho-de-função em genes que controlam a via de sinalização de PI3K (*PTEN*; *PI3KCA* e *IMPP4B*) (The Cancer Genome Atlas Network 2012), aproximadamente 10% apresentam mutações pontuais de ganho-de-função que são mais comuns em outros cânceres, tais como encontrada nos genes *BRAF*, *KRAS* e *EGFR* (SHAH et al. 2012; STEPHENS et al. 2012; NIK-ZAINAL et al. 2016). Além das mutações de perda de função de *TP53*, eventos de perda de função em outros genes envolvidos com a maquinaria de reparo do DNA, como em *RBI* e *BRCA1*,

também são comuns em CMTN (ANDRE et al. 2009; ESTELLER et al. 2000; The Cancer Genome Atlas Network 2012; NIK-ZAINAL et al. 2016).

Outra característica genética do CMTN é a alta instabilidade genômica. Isso pode ser devido a diversos fatores, tais como: inativação precoce de p53; aberrante expressão de c-myc; mutação ou expressão aberrante de genes do *checkpoint* do ciclo celular (*RBI*), de genes envolvidos com a divisão celular (amplificação de ciclinas) e de genes responsáveis pela manutenção do genoma (reparo do DNA por recombinação homóloga; perda de *PTEN*), os quais podem contribuir significativamente para esse cenário (XU et al. 2014). Além disso, funções como ciclo celular, recombinação homóloga e motilidade celular estão supostamente desreguladas em CMTN devido a outras alterações genéticas como expressão monoalélica associada à perda de heterozigosidade e alteração em número de cópia (HA et al. 2012). Estudos mostram que alterações em número de cópia afetando a expressão gênica não são eventos comuns nesse subtipo de câncer de mama, embora essas alterações já tenham sido observadas em genes supressores de tumor como *RBI* (5%) e *PTEN* (3%) e em genes que podem ser alvos terapêuticos como o receptor do fator de crescimento epidermal (*EGFR*) (5%) e o receptor do fator de crescimento do fibroblasto tipo 2 (*FGFR2*) (3%) (HAN et al. 2008; TURNER et al. 2010; SHAH et al. 2012; HA et al. 2012).

Dentre os genes citados acima e que apresentam um papel importante na manutenção da estabilidade genômica nos tumores TN, o *BRCA1* tem uma posição de destaque por ser um fator genético de risco associado ao desenvolvimento do CMTN (GONZALEZ-ANGULO et al. 2011; SHARMA et al. 2014; COUCH et al. 2015; AFGHAHI et al. 2016; BUYS et al. 2017). Dados mostram que 70% das

pacientes com mutação germinativa em *BRCA1* desenvolvem preferencialmente CMTN e, dentre as pacientes diagnosticadas com CMTN, 10 a 20% carregam mutação germinativa em *BRCA1* (FOUKLES et al. 2010; GONZALEZ-ANGULO et al. 2011; MAVADDAT et al. 2013; CARRARO et al. 2013; RUMMEL et al. 2013; PELLEGRINO et al. 2016; BRIANESE et al. 2018).

Em um estudo prévio, realizado pelo nosso grupo, foi observado 50% de mutação germinativa em *BRCA1* em pacientes diagnosticadas com CMTN em idade jovem (menor que 35 anos de idade), e a taxa de detecção de mutação aumentava para 83% quando associava com história familiar positiva para câncer de mama hereditário (CARRARO et al. 2013). Mais recentemente, outro estudo do nosso grupo, que avaliou 130 amostras de TN, detectou mutação germinativa em *BRCA1* em 11,5% em mulheres diagnosticadas com CMTN em qualquer idade, sendo este percentual aumentado, significativamente, em mulheres diagnosticadas em idade jovem, passando para 26%. Esse dado corroborou com a alta frequência de mutação germinativa em *BRCA1* em mulheres diagnosticadas com CMTN antes dos 35 anos de idade (CARRARO et al. 2013), confirmando o alto percentual de câncer hereditário em CMTN (BRIANESE et al. 2018). Esses dados mostram que o CMTN diagnosticado em mulheres jovens é fator de risco para presença de mutação germinativa em *BRCA1* (ROBERTSON et al. 2012; RUMMEL et al. 2013; CARRARO et al. 2013; BRIANESE et al. 2018).

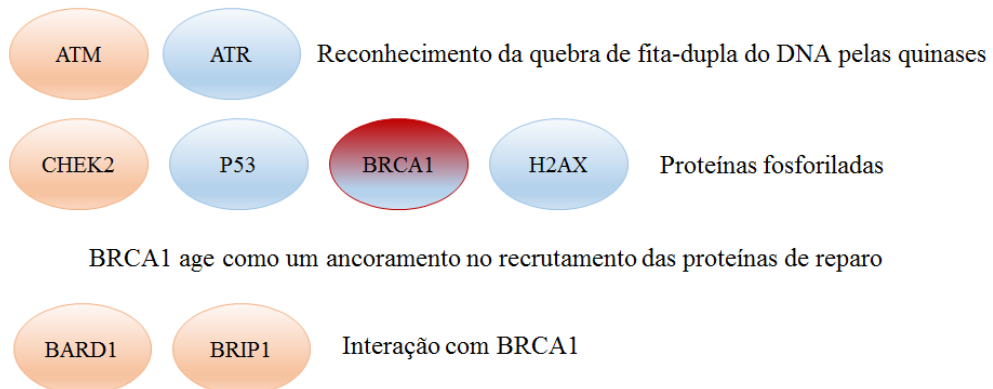
### **1.3 GENE *BRCA1* E A DEFICIÊNCIA NA RECOMBINAÇÃO HOMÓLOGA**

O gene *BRCA1* é composto por 24 éxons, dos quais 22 codificam a proteína nuclear BRCA1 de 1.863 aminoácidos, e está localizado no braço longo do cromossomo 17 (MIKI et al. 1994). BRCA1 é funcionalmente caracterizada por conter três domínios principais: o domínio N-terminal, o domínio de localização nuclear (NLS) e um domínio C-terminal (ROSEN et al. 2003; HUEN et al. 2010). É uma proteína cujas funções são primordiais nos principais processos celulares, incluindo reparo a danos causados na fita dupla do DNA e regulação dos genes envolvidos nesse processo, controle do ciclo celular e manutenção da estabilidade cromossômica (YOSHIDA e MIKI 2004; VENKITARAMAN 2014). Um dos principais mecanismos responsáveis pelo reparo do DNA a danos causados na fita simples do DNA, que levam à quebra da dupla fita do DNA (QDF) exercida por BRCA1, juntamente com diversas outras proteínas, é a recombinação homóloga (RH). A RH é um processo essencial utilizado pelas células para a manutenção da estabilidade genômica durante a replicação do DNA, uma vez que a QDF pode induzir o desenvolvimento de rearranjos cromossômicos e, conseqüentemente, levar à instabilidade genômica (ROY et al. 2012; CAESTECKER e VAN DE WALLE 2013). Basicamente, a RH acontece em muitas etapas e por diversas proteínas. Inicialmente, ocorre o reconhecimento da QDF pelas quinases ATM e ATR, que fosforilam e recrutam as proteínas CHEK2, TP53, BRCA1 e H2AX. BRCA1, juntamente com as proteínas BARD1 e BRIP1, formam um complexo de ancoramento que organiza a montagem das proteínas de reparo exercendo, portanto,



um papel essencial para a ativação da ressecção do DNA. Na etapa seguinte, o complexo MRN, formado pelas proteínas MRE11, RAD50 e NBS1, é responsável pela ressecção das extremidades da QDF que são ligadas e revestidas pela proteína RPA. Posteriormente, BRCA2 é recrutada com o auxílio de PALB2 para recrutar e fazer o carregamento da proteína RAD51 na extremidade da QDF revestida pela RPA, com a participação das proteínas RAD51B, RAD51C e RAD51D. Depois, a RAD51 executa a invasão da fita homóloga do DNA, permitindo o pareamento da região danificada com a cromátide-irmã não danificada, a qual servirá como um molde para o processo de reparo do DNA e efetiva manutenção da replicação. Todas essas etapas estão resumidas na Figura 2. Além do seu papel na etapa inicial da RH, BRCA1 também pode ativar a transcrição de vários genes envolvidos em outros mecanismos de reparo do DNA como o reparo por excisão de bases (ZHU et al. 2011).

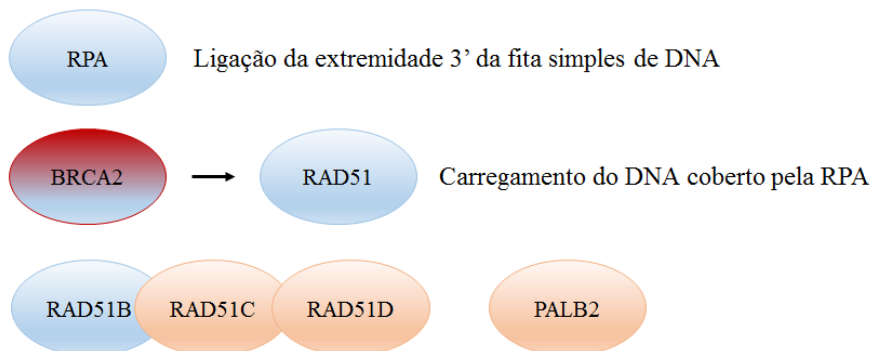
### 1. Reconhecimento e montagem das proteínas de reparo da quebra de fita-dupla do DNA



### 2. Ressecção do DNA



### 3. Carregamento da RAD51



### 4. Invasão da fita homóloga do DNA



### 5. Síntese e reparo do DNA

Fonte: Adaptado de WALSH (2015)

**Figura 2** - Representação esquemática da função das proteínas envolvidas com o reparo de quebra de dupla fita do DNA através da recombinação homóloga

Sendo assim, uma vez que *BRCA1* é um gene supressor de tumor primordial que exerce um papel crucial no reparo de quebra de fita dupla do DNA por recombinação homóloga (WALSH 2015), alterações genéticas, como inativação germinativa em heterozigose acompanhada por uma inativação somática do segundo alelo, como perda da heterozigosidade, silenciamento epigenético e outros mecanismos que levam à perda de função da proteína, resultarão em limitadas capacidades de reparo do DNA pela célula, uma RH deficiente e, conseqüentemente, maior susceptibilidade a uma transformação maligna (SEVERSON et al. 2015, LORD e ASHWORTH 2016).

Um achado clinicamente relevante em tumores TN é alta taxa de deficiência do reparo do DNA por recombinação homóloga similarmente observada nos tumores de pacientes que carregam mutação em *BRCA1* (FOULKES et al. 2010). Essa característica genética tem sua importância clínica pois a deficiência na recombinação homóloga leva a um padrão mutacional específico que pode beneficiar os pacientes ao uso de alguns agentes terapêuticos, como quimioterapia à base de platina ou inibidores de PARP (GEYER et al. 2017).

Além disso, estudos mostram que tumores TN esporádicos (sem mutação germinativa em *BRCA1*) podem exibir características moleculares semelhantes aos tumores hereditário (que apresentam mutação germinativa em *BRCA1*), sugerindo que outros mecanismos podem estar envolvidos com o processo tumorigênico do CMTN (TURNER et al. 2004; JOOSSE et al. 2011; HA et al. 2012; SHAH et al. 2012; LIPS et al. 2013). No recente trabalho do nosso grupo, BRIANESE et al. (2018) observaram que o promotor do gene *BRCA1* está hipermetilado em 20,6% de toda a casuística, e 39% em tumores de mulheres diagnosticadas antes dos 40 anos,

novamente apresentando maior percentual no grupo de mulheres jovens. Destaque-se que nossos dados mostraram que todos os TN classificados como hipermetilados eram classificados como esporádicos (sem mutação germinativa em *BRCA1*), sugerindo que mutação germinativa ou hipermetilação do promotor no tumor são eventos excludentes. Esses dados revelam ainda que a deficiência de *BRCA1* parece ser um evento desencadeador do tumor para um percentual significativo dos casos de TN esporádicos.

Dentro desse contexto, esse trabalho mostra a importância de investigar aspectos moleculares, como o perfil transcricional completo, do CMTN diagnosticado em pacientes jovens, antes dos 50 anos de idade, sob diferentes aspectos da deficiência de *BRCA1*. Para isso, avaliamos as alterações qualitativas e quantitativas dos tumores TN Hereditário (com mutação germinativa em *BRCA1*) e TN Esporádico (negativos para mutação germinativa em *BRCA1/2* e com deficiência ou não em *BRCA1* por hipermetilação no promotor) e do tecido normal adjacente ao tumor com o intuito de melhor entender os processos biológicos envolvidos com a tumorigênese do CMTN Hereditário e do CMTN Esporádico.

Portanto, os resultados deste estudo, juntamente com outros aspectos genéticos e moleculares investigados no nosso grupo destes grupos de amostras, contribuirá para melhor compreensão molecular do CMTN no contexto de *BRCA1* e, por conseguinte, possibilitarão novas descobertas nos diferentes subgrupos de TN, colaborando com o entendimento da heterogeneidade desse subtipo de câncer de mama.

## **2 OBJETIVOS**

### **2.1 OBJETIVO GERAL**

Investigar o perfil transcricional do CMTN diagnosticado em pacientes jovens, antes dos 50 anos de idade, sob diferentes aspectos da deficiência de *BRCA1*.

### **2.2 OBJETIVOS ESPECÍFICOS**

- Investigar diferenças no padrão de expressão global de tumores TN Hereditário (com mutação germinativa em *BRCA1*) e TN Esporádico (com deficiência ou não em *BRCA1* por hipermetilação no promotor) e do tecido normal adjacente ao tumor;
- Verificar os processos biológicos envolvidos com a tumorigênese do CMTN Hereditário e do CMTN Esporádico.

### 3 MATERIAIS E MÉTODOS

#### 3.1 ASPECTOS ÉTICOS

Este estudo teve a aprovação do Comitê de Ética em Pesquisa (CEP) do A.C.Camargo Cancer Center sob o número 1746/13C (Anexo 1) e está vinculado ao projeto temático aprovado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (nº 2013/23277-8), sob a responsabilidade da Dra. Dirce Maria Carraro.

#### 3.2 CASUÍSTICA

Foram selecionados casos de CMTN de pacientes diagnosticadas no A.C.Camargo Cancer Center em idade de até 50 anos a partir do levantamento realizado pelo nosso grupo dos 131 tumores de mama triplo-negativo (TN) (Quadro 2). No mais recente trabalho do grupo, BRIANESE et al. 2018, esses tumores foram rastreados para mutação patogênica germinativa dos genes *BRCA1* e *BRCA2* através da metodologia de *Next Generation Sequencing* (NGS) e para alteração em número de cópias de *BRCA1* através da abordagem de *Multiplex Ligation-dependent Probe Amplification* (MLPA). O silenciamento somático de *BRCA1* também foi avaliado pela análise de metilação do promotor de *BRCA1* através do método de conversão por bissulfito de sódio do DNA tumoral. Então, todos esses casos de CMTN foram classificados em CMTN Hereditário e CMTN Esporádico como sendo tumores

diagnosticados em pacientes portadoras de mutação patogênica germinativa em *BRCA1* e não portadoras de mutação germinativa em ambos genes *BRCA1/2*, respectivamente. Os tumores TN Esporádicos foram ainda classificados em *BRCA1*-deficiente e *BRCA1*-proficiente, correspondendo, respectivamente, a tumores com o promotor de *BRCA1* hipermetilado e tumores com o promotor de *BRCA1* não hipermetilado (BRIANESE et al. 2018).

Para aumentar o número de amostras de CMTN em idade jovem, também incluímos 3 amostras (MJ2021; MJ2026; e MJ2003) de um outro projeto do grupo, diagnosticadas com câncer de mama em idade jovem com histórico familiar de câncer de mama, previamente rastreadas para mutação pontual (por sequenciamento capilar) e rearranjos genômicos (por MLPA) nos genes *BRCA1*, *BRCA2*, *TP53* e *CHEK2* (CARRARO et al. 2013; SILVA et al. 2014) (Quadro 2).

**Quadro 2 - Características moleculares dos casos de CMTN selecionados para o estudo**

Grupo	Casos	Idade	Variante patogênica germinativa <i>BRCA1</i>	Frequência alélica da variante no tumor (%)	Metilação no promotor de <i>BRCA1</i>	Classificação final de <i>BRCA1</i>
Hereditário	M256T	35	c.1A>G; p.(Met1?)	>60 (sanger)	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M280T	36	c.1123_1124delinsA; p.(Leu375Lysfs*19)	67.0	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M282T	38	c.2405_2406delTG; p.(Val802Glufs*7)	67.42	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M287T	44	c.4183C>T; p.(Gln1395*)	51.01	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M292T	34	c.4096+1G>A; splice site donor	60.53	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M303T	41	c.4484G>T; p.(Arg1495Met)	89.52	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	M355T	43	c.181T>G; p.(Cys61Gly)	91.37	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	MJ2021	27	c.181T>G; p.(Cys61Gly)	NA	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Hereditário	MJ2026	31	c.3331_3334del; p.(Gln1111Asnfs*5)	NA	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -mutado
Esporádico <i>BRCA1</i> -Def	M15T	46	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M299T	38	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M347T	44	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M357T	36	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M365T	50	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M366T	34	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M377T	46	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M379T	32	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Def	M385T	33	-	-	<i>BRCA1</i> -hipermetilado	<i>BRCA1</i> -deficiente
Esporádico <i>BRCA1</i> -Prof	M7T	36	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M214T	43	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M217	41	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M269T	29	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M275T	36	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M277T	45	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M278T	41	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M283T	42	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M284T	44	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M286T	39	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M290T	31	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M291T	50	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M298T	38	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M312T	48	-	-	<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente



## Cont/ Quadro 2

Grupo	Casos	Idade	Variante patogênica germinativa <i>BRCA1</i>	Frequência alélica da variante no tumor (%)	Metilação no promotor de <i>BRCA1</i>	Classificação final de <i>BRCA1</i>
Esporádico <i>BRCA1</i> -Prof	M348T	46	-		<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M351T	47	-		<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	M358T	47	-		<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	MIC20T	50	-		<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente
Esporádico <i>BRCA1</i> -Prof	MJ2003T	26	-		<i>BRCA1</i> -não-hipermetilado	<i>BRCA1</i> -proficiente

Hereditário: CMTN Hereditário *BRCA1*-mutado; Esporádico *BRCA1*-Def.: CMTN Esporádico *BRCA1*-hipermetilado; Esporádico *BRCA1*-Prof.: CMTN Esporádico *BRCA1*-não-hipermetilado; - Não foi detectada variante patogênica germinativa em *BRCA1*; NA: frequência alélica no tumor não avaliada.

**Fonte:** CARRARO et al. (2013); SILVA et al. (2014); BRIANESE et al. (2018).

### 3.3 EXTRAÇÃO DE RNA

As moléculas de RNA total, disponíveis no sistema do Biobanco do A.C.Camargo Cancer Center, foram selecionadas a partir de todos os casos de CMTN, tanto dos casos pareados com o tecido normal adjacente ao tumor como dos casos não pareados. A extração do RNA total do tecido tumoral e normal adjacente ao tumor, ambos tecidos congelados, foi realizada pelo Banco de Macromoléculas do A.C.Camargo Cancer Center de acordo com o protocolo operacional padrão do Banco de Macromoléculas (OLIVIERI et al. 2013). Resumidamente, até 30mg de tecido foram homogeneizados utilizando o equipamento Precellys® (Bertin Technologies, Montigny-le-Bretonneux, França) com 400 ul do tampão RLT com  $\beta$ -Mercaptoethanol ( $\beta$ -ME) de acordo com as recomendações do fabricante para o kit de extração QIASymphony RNA (Qiagen, CA, EUA). A extração do RNA foi feita de forma automatizada no equipamento QIASymphony (Qiagen, CA, EUA) utilizando o protocolo miRNA CT 400 V8. Neste protocolo é recuperado o RNA total, inclusive o miRNA, e é feito o tratamento do RNA total com a enzima DNase fornecida pelo kit durante o processo de purificação. A concentração e a pureza do RNA foram avaliadas no equipamento Nanodrop™ ND-1000 (Thermo Scientific, DE, EUA) e a integridade foi determinada pelo equipamento Bioanalyzer Agilent 2100 (Agilent Technologies, CA, EUA). O Bioanalyzer fornece um número de integridade do RNA (RIN-RNA *Integrity Number*) que varia de 1 a 10, onde 1 refere-se a um RNA degradado e 10 a um RNA íntegro (SCHROEDER et al. 2006). A quantificação do RNA foi realizada pelo método fluorométrico *Qubit® Fluorometer*, usando o kit *Qubit RNA BR Assay* (Thermo Fisher), que é um método

de quantificação baseado em fluorescência que emprega fluoróforos específicos que se ligam ao RNA simples fita, levando a uma quantificação mais específica e exata quando comparada com métodos baseados em espectrometria.

Selecionamos as moléculas com uma boa integridade medida pelo valor do RIN (do inglês, *RNA Integrity Number*), priorizando aquelas com o RIN maior ou igual a 4 ( $RIN \geq 4.0$ ). Todas as moléculas selecionadas para o estudo estão descritas no Quadro 3, o qual mostra as características da avaliação anatomopatológica e os dados de extração e quantificação das moléculas de RNA total dos casos utilizados no estudo.

**Quadro 3** - Dados da avaliação anatomopatológica de todas amostras tumorais e normais e da extração e quantificação das moléculas de RNA total a partir dos casos de CMTN utilizados no estudo.

<b>Amostra</b>	<b>Avaliação H&amp;E</b>	<b>RIN</b>	<b>R260 / 280</b>	<b>R260 / 230</b>	<b>Concentração (ng/uL)</b>
ID_M256N	Não neoplásico em 100% da amostra	8,4	2,3	0,9	101
ID_M256T	Carcinoma ductal invasivo em 100% da amostra	9,9	2,1	1,6	985
ID_M280N	Não neoplásico em 100% da amostra	7,2	2,0	0,8	49,7
ID_M280T	Carcinoma ductal invasivo em 100% da amostra	8,1	2,1	1,8	481
ID_M282N	Não neoplásico em 100% da amostra	7,6	2,1	0,8	39,3
ID_M282T	Carcinoma ductal invasivo em 90% da amostra	9,6	2,0	1,8	27
ID_M287N	Não neoplásico em 100% da amostra	7,7	2,2	1,3	23,4
ID_M287T	Carcinoma ductal invasivo em 30% da amostra	9,6	2,1	1,8	102,3
ID_M292N	Não neoplásico em 100% da amostra	7,8	2,1	1,6	40,9
ID_M292T	Carcinoma mamário invasivo em 100% da amostra	7,1	2,1	2,0	372
ID_M303N	Não neoplásico em 100% da amostra	7,9	2,1	1,4	73,7
ID_M303T	Carcinoma ductal invasivo em 100% da amostra	7,5	2,1	1,9	887
ID_M355N	Não neoplásico em 100% da amostra	7,6	2,1	1,7	67,7
ID_M355T	Carcinoma mamário invasivo em 30% da amostra	8,7	2,0	1,8	82,8
ID_MJ2026N	Não neoplásico em 100% da amostra	6,3	2,1	1,6	241
ID_MJ2026T	Carcinoma ductal invasivo em 90% da amostra	7,2	1,8	1,9	853
ID_MJ2021T	Carcinoma ductal invasivo em 100% da amostra	8,7	2,0	1,9	83,7
ID_M15N	Não neoplásico em 100% da amostra	6,4	2,1	1,4	83,2
ID_M15T	Carcinoma ductal invasivo em 100% da amostra	9,2	2,2	2,0	153
ID_M299N	Não neoplásico em 100% da amostra	7,5	2,0	0,5	45,4
ID_M299T	Carcinoma lobular pleomorfo em 100% da amostra	9,8	2,1	1,8	1270
ID_M365N	Não neoplásico em 100% da amostra	4,2	2,2	1,2	35,6
ID_M365T	Carcinoma mamário invasivo em 90% da amostra	6,9	2,1	1,6	144

Cont/ Quadro 3

Amostra	Avaliação H&E	RIN	R260 / 280	R260 / 230	Concentração (ng/uL)
ID_M377N	Não neoplásico em 100% da amostra	7,7	2,1	1,6	15
ID_M377T	Carcinoma ductal invasivo em 100% da amostra	9,7	2,1	1,8	40,2
ID_M347T	Carcinoma ductal invasivo em 100% da amostra	8,9	2,1	1,8	200
ID_M357T	Carcinoma ductal invasivo em 100% da amostra	9,7	2,0	2,0	1300
ID_M366T	Carcinoma ductal invasivo em 100% da amostra	9,1	2,1	1,7	481
ID_M379T	Carcinoma ductal invasivo em 100% da amostra	9,7	2,1	2,0	1180
ID_M385T	Carcinoma ductal invasivo em 100% da amostra	9,9	2,1	1,8	414
ID_M7N	Não neoplásico em 100% da amostra	5,6	2,1	1,8	126
ID_M7T	Carcinoma ductal invasivo em 90% da amostra	8,2	2,1	2,2	850
ID_M214N	Não neoplásico em 100% da amostra	9,8	2,1	1,5	175
ID_M214T	Carcinoma ductal invasivo em 50% da amostra	9,1	2,1	1,6	288
ID_M269N	Não neoplásico em 100% da amostra	9,1	2,1	1,6	278
ID_M269T	Carcinoma ductal invasivo em 100% da amostra	10,0	2,0	1,7	876
ID_M275N	Não neoplásico em 100% da amostra	7,2	1,9	0,6	40,8
ID_M275T	Carcinoma ductal invasivo em 60% da amostra	5,6	2,0	1,8	150
ID_M278N	Não neoplásico em 100% da amostra	4,9	2,1	1,1	121
ID_M278T	Carcinoma ductal invasivo em 30% da amostra	9,9	2,0	1,7	202,5
ID_M284N	Não neoplásico em 100% da amostra	5,9	1,9	1,3	20,8
ID_M284T	Carcinoma medular atípico em 90% da amostra	9,0	2,0	1,8	1,13
ID_M286N	Não neoplásico em 100% da amostra	7,2	2,1	0,6	52,9
ID_M286T	Carcinoma ductal invasivo em 100% da amostra	9,2	2,0	1,6	377
ID_M290N	Não neoplásico em 100% da amostra	7,7	2,1	1,7	104
ID_M290T	Carcinoma ductal invasivo em 100% da amostra	9,7	2,1	1,9	381
ID_M291N	Não neoplásico em 100% da amostra	8,4	2,03	1,72	31,1

**Cont/ Quadro 3**

<b>Amostra</b>	<b>Avaliação H&amp;E</b>	<b>RIN</b>	<b>R260 / 280</b>	<b>R260 / 230</b>	<b>Concentração (ng/uL)</b>
ID_M291T	Carcinoma ductal invasivo em 40% da amostra	9,6	2,11	1,71	145,6
ID_M298N	Não neoplásico em 100% da amostra	8,2	2,1	1,5	39,1
ID_M298T	Carcinoma ductal invasivo em 100% da amostra	8,7	2,1	1,62	239
ID_M312N	Não neoplásico em 100% da amostra	8,1	2,1	1,76	28,9
ID_M312T	Carcinoma ductal invasivo de 60% da amostra	8,3	2,1	1,3	43,2
ID_M358N	Não neoplásico em 100% da amostra	4,6	1,97	0,93	25,5
ID_M358T	Carcinoma mamário invasivo em 100%	9,0	2,11	1,93	175,2
ID_MJ2003N	Não neoplásico em 100% da amostra	9,7	2,1	1,7	451
ID_MJ2003T	Carcinoma ductal invasivo em 100% da amostra	6,3	2,1	2,2	226
ID_M217T	Carcinoma ductal invasivo em 100% da amostra	9,2	1,98	1,88	212
ID_M277T	Carcinoma ductal invasivo em 100% da amostra	7,6	2,11	0,8	221,6
ID_M283T	Carcinoma ductal invasivo em 100% da amostra	9,5	2,03	1,61	190
ID_M348T	Carcinoma ductal invasivo em 100% da amostra	8,6	2,05	1,42	44,5
ID_M351T	Carcinoma ductal invasivo em 100% da amostra	9,0	2,05	1,86	416
ID_MIC20T	Carcinoma ductal invasivo em 100% da amostra	6,8	2,12	2,01	110

### 3.4 PREPARO DAS BIBLIOTECAS DE cDNA

Para o sequenciamento do RNA de todas as amostras do estudo, foram construídas bibliotecas de cDNA utilizando o kit *TruSeq Stranded Total RNA LT - Set A* (Illumina), o qual contém 12 distintos adaptadores, e seguindo o protocolo *TruSeq Stranded Total RNA Sample Preparation LS* (Illumina), cujas etapas estão descritas, resumidamente, a seguir e esquematizadas na Figura 3. Considerando que mais de 90% da massa de RNA total de uma típica célula humana é composta de RNA ribossomal (rRNA) (PALAZZO e LEE 2015), nós escolhemos a metodologia da depleção do rRNA antes de iniciar a construção das bibliotecas de cDNA. Essa metodologia de depleção do rRNA é uma abordagem de maior aplicabilidade para se obter uma quantificação mais acurada do transcriptoma completo humano pois permite a identificação e quantificação de uma grande variedade de espécies de RNAs, como dos RNA poli-adenilados codificadores de proteínas (mRNAs) e dos RNAs não poli-adenilados não codificadores de proteínas (ncRNAs), dentre outras aplicabilidades (HUANG et al. 2011).

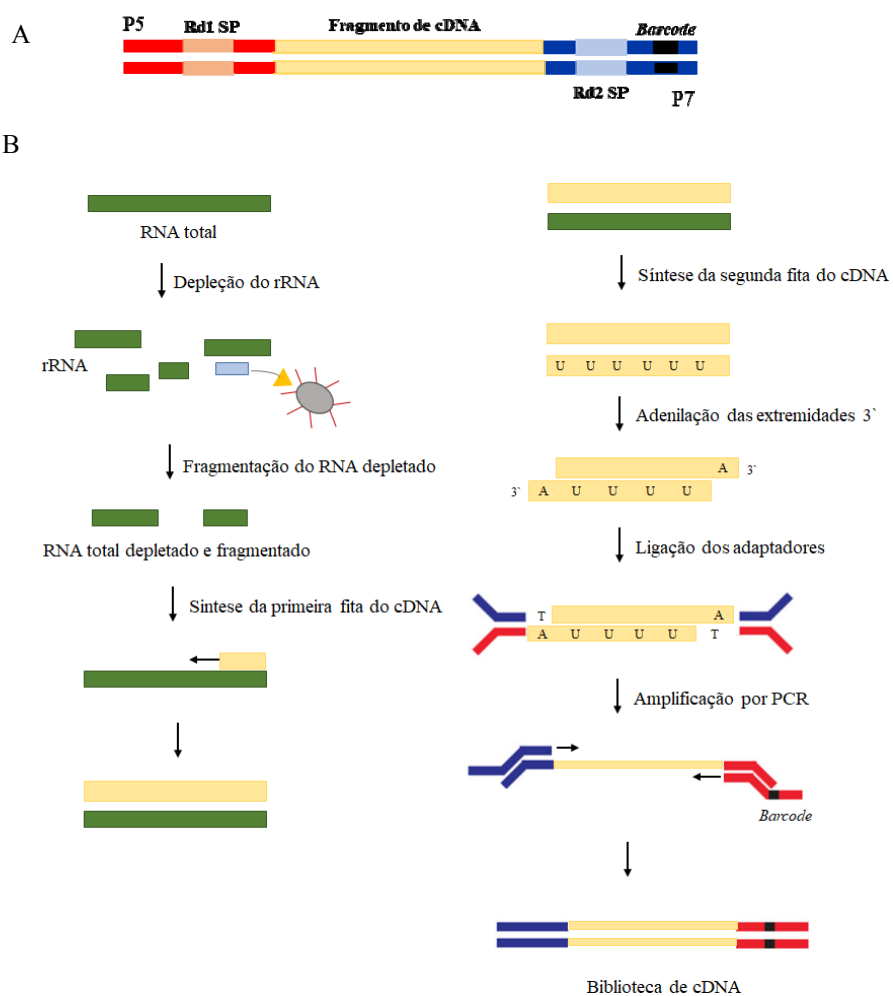
Para um rendimento satisfatório das bibliotecas, a recomendação do protocolo para a massa inicial de RNA total de ótima integridade ( $RIN \geq 7.0$ ) está entre 100 e 1000 nanogramas (ng). Neste estudo, nós utilizamos 300 ng de RNA total inicial que foram submetidos à depleção do rRNA através do kit *Ribo-Zero Human Gold* (Illumina). Essa etapa baseia-se na hibridização dos rRNAs (28S; 18S e 5S) a sequências complementares ligadas às esferas magnéticas (*rRNA Removal Beads* - Illumina). A eficiência da depleção foi analisada pela plataforma *Agilent Bioanalyzer 2100* utilizando o kit *Agilent RNA 6000 Pico*. Para as amostras que não tiveram uma

depleção eficiente, esse procedimento foi repetido a partir do RNA total inicial. Porém, caso ainda fosse verificada a presença do rRNA, foi realizada uma purificação em coluna da molécula de RNA total inicial com o intuito de melhorar a qualidade da mesma através do kit *RNeasy MinElute Cleanup* (Qiagen).

Posterior mensuração da depleção total do rRNA, seguiu-se o protocolo com a purificação do RNA livre de rRNA com esferas magnéticas específicas para RNA, *Agencourt RNAClean XP* (Beckman Coulter), seguindo com a fragmentação a 94°C por 8 min para a obtenção dos fragmentos com tamanho médio de 155pb, conforme recomendação do fabricante. Após a fragmentação, seguiu-se a síntese da primeira fita do cDNA, usando a transcriptase reversa e *primers* randômicos e a síntese da segunda fita do cDNA com a incorporação do deoxy-UTP (dUTP) no lugar de deoxy-TTP (dTTP) a 16°C por 1 hora. O uso do dUTP é uma metodologia utilizada para preservar a informação da fita original do RNA e, assim, manter a especificidade da fita (LEVIN et al. 2010). Isso é garantido durante a etapa de PCR da biblioteca final onde a polimerase não incorpora o dUTP e, assim, não ocorre a amplificação da segunda fita do cDNA. Ao final da etapa da síntese da segunda fita do cDNA, foram gerados fragmentos com extremidades 3' abruptas (*blunt end*). Posteriormente, as extremidades 3' foram adeniladas com a adição do nucleotídeo adenina (A) a 37°C por 30 minutos e 70°C por 5 minutos, para impedir que esses fragmentos se liguem entre si e os preparar para a etapa seguinte na ligação aos adaptadores, pois a adenina fornece complementariedade à timina presente nas extremidades 3' dos adaptadores. Assim, após a adenilação, foi feita a ligação dos adaptadores (P5 e P7) a 30°C por 10 minutos, os quais contêm sequências que identificam cada amostra (*barcodes*) e sequências complementares aos iniciadores da



etapa de amplificação por PCR e do sequenciamento da plataforma Illumina (Rd1SP e Rd2SP) (Figura 3A). Em seguida, ocorreu a amplificação por PCR da biblioteca de cDNA. Entre cada uma dessas etapas, a purificação e seleção de tamanho dos fragmentos foram realizadas utilizando esferas magnéticas para DNA, *Agencourt AMPure XP* (Beckman Coulter), com o objetivo de eliminar fragmentos de nucleotídeos de tamanho fora do esperado (abaixo de 200pb). Em seguida, as bibliotecas finais de cDNA foram analisadas quantitativamente pelo método fluorométrico *Qubit* com o kit *Qubit dsDNA HS Assay* (Thermo Fisher) e qualitativamente através da plataforma *Agilent Bioanalyzer 2100* com o kit *High Sensitivity DNA* (Agilent Technologies). Após essa análise quantitativa e qualitativa, as bibliotecas seguiram para a etapa de normalização para o sequenciamento. Todas essas etapas estão esquematizadas, resumidamente, na Figura 3B.



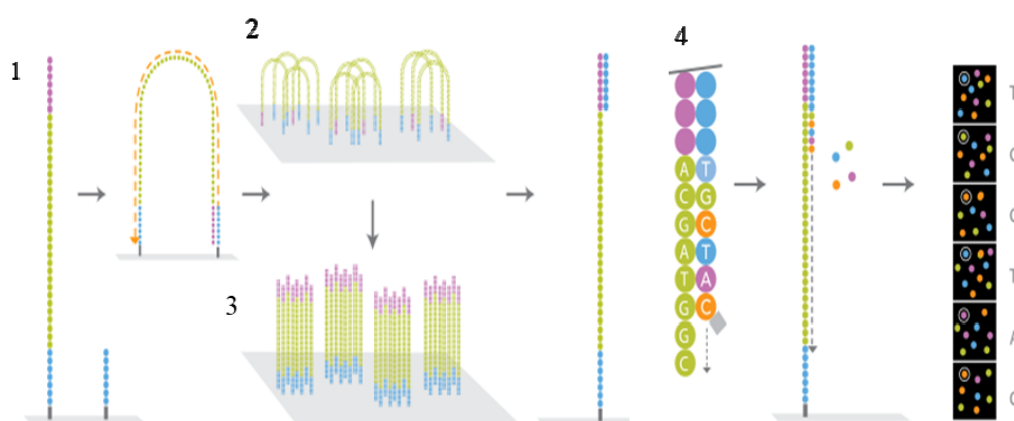
**Figura 3** - Esquema representativo da biblioteca final de cDNA. (A) Adaptadores *Illumina* P5 e P7. (B) Etapas da construção da biblioteca de cDNA de acordo com o protocolo *TruSeq Stranded Total RNA Sample Preparation LS* - *Illumina* (Adaptado de *TruSeq Stranded Total RNA Sample Preparation LS protocol*).

### 3.5 SEQUENCIAMENTO DO RNA (RNASeq)

Após a construção das bibliotecas de cDNA, o sequenciamento foi realizado através da plataforma *NextSeq 500* (Illumina). Inicialmente, as bibliotecas foram quantificadas por PCR em Tempo Real (qPCR), utilizando o kit *Kapa Library Quantification (KAPA Biosystems)* para a mensuração da molaridade. Assim, com o tamanho médio obtido pelo *Bioanalyzer* e a quantificação por qPCR, as quantidades equimolares de cada biblioteca foram calculadas a fim de normalizá-las a uma concentração molar inicial de 4nM necessária para iniciar o protocolo do sequenciamento *NextSeq System Denature and Dilute Libraries Guide Protocol A*. Uma vez normalizadas, desnaturadas e diluídas a 1.8pM, o *pool* das bibliotecas foi sequenciado (8 bibliotecas em cada *pool* por corrida de sequenciamento), no formato *paired-end*, usando o kit *NextSeq™ 500 Mid Output - 150* ciclos (Illumina), cujo rendimento esperado é de 260 milhões de sequências por corrida. Na abordagem de sequenciamento *paired-end* de 150 ciclos, são sequenciados até 75 nucleotídeos no sentido *forward* e até 75 nucleotídeos no sentido *reverse* além da região da sequência que identifica a origem de cada fragmento (*barcode*), garantindo que ambos sentidos de cada fragmento sejam sequenciados com alta qualidade.

O sequenciamento pela plataforma *NextSeq 500* baseia-se na hibridização das moléculas de cDNA fita simples a oligonucleotídeos que estão ligados à *flowcell*. Nesta, ocorre a amplificação por ponte dos fragmentos hibridizados. Na amplificação por ponte, a molécula molde de DNA é removida a partir de lavagens com tampões específicos, e a fita que foi amplificada faz uma “ponte” devido a diferentes temperaturas e reagentes adicionados, permitindo que a extremidade livre do

fragmento se ligue a um outro nucleotídeo complementar a esta região fixado na *flowcell* e iniciando um novo ciclo de amplificação que se repete, formando um conjunto de fragmentos clonais, os *clusters*, na mesma região da *flowcell*. A formação dos *clusters*, portanto, permite amplificar o sinal, possibilitando a leitura das fluorescências pelo sequenciador (Figura 4). A leitura dos fragmentos é realizada base a base, onde, em cada incorporação de um nucleotídeo, ocorre um bloqueio do sinal, fazendo com que o equipamento leia um nucleotídeo por vez. O equipamento NextSeq 500 é capaz de realizar a leitura dos nucleotídeos incorporados em duas cores: a incorporação de timinas (T) emite fluorescência de cor verde, e as citosinas (C) emite cor vermelha. Para adenina (A), ocorre emissão da cor amarela, que é a mistura das cores verde e vermelha, e, para a guanina (G), a identificação é feita pela ausência de fluorescência (BRONNER et al. 2014).



Fonte: [https://www.dkfz.de/gpcf/hiseq\\_technology.html](https://www.dkfz.de/gpcf/hiseq_technology.html)

**Figura 4** - Representação esquemática das etapas do sequenciamento na plataforma *Illumina*. 1: Biblioteca de cDNA simples fita delimitada em suas extremidades pelos adaptadores P5 (destacado pela cor lilás) e P7 (destacado pela cor azul), os quais se ligarão aos nucleotídeos fixados na *flowcell* (quadrado cinza); 2: Amplificação em ponte por PCR; 3: Formação dos clusters; 4: Representação da leitura das fluorescências dos nucleotídeos pelo equipamento HiSeq, o qual é capaz de identificar e ler quatro cores (o equipamento NextSeq, utilizado nesse estudo, é capaz de reconhecer e ler duas cores).

### 3.6 ANÁLISE DE BIOINFORMÁTICA

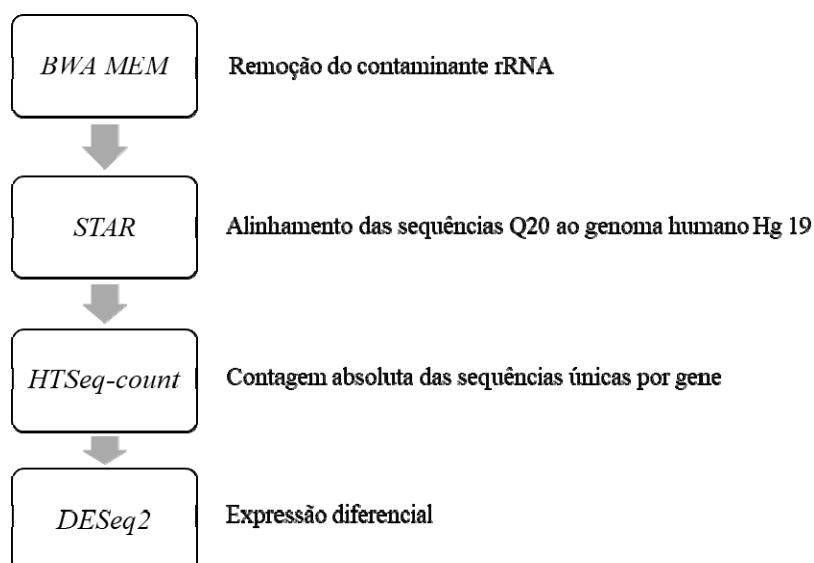
Após o sequenciamento das bibliotecas de cDNA, os arquivos FASTQ gerados pelo NextSeq 500 foram analisados pelo *pipeline* desenvolvido pela equipe do Laboratório de Biologia Computacional do CIPE, coordenada pelo Dr. Israel Tojal da Silva, e, resumidamente, descrito a seguir e representado na Figura 5.

Inicialmente, as sequências foram alinhadas contra sequências de RNA ribossomais (28S; 18S e 5S) utilizando o algoritmo *BWA MEM* 0.7.17 (LI, 2013). As sequências referentes aos rRNAs com MAPQ (qualidade de mapeamento)  $\geq 20$  foram removidas das análises seguintes. Após essa filtragem, as sequências foram alinhadas no genoma humano referência Hg 19 (*Genome Reference Consortium*

GRCh37), utilizando o programa *STAR 2.5.3a* (DOBIN et al. 2013) (do inglês, *Spliced Transcripts Alignment to a Reference*). Os arquivos gerados referentes ao alinhamento único das sequências foram convertidos do formato SAM para BAM utilizando o programa *SAMtools 1.9* (LI et al. 2009). A partir do mapeamento foi realizada a contagem de sequências únicas por gene e anotação dos genes utilizando o programa *HTSeq-count 0.9.1* (ANDERS et al. 2014) e anotações do banco de dados RefSeq (GRCh37.p13 de 2017-04-19) (O'LEARY et al. 2016). Com relação à contagem das sequências únicas por gene, identificamos aqueles que foram representados com pelo menos 1 sequência (1x), 10 sequências (10x), 50 sequências (50x) e 100 sequências (100x).

Para análise de genes diferencialmente expressos (GDEs) foi usado o programa *DESeq2 1.16.1* (ANDERS e HUBER 2010), o qual utiliza uma normalização interna baseada no tamanho da biblioteca de cada amostra. Após obtenção dos valores de expressão de cada gene, foram considerados como GDEs nas amostras tumorais em relação às amostras normais adjacentes àqueles cuja análise dos níveis de expressão apresentou aumentada ou diminuída em mais de 4 vezes (*fold change*  $\geq 4,0$  e  $\leq -4,0$ , respectivamente) e cuja análise estatística corrigida (pelo método *Benjamini–Hochberg*) (BENJAMINI e HOCHBERG 1995) apresentou menor ou igual a 0,05 (p valor ajustado  $\leq 0,05$ ).

Para a classificação dos GDEs em codificadores de proteína, não codificadores de proteína e pseudogene foi utilizado o banco de dados *Ensembl* (GRCh37 release 87) (AKEN et al. 2016).



**Figura 5** - Fluxograma das etapas do *pipeline* desenvolvido pela equipe do Laboratório de Biologia Computacional do CIPE, coordenada pelo Pesquisador Dr. Israel Tojal da Silva, para a análise das sequências geradas do RNA-Seq pela plataforma NextSeq 500 da Illumina.

Com o intuito de avaliar nossos dados de expressão diferencial em uma nova coorte de amostras de CMTN, realizamos uma análise de sobreposição com os dados de expressão diferencial das amostras de CMTN gerados por RNASeq pelo *The Cancer Genome Atlas (TCGA)* (The Cancer Genome Atlas Network 2012). Para isso, analisamos os GDEs concordantes, considerando toda casuística do nosso estudo (CMTN\_Estudo), bem como nos grupos separadamente (CMTN Hereditário\_Estudo e CMTN Esporádico\_Estudo), e os GDEs da casuística do TCGA, a qual separamos por grupo de idade (CMTN\_TCGA; CMTN\_TCGA < 40 anos; CMTN\_TCGA ≥ 41 ≤ 50 anos; CMTN\_TCGA > 51 anos). Os critérios para a filtragem dos GDEs do TCGA foram os mesmos utilizados neste estudo ( $fold\ change \geq e \leq |4|$  e p valor ajustado ≤ 0,05).

### 3.7 CLASSIFICAÇÃO MOLECULAR DAS AMOSTRAS DE CMTN BASEADA NA CLASSIFICAÇÃO DE LEHMANN

Para classificação molecular dos tumores TN, foi utilizada uma ferramenta online, *TNBCtype* (<http://cbc.mc.vanderbilt.edu/tNBC/>). Essa ferramenta correlaciona os dados de expressão gênica das amostras a serem analisadas com as assinaturas gênicas previamente estabelecidas por LEHMANN et al. (2011), que são baseadas em genes diferencialmente expressos e ontologias gênicas de cada subtipo de CMTN. Na pesquisa realizada por CHEN et al. (2012), foi observada que a expressão positiva do receptor de estrógeno (RE) avaliado por *Microarray* de amostras RE positivo foi falsamente identificada como negativa pela imunohistoquímica, o que afetou na normalização da expressão gênica de amostras de CMTN. Assim, considerando que a expressão positiva do RE a nível transcricional pode impactar na classificação molecular dos subtipos de CMTN, os autores desenvolveram o filtro RE-positivo que é baseado no cálculo do percentil de expressão do RE entre todos os genes para cada amostra. Amostras cuja expressão do RE está maior que o percentil 75 de expressão gênica de todos os genes são removidas pois são identificadas como amostras com expressão positiva do RE a nível transcricional. Essa etapa de controle de qualidade realizada pela ferramenta *TNBCtype* é uma efetiva abordagem para identificar e remover amostras RE positivo e, assim, garantir a acurácia da classificação molecular das amostras CMTN (CHEN et al. 2012). Logo, passando por este controle de qualidade, as amostras tornam-se aprovadas para serem classificadas nos 6 subtipos moleculares se tiverem um coeficiente de correlação com cada subtipo molecular  $> 0,2$  e p valor ajustado  $< 0,05$ ,



os quais são: imunomodulador (IM), *basal-like 1* (BL1), *basal-like 2* (BL2), mesenquimal (M), mesenquimal *stem like* (MSL), luminal receptor de andrógeno (RA). Amostras cujo coeficiente de correlação é  $< 0,2$  e sem significância estatística ( $p$  valor  $> 0,05$ ) são classificadas como instáveis (INS). Portanto, a partir dos valores de expressão normalizados de todos os genes identificados nas amostras tumorais da nossa casuística, realizamos a classificação molecular das mesmas através dessa ferramenta.

Realizamos também análise de sobrevida global pelo método de Kaplan-Meier, usando o test de Log-rank para calcular as curvas de sobrevida das amostras de CMTN em relação aos subtipos moleculares nos quais elas foram classificadas. Consideramos a sobrevida global como sendo o tempo, em dias, decorrido do dia do diagnóstico histopatológico até a data do óbito, por alguma causa, ou até a data do último segmento. Inicialmente, geramos as curvas de sobrevida global de cada subtipo molecular classificado para todas as amostras. Assim, a comparação foi feita entre as curvas de sobrevida de todas as amostras que foram classificadas no subtipo imunomodulador (IM), nos subtipos *basal-like 1* e *basal-like 2* (BL1 e BL2) e nos outros subtipos (M, MSL, LRA, INS). Comparamos também as curvas de sobrevida global das amostras classificadas no subtipo IM e nos outros subtipos (M, BL1, BL2, MSL, LRA, INS) considerando dois grupos de amostras: tumores TN com deficiência em *BRCA1* (TN *BRCA1*-mutado ou TN *BRCA1*-hipermetilado) e tumores TN com *BRCA1* proficiente (TN *BRCA1*-não-hipermetilado). Os resultados foram considerados estatisticamente significativos quando  $p$  valor  $\leq 0,05$ .

### 3.8 ANÁLISE FUNCIONAL *IN SILICO* DOS GENES DIFERENCIALMENTE EXPRESSOS (GDEs)

Para melhor compreensão da tumorigênese do CMTN, todos os GDEs nas amostras tumorais em relação às amostras normais adjacentes foram submetidos a uma análise funcional *in silico* através do programa *Ingenuity Pathway Analysis* (v8.0 IPA, *Sistemas Ingenuity, Redwood City, CA, EUA*; <http://www.ingenuity.com>), que utiliza o banco de dados *Knowledge Base* como repositório de dados biológicos. Com o objetivo de anotar os processos biológicos estatisticamente significativos, realizamos uma abordagem exploratória que incluiu o enriquecimento gênico em vias canônicas, redes de interação gênicas e análise dos reguladores *upstream*. Toda essa análise foi feita em colaboração com o Dr. Fábio A. Marchi do A.C. Camargo Cancer Center.

Para essa análise, os critérios para considerar os estados de ativação e inibição preditos em cada análise foram considerados significativos quando o *z-score* foi  $\geq 2$  e  $\leq -2$ , respectivamente. Resumidamente, as análises realizadas estão descritas a seguir:

- 1 Enriquecimento das vias canônicas: esta análise identificou as vias canônicas significativamente ativadas e inibidas a partir do conjunto de genes analisados;
- 2 Redes de interação gênica: esta análise mostrou as redes de interação biológica entre os genes identificados de acordo com dois critérios definidos pelo programa: genes focais e pontuação. Genes focais são os genes analisados do estudo e que foram identificados na rede, e a pontuação é definida por um valor (*p-score*) derivado do p-valor do enriquecimento da

rede para determinadas categorias biológicas, considerando o número de genes da rede e o número de genes analisados que constam na rede. Como os p-valores identificados são tipicamente baixos (por exemplo:  $10^{-8}$ ), a pontuação (*p-score*) é definida como  $-\log_{10}(\text{p-valor})$ . Assim, essa análise pode ajudar a entender as atividades biológicas ativadas e inibidas no processo da tumorigênese do CMTN;

- 3 Análise reguladores *upstream*: esta análise possibilitou a identificação de possíveis reguladores transcricionais que poderiam explicar as mudanças nos níveis de expressão dos genes analisados. Ela nos permite observar a relação regulador-alvo (*Target molecules in dataset*) bem como uma relação causal entre importantes reguladores com seus alvos (*Mechanistic network*) através da qual é possível prever cascatas de sinalização ativadas ou inibidas onde os reguladores com seus alvos estão envolvidos. Tudo isso é baseado no valor de expressão dos genes reguladores e dos alvos analisados no estudo.

## 4 RESULTADOS

### 4.1 SELEÇÃO DAS AMOSTRAS E EXTRAÇÃO DE RNA

Foram utilizados neste estudo 37 casos de pacientes diagnosticadas com CMTN em idade jovem ( $\leq 50$  anos), sendo 9 casos de CMTN Hereditário (com mutação patogênica germinativa em *BRCA1*) e 28 casos de CMTN Esporádico, sendo 9 casos *BRCA1*-deficiente (com o promotor de *BRCA1* hipermetilado) e 19 casos *BRCA1*-proficiente (com o promotor de *BRCA1* não hipermetilado), mostrados no Quadro 2 descritos na metodologia.

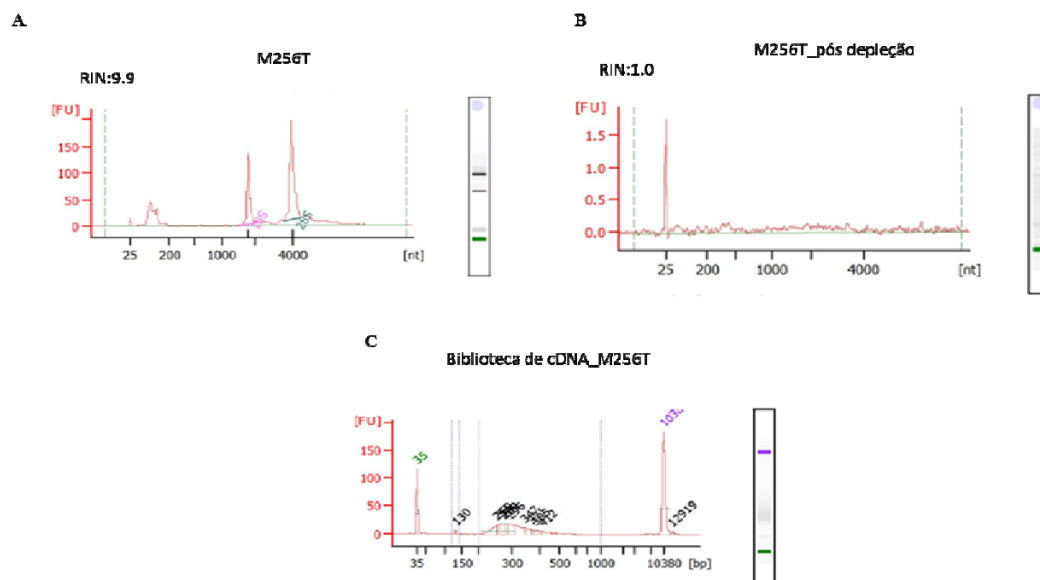
Como podemos observar no Quadro 3, na sessão 3.3 de Materiais e Métodos, a grande maioria das moléculas de RNA total foi de ótima integridade, como mostrado pelo  $RIN \geq 7$ , e puro (livre de contaminantes, como proteínas, sais, fenol, e outros compostos), como observado pelas razões 260/280 e 260/230 em torno de 2.0, valor de referência que representa um RNA puro (Thermo Fisher Scientific 2016).

### 4.2 SEQUENCIAMENTO DO RNA

#### 4.2.1 Construção das Bibliotecas de cDNA

Conforme descrito na sessão 3.3 de Materiais e Métodos, antes de seguirmos o protocolo da construção das bibliotecas, a eficiência da depleção do RNA ribossomal (rRNA) foi avaliada pela plataforma Agilent Bioanalyzer 2100.

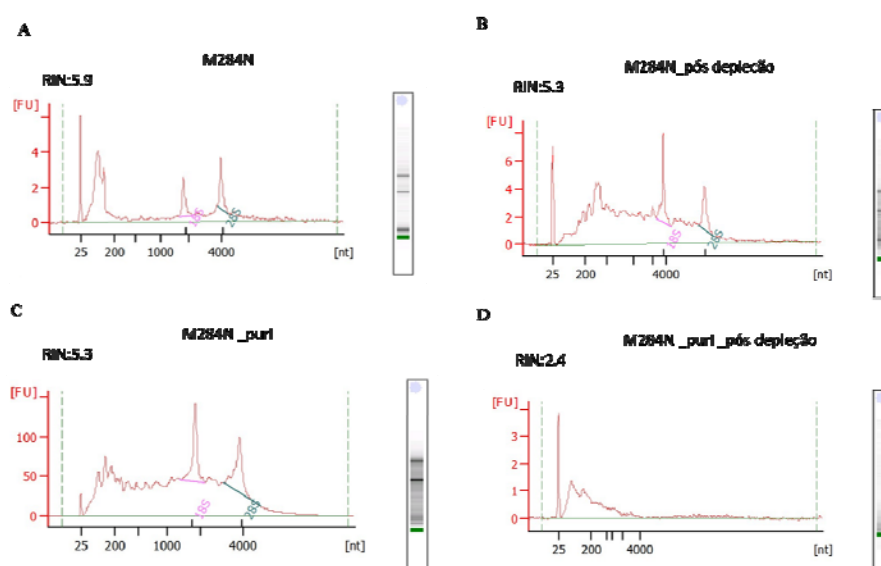
Para 45,2% (28/62) das amostras de RNA total, a depleção do rRNA foi eficiente, o que permitiu a construção das bibliotecas seguidamente (Figura 6).



**Figura 6** - Perfil da amostra de RNA tumoral ID\_M256T analisado pelo Agilent Bioanalyzer 2100, demonstrando as etapas de construção da biblioteca realizadas eficientemente, como exemplo. (A) Eletroferograma do RNA total inicial, revelando ótima relação entre os picos de rRNA 18s e 28s (RIN 9.9); (B) Eletroferograma do RNA total após etapa de depleção, mostrando depleção total do rRNA; (C) Eletroferograma do produto final da biblioteca de cDNA, mostrando maior porcentagem dos fragmentos de 255 nucleotídeos e baixa porcentagem (1%) de produtos com tamanho de 130 nucleotídeos (dímero de adaptador).

Porém, 54,8% (34/62) das amostras não apresentaram depleção total do rRNA. Para essas amostras, a etapa inicial da depleção foi repetida a partir do RNA total e, para a maioria delas (26/34), a depleção do rRNA foi eficiente. Contudo, para 8 amostras (destacadas no quadro 4 com um asterisco\*), a depleção do rRNA continuou ineficiente. Para estas amostras, realizamos uma purificação em coluna das moléculas de RNA total inicial antes de repetir o protocolo da depleção do rRNA com o objetivo de melhorar a qualidade das moléculas para obter a eficiência na depleção, como descrito na sessão 3.3 de Materiais e Métodos. Após essa

purificação, nova depleção do rRNA foi realizada e, através da avaliação pelo *Bioanalyzer*, as mesmas tiveram depleção total do rRNA. Após a verificação da eficiência da depleção do rRNA para todas as amostras de RNA total pelo *Agilent Bioanalyzer 2100*, seguimos com as etapas do protocolo da construção das bibliotecas de cDNA, como exemplificado na Figura 7.



**Figura 7** - Perfil da amostra de RNA total ID\_M284N analisado pelo *Agilent Bioanalyzer 2100* da não depleção total do rRNA, como exemplo. (A) Eletroferograma do RNA total inicial, revelando a relação entre os picos de rRNA 18s e 28s (RIN: 5.9); (B) Eletroferograma do RNA total após etapa de depleção, mostrando a não depleção total do rRNA pela presença dos picos de rRNA 18s e 28s (RIN:5.3); (C) Eletroferograma do RNA total após a purificação, mostrando a integridade do RNA total através da presença dos picos 18 s e 28s (RIN:5.3); (D) Eletroferograma do RNA total purificado após etapa de depleção, mostrando a depleção total do rRNA.

Como podemos observar no Quadro 4, as 62 bibliotecas finais de cDNA apresentaram, em média, fragmentos com tamanho médio de 274pb, próximo do esperado conforme protocolo (260 pb). O perfil do tamanho dos fragmentos de algumas bibliotecas está representado na Figura 8, como exemplo.

**Quadro 4** - Dados das bibliotecas finais de cDNA construídas a partir do RNA total de todas as amostras tumorais e normais.

<b>Amostra</b>	<b>Adaptador</b>	<b>Concentração da biblioteca final (ng/uL - Qubit)</b>	<b>Tamanho médio da biblioteca final (pb - Bioanalyzer)</b>	<b>Molaridade biblioteca final (nM - Qubit)</b>	<b>Molaridade biblioteca final (nM - qPCR)</b>
ID_M 256N	AR002	4,38	247	21,26	44,68
ID_M 256T	AR015	0,12	255	6,02	3,76
ID_M 280N	AR013	4,99	293	23,99	48,00
ID_M 280T	AR016	0,25	252	1,09	3,38
ID_M 282N	AR006	11,10	260	51,44	128,80
ID_M 282T*	AR007	26,50	239	127,40	222,54
ID_M 287N*	AR004	3,39	257	17,70	10,18
ID_M 287T	AR012	0,47	247	2,42	23,55
ID_M 292N	AR014	3,98	252	0,56	13,84
ID_M 292T	AR019	1,41	294	6,98	24,21
ID_M 303N	AR002	6,93	293	32,50	61,69
ID_M 303T	AR013	3,56	323	17,01	21,16
ID_M 355N	AR018	1,4	311	6,90	12,47
ID_M 355T*	AR019	1,65	257	8,52	12,26
ID_MJ 2026N	AR013	1,04	266	5,06	7,78
ID_MJ 2026T	AR016	0,30	244	1,54	7,84
ID_MJ 2021T	AR004	43,80	277	200,55	175,91
ID_M 15N	AR002	13,50	258	80,90	69,91
ID_M 15T	AR004	12,00	318	55,11	84,32
ID_M 299N	AR004	37,40	258	155,09	347,54
ID_M 299T	AR007	6,88	255	32,67	69,72
ID_M 365N	AR006	30,50	293	139,24	236,83
ID_M 365T	AR007	16,20	292	79,88	101,54
ID_M 377N*	AR002	8,11	294	40,12	66,62
ID_M 377T*	AR004	9,90	290	48,71	66,19

**Cont/ Quadro 4**

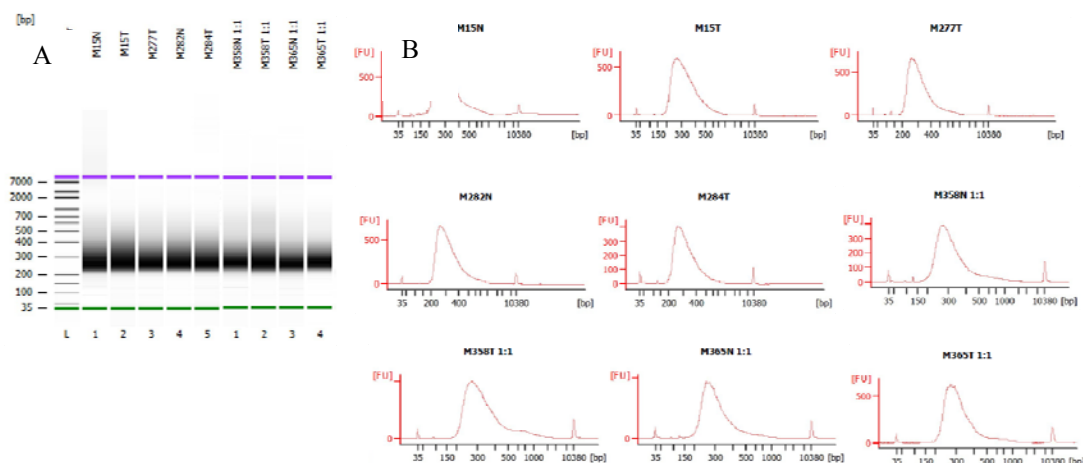
<b>Amostra</b>	<b>Adaptador</b>	<b>Concentração da biblioteca final (ng/uL - Qubit)</b>	<b>Tamanho médio da biblioteca final (pb - Bioanalyzer)</b>	<b>Molaridade biblioteca final (nM - Qubit)</b>	<b>Molaridade biblioteca final (nM - qPCR)</b>
ID_M 347T	AR007	27.4	275	130,10	166,09
ID_M 357T	AR014	29.1	298	134,85	132,06
ID_M 366T	AR015	15.1	253	75,42	64,85
ID_M 379T	AR016	38.9	275	178,11	194,60
ID_M 385T	AR018	34.7	279	156,55	184,00
ID_M7N	AR016	3,27	259	16,18	22,13
ID_M7T	AR013	2,86	257	14,47	16,36
ID_M214N	AR006	6,39	240	33,10	62,94
ID_M214T	AR007	9,35	288	46,55	66,90
ID_M269N	AR015	2,28	270	11,97	10,93
ID_M269T	AR019	6,86	249	34,72	40,00
ID_M 275N	AR014	2,84	279	14,09	15,03
ID_M 275T	AR018	0,30	258	1,48	3,85
ID_M278N	AR005	6,17	250	31,12	43,64
ID_M278T	AR012	5,75	247	29,39	34,46
ID_M 284N*	AR002	10,80	272	53,25	97,50
ID_M 284T	AR007	6.11	271	29,01	60,42
ID_M 286N	AR002	11,00	260	53,22	38,28
ID_M 286T	AR006	5,70	247	28,20	54,79
ID_M 290N	AR013	2,02	268	9,80	18,54
ID_M 290T	AR014	0,75	294	3,70	4,98
ID_M 291N*	AR005	5,95	271	30,61	39,32
ID_M 291T	AR015	0.74	280	3,58	22,74
ID_M 298N	AR005	2,88	235	13,55	21,16
ID_M 298T	AR012	12,00	291	56,29	135,15



**Cont/ Quadro 4**

<b>Amostra</b>	<b>Adaptador</b>	<b>Concentração da biblioteca final (ng/uL - Qubit)</b>	<b>Tamanho médio da biblioteca final (pb - Bioanalyzer)</b>	<b>Molaridade biblioteca final (nM - Qubit)</b>	<b>Molaridade biblioteca final (nM - qPCR)</b>
ID_M 312N*	AR006	4,97	260	25,40	58,52
ID_M 312T	AR016	0,74	248	3,77	18,82
ID_M 358N	AR019	12,40	275	57,12	81,06
ID_M 358T	AR005	16,60	316	73,60	69,49
ID_MJ 2003N	AR015	1,00	291	4,93	8,70
ID_MJ 2003T	AR019	0,14	305	0,68	5,77
ID_M 217T	AR005	17,6	296	82,05	79,80
ID_M277T	AR006	10,50	313	49,25	89,59
ID_M 283T	AR006	2,56	278	12,46	11,24
ID_M 348T	AR012	26,4	296	119,46	100,11
ID_M 351T	AR013	17,5	315	80,61	65,14
ID_MIC 20T	AR019	35,5	275	156,50	159,36

\*amostras de RNA total que foram purificadas para obter eficiente depleção do rRNA durante etapa da construção da biblioteca.



**Figura 8** - Perfil da distribuição dos fragmentos de 9 bibliotecas finais de cDNA analisado pelo Agilent Bioanalyzer 2100, como exemplo. (A) Imagem digital do gel, mostrando a maior porcentagem das bandas dos fragmentos em torno de 260 pb. (B) Eletroferogramas mostrando o perfil de distribuição dos fragmentos por tamanho (eixo X) e quantidade por unidades de fluorescência (eixo Y).

#### 4.2.2 Sequenciamento das bibliotecas de cDNA na plataforma *NEXTSEQ 500*

Após análise quantitativa por qPCR e qualitativa pelo *Agilent Bioanalyzer 2100* das 62 bibliotecas finais (Quadro 4), as mesmas foram sequenciadas na plataforma *NextSeq 500*, totalizando 11 corridas de sequenciamento, pois para algumas bibliotecas o sequenciamento foi repetido com o objetivo de aumentar o número de sequências mapeadas por amostra.

O sequenciamento das 62 bibliotecas de cDNA gerou, em média, 48.627.204 milhões de sequências por amostra. Após filtragem para retirada das sequências referentes ao rRNA e análise de qualidade, 33.611.903 milhões de sequências sem contaminantes e  $Q \geq 20$  foram obtidas, em média, por amostra, das quais 79,5% (26.736.816 / 33.611.903) foram mapeadas no genoma humano referência versão Hg 19 com o *software* STAR. Através do programa *HTSeq-count*, identificamos 18.928 genes expressos, em média, por amostra, com pelo menos 1 sequência única mapeada, de acordo com o bando de dado *RefSeq* (Tabela 1 e Figura 9A e B).

**Tabela 1** - Dados do sequenciamento das bibliotecas de cDNA gerado pela plataforma NextSeq 500.

Amostra	Adaptador	Total de sequências geradas	Total de sequências livres rRNA	% livre rRNA	Total de sequências mapeadas	% map.	Total de genes >= 1x	Total de genes >= 10x	Total de genes >= 20x	Total de genes >= 50x	Total de genes >= 100x
ID_M256N	AR006	74.987.072	42.050.782	56,1	27.935.216	66,4	19.553	15.111	13.592	10.980	8.434
ID_M256T	AR007	35.806.248	32.037.458	89,5	26.251.490	81,9	18.785	15.084	13.616	11.537	9.450
ID_M280N	AR013	78.709.778	39.662.222	50,4	26.956.622	68,0	20.090	15.476	13.799	11.121	8.335
ID_M280T	AR016	50.430.580	27.114.646	53,8	23.021.794	84,9	19.382	15.262	13.805	11.661	9.532
ID_M282N	AR018	39.092.374	35.947.986	92,0	31.086.414	86,5	20.360	15.900	14.421	12.256	10.098
ID_M282T	AR019	82.435.610	17.474.968	21,2	15.064.520	86,2	19.011	14.203	12.803	10.505	8.096
ID_M287N	AR013	63.132.808	29.615.920	46,9	25.577.320	86,4	19.101	15.477	14.081	12.001	9.723
ID_M287T	AR005	47.090.388	36.626.284	77,8	23.460.146	64,1	13.914	13.326	12.212	9.563	7.109
ID_M292N	AR006	54.341.496	45.234.938	83,2	24.798.524	54,8	18.300	13.956	12.189	9.114	6.385
ID_M292T	AR007	43.716.946	41.480.166	94,9	31.915.830	76,9	17.499	14.728	13.575	12.093	10.716
ID_M303N	AR005	69.137.234	48.273.604	69,8	42.437.758	87,9	20.343	16.517	15.084	13.030	11.020
ID_M303T	AR012	86.147.500	80.796.270	93,8	69.342.214	85,8	19.055	16.203	15.007	13.481	12.298
ID_M355N	AR015	39.699.436	37.988.672	95,7	33.052.226	87,0	16.396	16.060	14.543	12.247	10.007
ID_M355T	AR019	42.001.748	39.519.136	94,1	32.235.236	81,6	18.425	15.265	13.984	12.294	10.811
ID_MJ2026 N	AR002	100.825.210	31.624.408	31,4	25.443.596	80,5	20.217	15.198	13.543	10.961	8.300

**Cont/ Tabela 1**

<b>Amostra</b>	<b>Adaptador</b>	<b>Total de sequências geradas</b>	<b>Total de sequências livres rRNA</b>	<b>% livre rRNA</b>	<b>Total de sequências mapeadas</b>	<b>% map.</b>	<b>Total de genes &gt;= 1x</b>	<b>Total de genes &gt;= 10x</b>	<b>Total de genes &gt;= 20x</b>	<b>Total de genes &gt;= 50x</b>	<b>Total de genes &gt;= 100x</b>
ID_MJ2026T	AR006	106.928.898	73.015.240	68,3	59.394.818	81,3	17.257	16.951	15.530	13.618	11.982
ID_MJ2021T	AR004	16.571.851	12.485.181	75,0	11.429.032	92,0	19.419	14.898	13.626	11.893	10.059
ID_M15N	AR002	41.331.216	18.465.602	44,7	13.113.078	71,0	18.505	12.872	10.719	6.898	3.862
ID_M15T	AR015	31.537.812	28.744.108	91,1	24.573.506	85,5	19.532	14.904	13.514	11.602	9.672
ID_M299N	AR014	204.390.928	14.996.502	7,3	4.119.976	27,5	16.808	10.275	7.562	3.625	1.548
ID_M299T	AR018	53.856.778	39.330.132	73,0	27.992.466	71,2	19.472	14.963	13.595	11.645	9.782
ID_M365N	AR005	55.853.514	19.140.742	34,3	15.325.504	80,1	19.111	13.914	12.052	8.720	5.642
ID_M365T	AR015	67.321.422	43.422.838	64,5	29.169.998	67,2	19.341	14.497	13.076	11.024	8.928
ID_M377N	AR005	37.355.300	32.888.790	88,0	27.163.456	82,6	19.719	15.756	14.224	11.966	9.696
ID_M377T	AR012	28.490.176	23.432.642	82,2	19.939.538	85,1	18.909	14.582	13.359	11.572	9.763
ID_M347T	AR007	67.192.140	60.249.470	89,7	51.688.667	85,8	20.791	17.496	16.349	14.736	13.516
ID_M357T	AR014	27.846.659	24.298.092	87,3	20.685.687	85,1	20.536	16.151	14.822	13.043	11.558
ID_M366T	AR015	40.625.197	36.125.316	88,9	23.579.337	65,3	17.694	13.352	11.980	9.821	7.619
ID_M379T	AR016	25.041.327	23.365.698	93,3	20.311.828	86,9	20.680	16.470	15.052	13.217	11.697
ID_M385T	AR018	25.565.595	23.693.409	92,7	21.095.739	89,0	20.423	16.213	14.900	13.074	11.507
ID_M7N	AR006	34.007.418	32.493.698	95,5	28.374.744	87,3	19.987	15.805	14.221	11.904	9.463

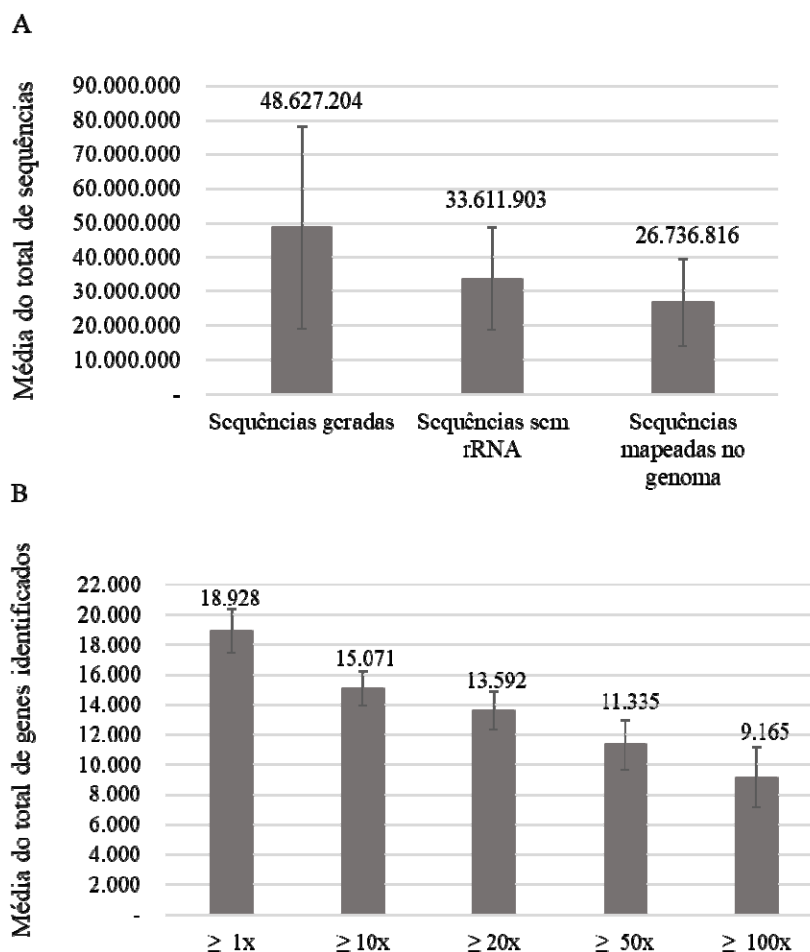
**Cont/ Tabela 1**

<b>Amostra</b>	<b>Adaptador</b>	<b>Total de sequências geradas</b>	<b>Total de sequências livres rRNA</b>	<b>% livre rRNA</b>	<b>Total de sequências mapeadas</b>	<b>% map.</b>	<b>Total de genes &gt;= 1x</b>	<b>Total de genes &gt;= 10x</b>	<b>Total de genes &gt;= 20x</b>	<b>Total de genes &gt;= 50x</b>	<b>Total de genes &gt;= 100x</b>
ID_M7	AR007	39.086.862	32.997.898	84,4	29.483.850	89,4	18.855	14.901	13.542	11.620	10.029
ID_M214N	AR013	32.286.954	28.937.912	89,6	25.079.368	86,7	19.934	15.441	13.913	11.614	9.314
ID_M214T	AR016	40.569.356	35.454.230	87,4	28.243.852	79,7	20.078	15.563	14.094	11.952	9.880
ID_M269N	AR004	51.651.190	46.646.002	90,3	33.932.214	72,7	18.817	15.310	14.020	12.107	10.216
ID_M269T	AR012	41.170.798	39.965.600	97,1	32.306.410	80,8	18.796	14.817	13.654	12.183	10.783
ID_M275N	AR014	32.269.844	31.154.864	96,5	24.129.274	77,4	19.750	15.195	13.636	11.056	8.444
ID_M275T	AR019	28.387.588	17.935.544	63,2	13.615.184	75,9	15.472	13.843	12.127	9.119	6.125
ID_M278N	AR002	41.242.244	35.451.716	86,0	29.480.146	83,2	19.908	15.213	13.611	10.817	8.067
ID_M278T	AR013	48.249.254	46.501.916	96,4	39.437.480	84,8	20.291	16.206	14.799	12.852	11.078
ID_M284N	AR002	30.933.674	28.037.672	90,6	24.611.918	87,8	19.894	15.186	13.561	10.895	8.223
ID_M284T	AR004	46.598.382	44.536.964	95,6	37.917.172	85,1	19.619	15.709	14.375	12.746	11.259
ID_M286N	AR004	68.957.390	28.558.346	41,4	17.985.960	63,0	19.526	14.500	12.612	9.524	6.401
ID_M286T	AR007	82.016.376	42.551.580	51,9	22.104.304	51,9	17.605	13.355	11.998	9.872	7.662
ID_M290N	AR002	45.877.842	32.839.266	71,6	28.450.328	86,6	18.269	15.606	14.001	11.377	8.734
ID_M290T	AR004	40.663.154	31.891.224	78,4	26.593.674	83,4	15.484	14.619	13.381	11.617	10.020
ID_M291N	AR016	34.541.232	27.953.542	80,9	22.898.036	81,9	19.688	15.143	13.618	11.233	8.757

**Cont/ Tabela 1**

<b>Amostra</b>	<b>Adaptador</b>	<b>Total de seqüências geradas</b>	<b>Total de seqüências livre rRNA</b>	<b>% livre rRNA</b>	<b>Total de seqüências mapeadas</b>	<b>% map.</b>	<b>Total de genes &gt; = 1x</b>	<b>Total de genes &gt; = 10x</b>	<b>Total de genes &gt; = 20x</b>	<b>Total de genes &gt; = 50x</b>	<b>Total de genes &gt; = 100x</b>
ID_M291T	AR013	30.634.952	25.122.442	82,0	20.885.004	83,1	16.840	14.975	13.459	11.138	8.637
ID_M298N	AR015	93.185.426	64.059.020	68,7	48.874.450	76,3	19.788	16.617	15.133	13.005	11.150
ID_M298T	AR019	59.275.816	35.568.148	60,0	24.416.084	68,6	19.981	15.269	13.696	11.407	8.926
ID_M312N	AR002	38.192.916	32.725.256	85,7	28.643.132	87,5	20.146	15.945	14.352	12.111	9.710
ID_M312T	AR007	22.168.008	17.845.666	80,5	11.651.266	65,3	15.461	13.066	11.292	8.127	5.092
ID_M358N	AR013	46.361.322	41.422.644	89,3	36.639.710	88,5	20.643	16.072	14.412	11.753	9.065
ID_M358T	AR014	78.945.142	74.185.276	94,0	66.581.814	89,8	20.383	16.200	14.863	13.051	11.441
ID_MJ2003N	AR019	41.413.612	40.305.460	97,3	34.468.812	85,5	20.060	16.445	15.121	13.238	11.618
ID_MJ2003T	AR005	39.513.042	36.310.672	91,9	29.380.742	80,9	16.237	15.224	13.803	11.277	8.836
ID_M217T	AR005	30.063.975	20.090.531	66,8	16.271.512	81,0	19.563	15.184	13.709	11.490	9.272
ID_M277T	AR006	10.707.586	10.102.492	94,3	9.071.461	89,8	18.748	14.240	12.922	10.842	8.595
ID_M283T	AR006	11.938.249	9.279.568	77,7	8.129.116	87,6	17.935	14.400	12.979	10.753	8.325
ID_M348T	AR012	11.114.204	10.281.923	92,5	8.813.350	85,7	19.307	14.834	13.478	11.487	9.276
ID_M351T	AR013	11.300.523	10.195.521	90,2	8.840.935	86,7	18.747	14.143	12.804	10.660	8.285
ID_MIC20T	AR019	14.099.048	13.434.168	95,3	12.209.732	90,9	19.082	14.345	12.887	10.650	8.373

Total de seqüências geradas: número total das seqüências geradas pelo *NextSeq 500*; Total de seqüências livres rRNA: número total de seqüências sem o contaminante RNA ribossomal (rRNA); % livre rRNA: porcentagem das seqüências geradas que não tinham seqüências referente ao rRNA; Total de seqüências mapeadas: número total de seqüências livres de rRNA e mapeadas contra o genoma humano referência (Hg19); % map.: porcentagem de seqüências livres de rRNA que foram mapeadas; Total de genes  $\geq$  1x: número de genes identificados com pelo menos 1 seqüência única mapeada (1x); Total de genes  $\geq$  10x: número de genes identificados com pelo menos 10 seqüências únicas mapeadas (10x); Total de genes  $\geq$  20x: número de genes identificados com pelo menos 20 seqüências únicas mapeadas (20x); Total de genes  $\geq$  50x: número de genes identificados com pelo menos 50 seqüências únicas mapeadas (50x); Total de genes  $\geq$  100x: número de genes identificados com pelo menos 100 seqüências únicas mapeadas (100x).



**Figura 9** - Dados do sequenciamento das 62 bibliotecas finais de cDNA gerados pela plataforma *NextSeq 500*. (A) Média do total das seqüências geradas após cada análise. (B) Média do total de genes que foram identificados com pelo menos 1 seqüência única mapeada ( $\geq 1x$ ); com pelo menos 10 seqüências únicas mapeadas ( $\geq 10x$ ); 20 seqüências únicas mapeadas ( $\geq 20x$ ); 50 seqüências únicas mapeadas ( $\geq 50x$ ) e com pelo menos 100 seqüências únicas mapeadas ( $\geq 100x$ ).

#### 4.3 CLASSIFICAÇÃO MOLECULAR DAS AMOSTRAS DE CMTN

A classificação molecular das 37 amostras tumorais de CMTN foi realizada utilizando a ferramenta *TNBCType* através do dado de expressão normalizado por FPKM (do inglês, *Fragments Per Kilobase Million*) de todos os genes identificados. Esta medida informa o valor de expressão de cada gene normalizado a partir do

número de sequências mapeadas por milhão (profundidade do sequenciamento) e do tamanho do gene em Kilobase (WANG et al. 2009).

Das 37 amostras de CMTN que foram submetidas à classificação molecular, 94,6% (35/37) passaram pelo controle de qualidade da filtragem de expressão do RE e, então, foram classificadas nos subtipos de CMTN, conforme descrito na sessão 3.7 do Materiais e Métodos. Porém, 5,4% (2/37) das amostras (ID\_M275T e ID\_M348T) não passaram por esse filtro sendo descartadas da análise. Essas duas amostras são do grupo *BRCA1*-proficiente.

Das 35 amostras classificadas, 80% (28/35) apresentaram coeficiente de correlação  $> 0.2$  e significância estatística (pvalor ajustado  $< 0.05$ ) para os 6 subtipos moleculares: imunomodulador (IM), mesenquimal (M), mesenquimal *stem-like* (MSL), *basal-like 1* (BL1) e *basal-like 2* (BL2) e luminal receptor de andrógeno (LRA). E 20% (7/35) das amostras não apresentaram coeficiente de correlação e nem significância estatística para nenhum dos subtipos, sendo, portanto, classificadas como instáveis (INS). Como apresentado no Quadro 5 e na Figura 10A, as amostras foram classificadas molecularmente nos subtipos IM, M, MSL, BL1/2 e LRA, apresentando, aproximadamente, os percentuais em ordem decrescente de 23 23 14, 9, 6 e 6%, respectivamente.

Em seguida, foi investigada a sobrevida global em 11 anos (4000 dias) das amostras de acordo com os subtipos classificados. Para essa análise e por uma questão do pequeno número de casos, os subtipos BL1 e BL2 (BL1/2) e M e MSL (M/MSL) foram agrupados. Assim, embora não termos encontrado significância estatística, observamos uma tendência de melhor sobrevida global para as classes IM e BL1/2 quando comparada com as demais classes (Figura 11A).



A classificação molecular das amostras de acordo com o status de *BRCA1*, *BRCA1*-deficiente (com mutação germinativa em *BRCA1* ou hipermetilação no promotor de *BRCA1*) e *BRCA1*-proficiente (negativos para mutação em *BRCA1/2* e com o promotor de *BRCA1* não hipermetilado), apresentada nos Quadros 6 e 7, respectivamente, mostrou que o primeiro grupo (18 amostras *BRCA1*-deficiente) apresenta maior percentagem de amostras classificadas como instáveis (INS), com 33% contra 6% do segundo grupo (17 amostras *BRCA1*-proficiente). Observamos, ainda, que no grupo *BRCA1*-deficiente, as outras classes moleculares foram representadas em IM, *BL1/2*, M e MSL em ordem decrescente de percentual, aproximadamente, em 28 17 11 e 11%, respectivamente (Quadro 6 e Figura 10B). Da mesma forma, as amostras do grupo *BRCA1*-proficiente foram classificadas em M, IM, MSL, *BL1/2* e LRA na mesma ordem decrescente de percentual, aproximadamente, em 35 18 18 12 e 12%, respectivamente (Quadro 7 e Figura 10C).

Comparando os dois grupos de amostras de acordo com o status de *BRCA1*, observa-se quem, entre os subtipos com coeficiente de correlação  $> 0,2$  e com significância estatística, o subtipo molecular M está enriquecido no grupo de amostras *BRCA1*-proficiente enquanto que o IM está mais enriquecido no grupo de amostras *BRCA1*-deficiente. Os demais subtipos moleculares estão representados de forma mais ou menos equilibrada nos dois grupos de amostras de acordo com o status de *BRCA1*. Vale destacar que o subtipo LRA, com poucas amostras classificadas, só foi representado no grupo de amostras *BRCA1*-proficiente (Figura 10C).

A análise de sobrevida global foi também realizada de acordo com o *status* de *BRCA1*. Observamos que a curva de sobrevida das amostras *BRCA1*-deficiente foi semelhante à curva de sobrevida das amostras *BRCA1*-proficiente (Figura 11B).

Realizamos, também, a análise de sobrevida global do subtipo IM e dos demais de acordo com o status de *BRCA1* (Figura 12 A e B). Para essa análise, nós consideramos o subtipo IM contra os demais que denominamos de grupo não-imunomodulador nos dois grupos de amostras, de acordo com o status de *BRCA1*: *BRCA1*-deficiente (5 amostras classificadas no perfil IM contra 13 classificadas no perfil não-imunomodulador) (Quadro 6 e Figura 12A) e *BRCA1*-proficiente (3 amostras classificadas no perfil imunomodulador contra 14 classificadas no perfil não-imunomodulador) (Quadro 7 e Figura 12B). Os resultados mostram que o subtipo molecular IM parece ter uma tendência de melhor sobrevida global em relação aos demais subtipos no grupo *BRCA1*-proficiente, apesar de ser sem significância estatística. Porém, o mesmo não foi observado no grupo *BRCA1*-deficiente, onde observamos que as amostras apresentaram tendência de melhor sobrevida global, independente do subtipo classificado (Figura 12).

**Quadro 5** - Amostras de CMTN baseada na classificação de LEHMANN et al. (2011) pela ferramenta *TNBCtype* (CHEN et al. 2012).

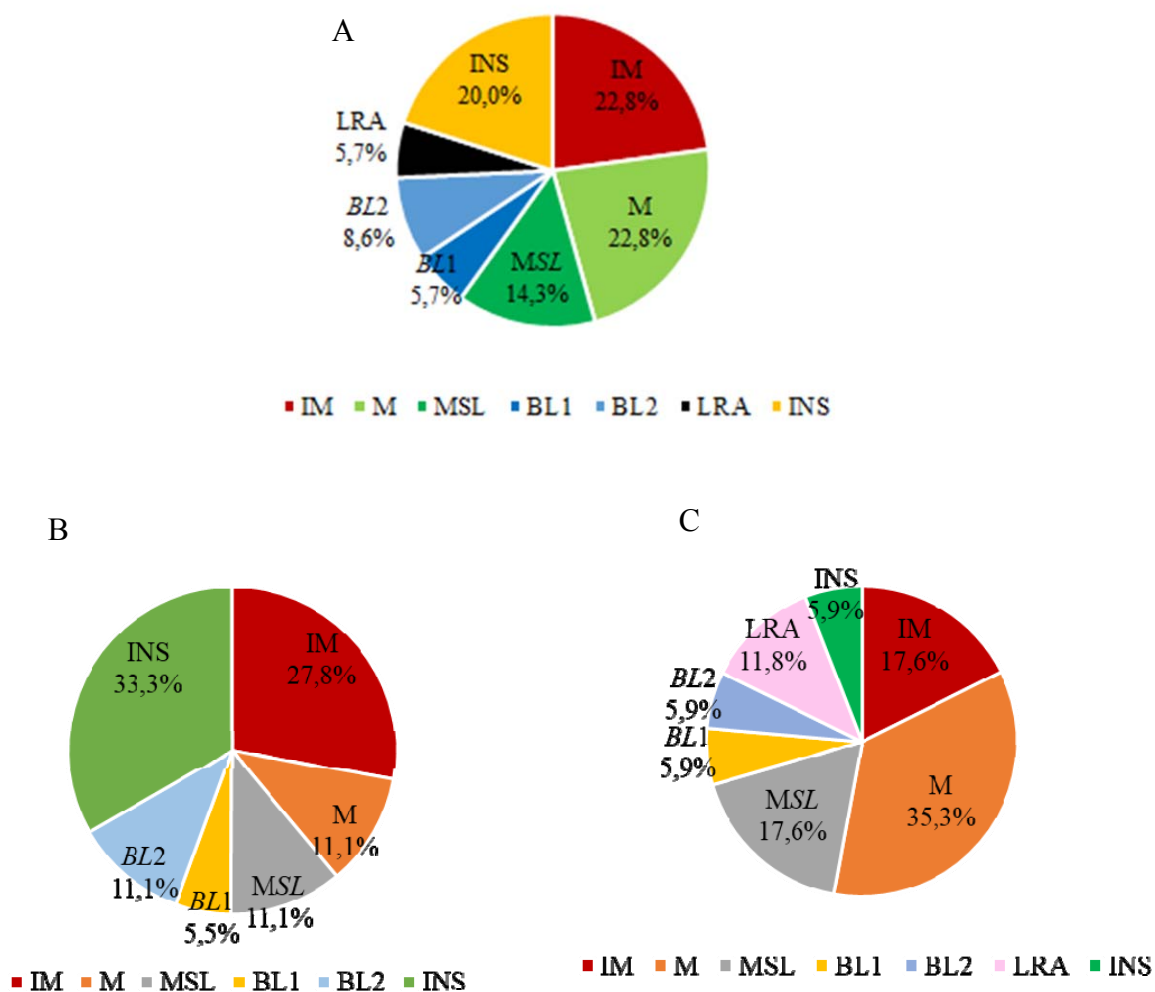
<b>Amostra</b>	<b>Subtipo molecular</b>	<b>Correlação</b>	<b>p valor</b>
ID_M214T	IM	0.428	<0.001
ID_M278T	IM	0.551	<0.001
ID_M282T	IM	0.67	<0.001
ID_M284T	IM	0.514	<0.001
ID_M292T	IM	0.625	<0.001
ID_M299T	IM	0.531	<0.001
ID_M347T	IM	0.284	<0.001
ID_MJ2026T	IM	0.658	<0.001
ID_M290T	M	0.659	<0.001
ID_M358T	M	0.475	<0.001
ID_M7T	M	0.537	<0.001
ID_MJ2003T	M	0.666	<0.001
ID_M217T	M	0.288	<0.001
ID_M283T	M	0.688	<0.001
ID_M366T	M	0.148	<0.001
ID_M385T	M	0.632	<0.001
ID_M287T	MSL	0.244	0.002
ID_M291T	MSL	0.674	<0.001
ID_M298T	MSL	0.404	<0.001
ID_M312T	MSL	0.585	<0.001
ID_MJ2021T	MSL	0.427	<0.001
ID_M269T	BL1	0.51	<0.001
ID_M280T	BL1	0.328	<0.001
ID_M277T	BL2	0.424	<0.001
ID_M357T	BL2	0.363	<0.001
ID_M379T	BL2	0.239	<0.001
ID_M351T	LRA	0.804	<0.001
ID_MIC20T	LRA	0.474	<0.001
ID_M15T	INS	NA	NA
ID_M256T	INS	NA	NA
ID_M286T	INS	NA	NA
ID_M303T	INS	NA	NA
ID_M355T	INS	NA	NA
ID_M365T	INS	NA	NA
ID_M377T	INS	NA	NA

**Quadro 6** - Amostras CMTN *BRCA1*-deficiente baseada na classificação de LEHMANN et al. (2011) pela ferramenta *TNBCtype* (CHEN et al 2012).

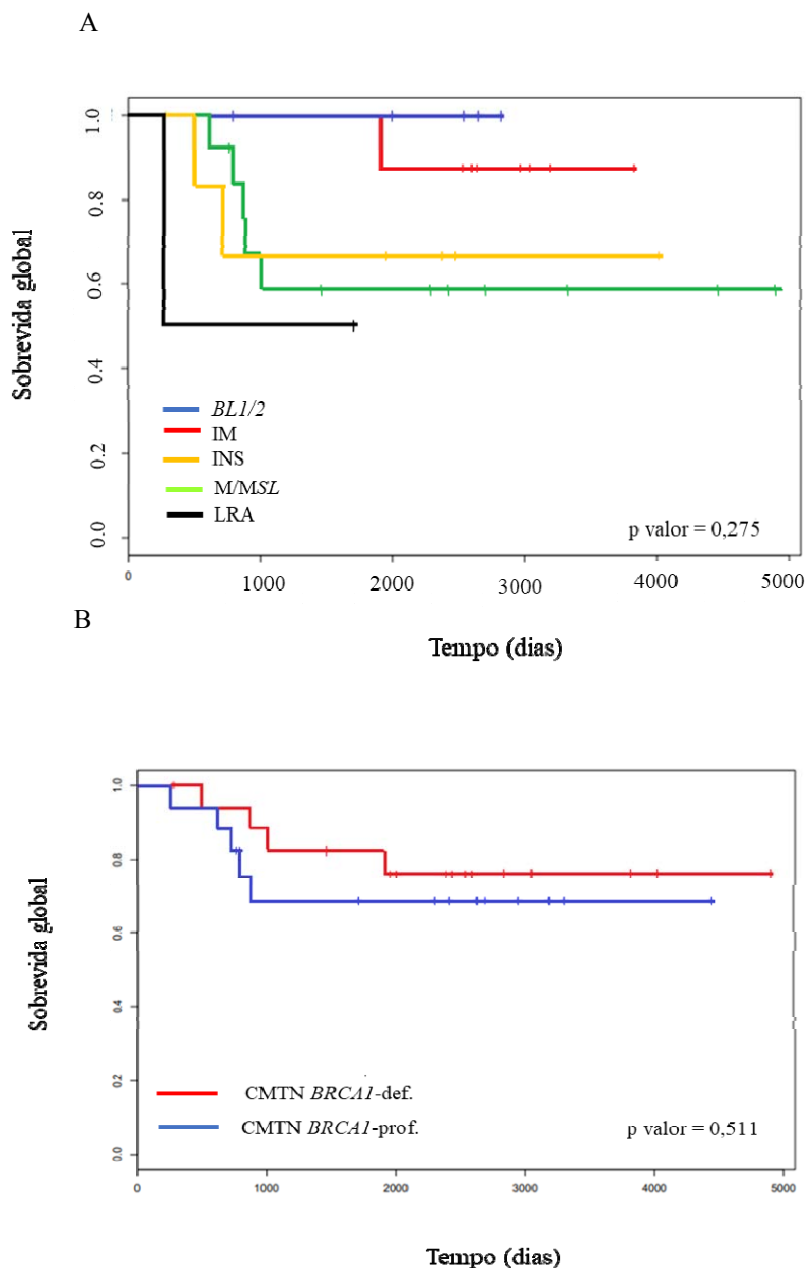
Amostra CMTN <i>BRCA1</i> -deficiente	Subtipo molecular	Correlação	p valor
M282T	IM	0.67	<0.001
M292T	IM	0.625	<0.001
MJ2026T	IM	0.658	<0.001
M299T	IM	0.531	<0.001
M347T	IM	0.284	<0.001
M366T	M	0.148	<0.001
M385T	M	0.632	<0.001
M287T	MSL	0.244	0.002
MJ2021T	MSL	0.427	<0.001
M280T	BL1	0.328	<0.001
M357T	BL2	0.363	<0.001
M379T	BL2	0.239	<0.001
M303T	INS	NA	NA
M355T	INS	NA	NA
M256T	INS	NA	NA
M15T	INS	NA	NA
M365T	INS	NA	NA
M377T	INS	NA	NA

**Quadro 7** - Amostras de CMTN *BRCA1*-proficiente baseada na classificação de LEHMANN et al. (2011) pela ferramenta *TNBCtype* (CHEN et al 2012).

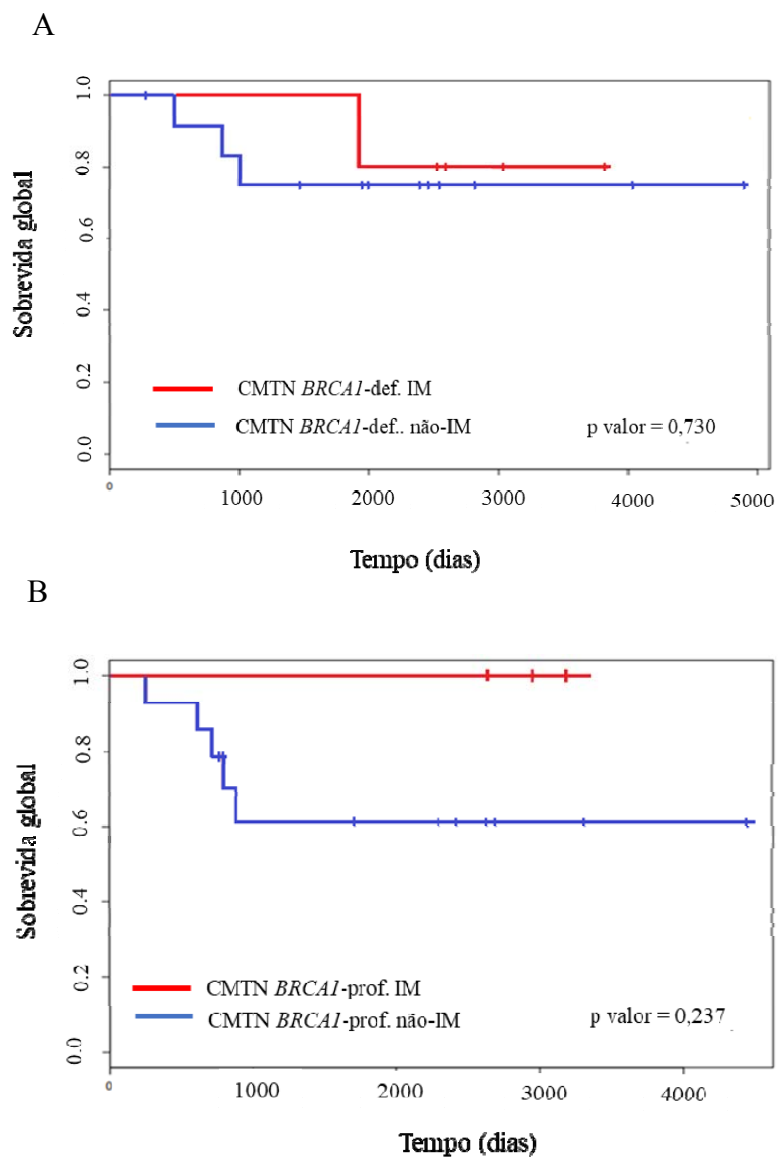
Amostra CMTN <i>BRCA1</i> -proficiente	Subtipo molecular	Correlação	p valor
M214T	IM	0.428	<0.001
M278T	IM	0.551	<0.001
M284T	IM	0.514	<0.001
M290T	M	0.659	<0.001
M358T	M	0.475	<0.001
M7T	M	0.537	<0.001
MJ2003T	M	0.666	<0.001
M217T	M	0.288	<0.001
M283T	M	0.688	<0.001
M291T	MSL	0.674	<0.001
M298T	MSL	0.404	<0.001
M312T	MSL	0.585	<0.001
M269T	BL1	0.51	<0.001
M277T	BL2	0.424	<0.001
M351T	LRA	0.804	<0.001
MIC20T	LRA	0.474	<0.001
M286T	INS	NA	NA



**Figura 10** – Distribuição dos subtipos moleculares nos quais as 35 amostras de CMTN foram classificadas. (A) Classificação molecular de todas as amostras de CMTN. (B) Classificação molecular de todas as amostras de CMTN *BRCA1*-deficiente. (C) Classificação molecular de todas as amostras de CMTN *BRCA1*-proficiente. Cada cor representa um subtipo molecular, sendo a cor vermelha o subtipo imunomodulador (IM); as cores verde escuro e claro os subtipos mesenquimal *stem like* (MSL) e mesenquimal (M), respectivamente; as cores azul escuro e claro os subtipos *basal-like 1* (BL1) e *basal-like 2* (BL2), respectivamente; a cor preta o subtipo luminal receptor de andrógeno (LRA); e a cor laranja o subtipo instável (INS).



**Figura 11** – Curvas de sobrevida global das amostras de CMTN que foram classificadas nos subtipos moleculares de acordo com a classificação de LEHMANN et al. (2011) através da ferramenta *TNBCType* (CHEN tal al. 2012) e de acordo com o status de *BRCA1* (*BRCA1*-deficiente e *BRCA1*-proficiente). Curvas de Kaplan-Meier geradas a partir da sobrevida global em 11 anos (4000 dias) das amostras de CMTN de acordo com os subtipos moleculares (A) e de acordo com o status de *BRCA1* (B). *BL1/2*: subtipos *basal-like 1* e *basal-like 2* (*BL1/2*); IM: subtipo imunodulador (IM); INS: subtipo instável; M/MSL: subtipos mesenquimal (M) e mesenquimal *stem like* (MSL); LRA: subtipo luminal receptor de andrógeno (LRA) (log-rank;  $p=0,275$ ;  $p=0,511$ ).



**Figura 12** – Curvas de sobrevida global das amostras de CMTN *BRCA1*-deficiente (A) e CMTN *BRCA1*-proficiente (B) que foram classificadas no subtipo molecular imunomodulador (IM) e nos demais subtipos que chamamos de não-imunomodulador (M, *MSL*, *BL1*, *BL2*, LRA, INS) através da ferramenta *TNBCType* (CHEN tal al. 2012). A linha vermelha representa a curva de sobrevida das amostras classificadas no subtipo IM, e a linha azul representa a curva de sobrevida das amostras classificadas nos outros subtipos. log-rank;  $p=0,730$ . log-rank;  $p=0,237$ .

#### 4.4 ANÁLISE DOS GENES DIFERENCIALMENTE EXPRESSOS

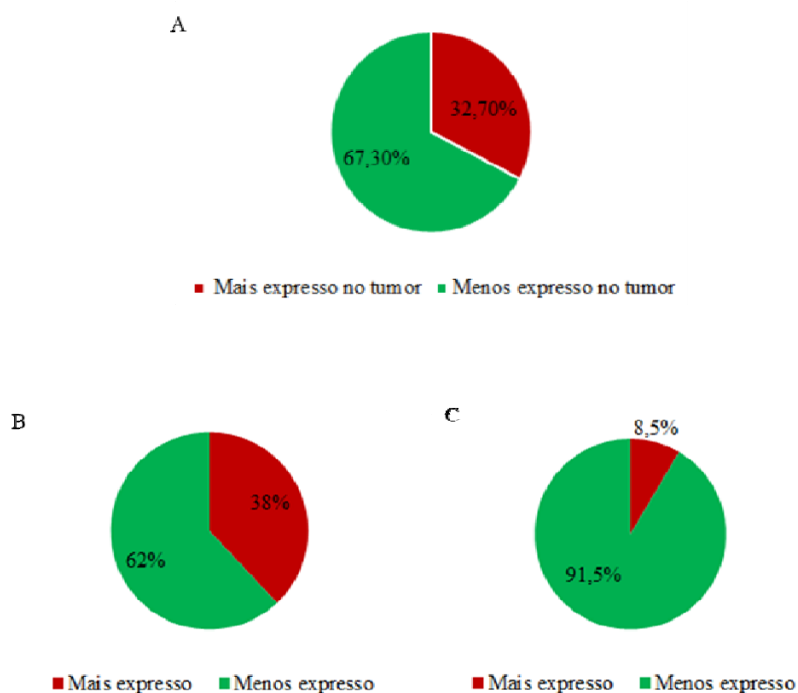
A análise da expressão diferencial nas amostras tumorais em relação às amostras normais foi realizada com todos os 25 casos de CMTN pareados assim como nos grupos, separadamente, 8 casos de CMTN Hereditário e 17 casos de CMTN Esporádico.

Após normalização feita pelo programa *DESeq2* dos valores de expressão pelo de todos os genes identificados em todas as amostras tumorais e normais com pelo menos 1 sequência única mapeada, a análise dos genes diferencialmente expressos (GDEs) nas amostras tumorais detectou 2039 GDEs, dos quais 32,7% (667/2039) estavam mais expressos e 67,3% (1372/2039) menos expressos no tumor, quando comparados com o tecido normal. Em relação à classificação de todos os GDEs pelo *Ensembl*, 81,5% (1661/2039) foram codificadores de proteína, 15,5% (318/2039) não codificadores de proteína e 3% (65/2039) pseudogene. Dos GDEs codificadores de proteína, 38% (630/1661) foram mais expressos e 62% (1031/1661) foram menos expressos no tumor. E, dos GDEs não codificadores de proteínas, 8,5% (27/318) foram mais expressos e 91,5% (291/318) foram menos expressos no tumor (Tabela 2). Assim, podemos observar que a maior porcentagem dos GDEs estavam menos expressos no tumor (Tabela 2; Figura 13). E que os GDEs não codificadores de proteína apresentaram uma maior porcentagem de genes menos expressos no tumor (91,5%) quando comparados com os GDEs codificadores de proteínas, os quais 62% estavam menos expressos no tumor (Tabela 2; Figura 13).



**Tabela 2** - Dados estatísticos dos 2039 GDEs nas amostras de CMTN pareadas.

CMTN x N	
<b>Todos os GDEs no tumor</b>	
Mais expressos no tumor	32,70% (667/2039)
Menos expressos no tumor	67,30% (1372/2039)
<b>GDEs codificador de proteína</b>	
Mais expressos no tumor	38% (630/1661)
Menos expressos no tumor	62% (1031/1661)
<b>GDEs não-codificador de proteína</b>	
Mais expressos no tumor	8,50% (27/318)
Menos expressos no tumor	91,50% (291/318)



**Figura 13** - Distribuição da porcentagem de todos os Genes Diferencialmente Expressos (GDEs) no tumor e dos GDEs classificados de acordo com o *RefSeq*. (A) Distribuição da porcentagem de todos os GDEs no tumor. (B) Distribuição da porcentagem de todos os GDEs codificadores de proteína. (C) Distribuição da porcentagem de todos os GDEs não codificadores de proteína (C) em todas as amostras CMTN, de acordo com a anotação do *RefSeq*. A cor verde representa os genes mais expressos no tumor, e a vermelha, os genes menos expressos no tumor.

Quando realizamos essa mesma análise com os casos dos grupos das amostras, CMTN Hereditário e CMTN Esporádico, observamos que foram detectados um total de 1965 GDEs nas amostras tumorais do grupo Hereditário em relação às normais, dos quais 30% (589/1965) foram mais expressos e 70% (1376/1965) menos expressos no tumor. E 85% (1672/1965) de todos os GDEs foram classificados pelo *RefSeq* como codificadores de proteína; 11,8% (233/1965) como não codificadores de proteína e 3% (60/1965) como pseudogene. Dentre todos os GDEs codificadores de proteína, 33,7% (564/1672) foram mais expressos e 66,3% (1108/1671) foram menos expressos no tumor. E, dos GDEs não codificadores de proteína, 7,7% (18/233) foram mais expressos e 92,3% (215/233) foram menos expressos no tumor (Tabela 3).

No grupo Esporádico, um total de 1837 GDEs no tumor foram identificados, dos quais 35,1% (645/1837) foram mais expressos e 64,9% (1192/1837) menos expressos no tumor. Em relação à classificação pelo *RefSeq*, 82,5% (1516/1837) de todos os GDEs foram classificados como codificadores de proteína; 14% (256/1837) como não codificadores de proteína e 3,5% (65/1837) como pseudogene. De todos os GDEs codificadores de proteína, 40% (610/1516) estavam mais expressos e 60% (906/1516) menos expressos no tumor. E de todos os GDEs não codificadores, 9,3% (24/256) estavam mais expressos e 90,7% (232/256) menos expressos no tumor (Tabela 4).

Em ambos os grupos das amostras, observamos que a maior porcentagem de todos os GDEs foram menos expressos no tumor (Figura 14A). O mesmo observado quando todos esses GDEs foram classificados em codificadores de proteínas (Figura 14B) e não codificadores de proteína (Figura 14C). Além disso, em ambos os grupos

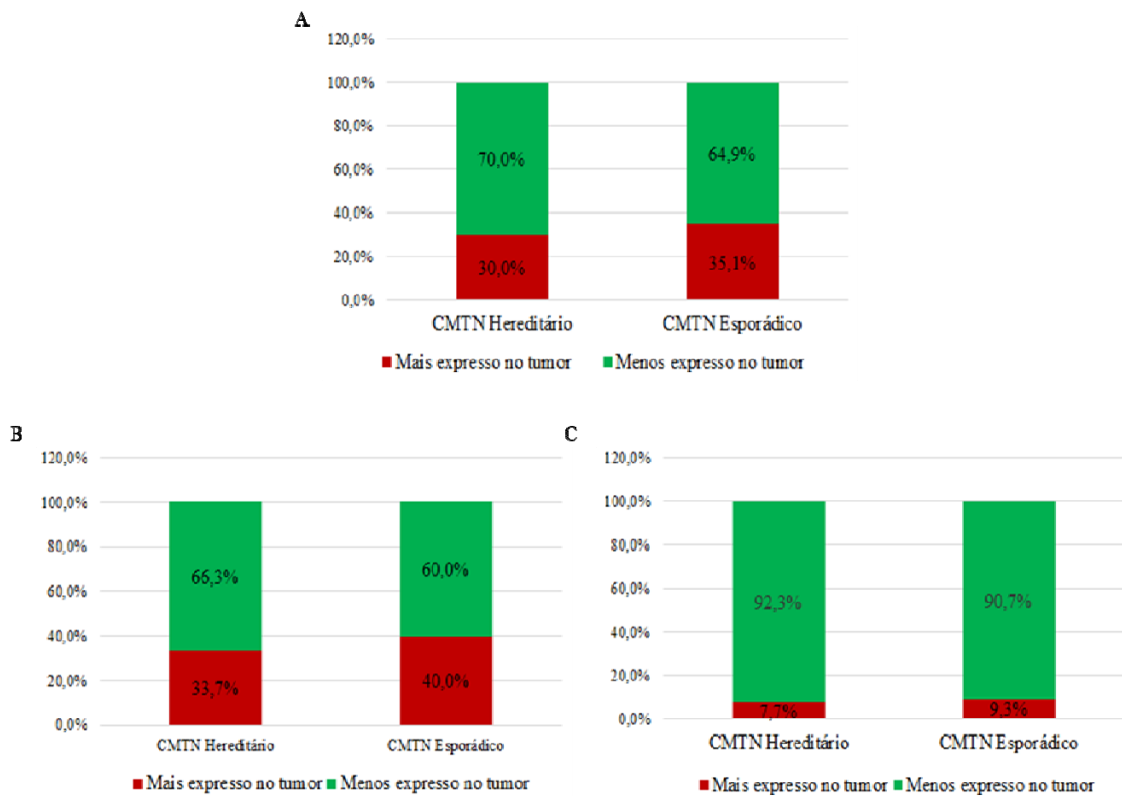
das amostras, os GDEs não codificadores de proteína apresentaram uma maior porcentagem de genes menos expressos no tumor que os GDEs codificadores de proteínas (Tabelas 3 e 4; Figura 14B e C).

**Tabela 3** - Dados estatísticos dos 1965 GDEs nas amostras de CMTN Hereditário

<b>CMTN Hereditário</b>	
<b>Todos os GDEs no tumor</b>	
Mais expressos no tumor	30% (589/1965)
Menos expressos no tumor	70% (1376/1965)
Codificador de proteína	85% (1672/1965)
Não codificador de proteína	11,8% (233/1965)
Pseudogene	3% (60/1965)
<b>GDEs codificador de proteína</b>	
Mais expressos no tumor	33,7% (564/1672)
Menos expressos no tumor	66,3% (1108/1671)
<b>GDEs não-codificador de proteína</b>	
Mais expressos no tumor	7,7% (18/233)
Menos expressos no tumor	92,3% (215/233)

**Tabela 4** - Dados estatísticos dos 1837 GDEs nas amostras de CMTN Esporádico.

<b>CMTN Esporádico</b>	
<b>Todos os GDEs no tumor</b>	
Mais expressos no tumor	35,1% (645/1837)
Menos expressos no tumor	64,9% (1192/1837)
Codificador de proteína	82,5% (1516/1837)
Não codificador de proteína	14% (256/1837)
Pseudogene	3,5% (65/1837)
<b>GDEs codificador de proteína</b>	
Mais expressos no tumor	40% (610/1516)
Menos expressos no tumor	60% (906/1516)
<b>GDEs não-codificador de proteína</b>	
Mais expressos no tumor	9,3% (24/256)
Menos expressos no tumor	90,7% (232/256)



**Figura 14** - Proporção da porcentagem de todos os Genes Diferencialmente Expressos (GDEs) no tumor para cada grupo das amostras de CMTN Hereditário e CMTN Esporádico (A), a proporção da porcentagem de todos os GDEs codificadores de proteína (B) e de todos os GDEs não codificadores de proteína (C).

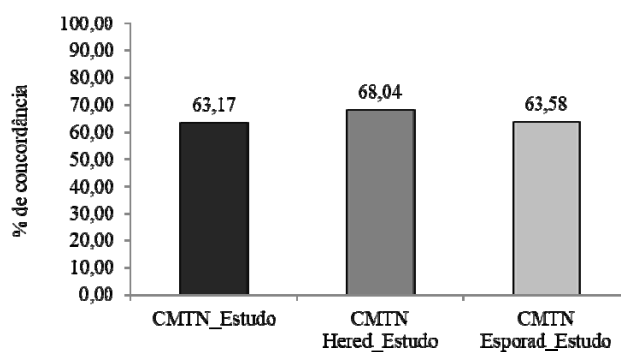
A partir desses nossos dados de expressão diferencial para as amostras de CMTN (CMTN\_Estudo) assim como para as amostras dos grupos Hereditário (CMTN Hereditário\_Estudo) e Esporádico (CMTN Esporádico\_Estudo), interrogamos quais GDEs da nossa casuística eram também GDEs nas amostras de CMTN do TCGA (CMTN\_TCGA). Antes de analisarmos a concordância dos nossos dados com os dados do TCGA, identificamos que os dados da expressão diferencial nas amostras CMTN\_TCGA (124 amostras), em relação às amostras normais não pareadas (13 amostras), revelou 1555 GDEs no tumor (Quadro 8). Quando

realizamos a busca dos GDEs nas amostras CTMN\_TCGA por agrupamento de idade, observamos 571 GDEs nas amostras CTMN\_TCGA com idade menor que 40 anos (21 amostras), 1591 GDEs nas amostras CTMN\_TCGA com idade entre 40 e 50 anos (29 amostras) e 1654 GDEs nas amostras CTMN\_TCGA com idade superior a 51 anos (74 amostras) (Quadro 8).

**Quadro 8** - Dados dos GDEs das amostras de CMTN do TCGA

<b>RNASeq_TCGA</b>	<b>Amostras TCGA</b>	<b>GDEs no tumor</b>
CMTN	124T / 13N	1552
CMTN < 40 anos	21T / 13N	571
CMTN $\geq 41 \leq 50$ anos	29T / 13N	1591
CMTN > 51 anos	74T / 13N	1652

Após isso, verificamos uma concordância de 63,17% (1288/2039) entre os GDEs nas amostras CMTN\_Estudo e os GDEs nas amostras CMTN\_TCGA. E, quando interrogamos os GDEs das amostras dos grupos Hereditário\_Estudo e Esporádico\_Estudo, observamos uma porcentagem de concordância de 68,04% (1337/1965) e 63,58% (1168/1837), respectivamente, com os GDEs nas amostras CMTN\_TCGA (Figura 15).



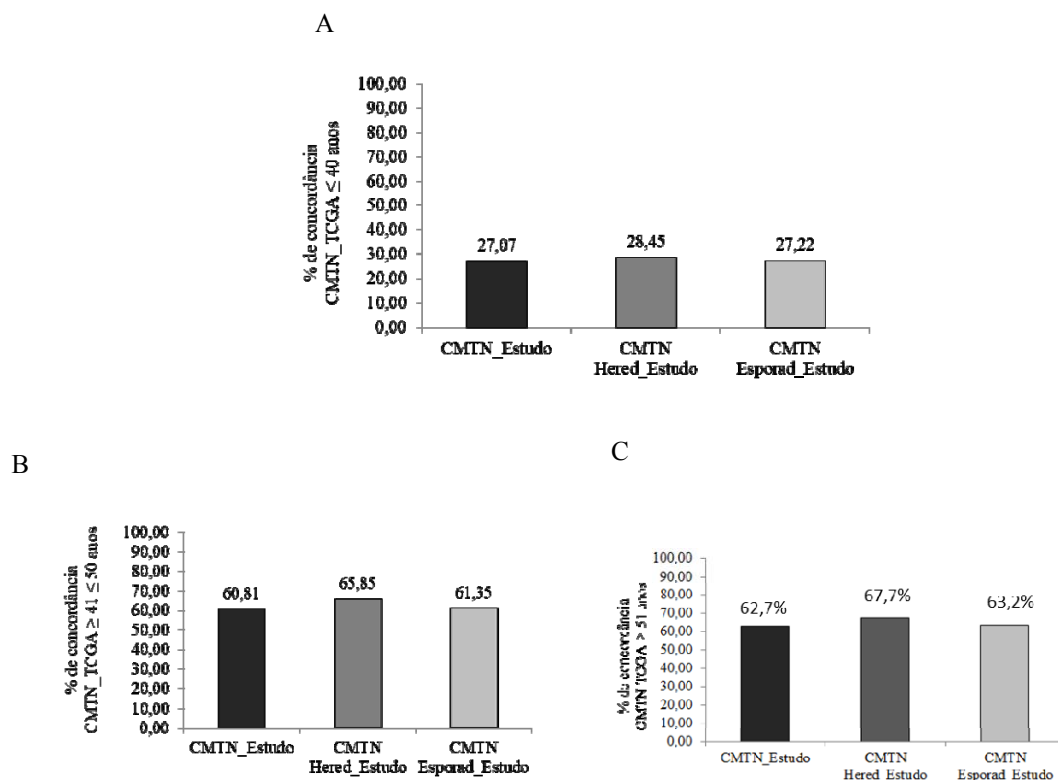
**Figura 15** - Porcentagem de concordância dos GDEs encontrados em todas as amostras CMTN do estudo (CMTN\_Estudo) com os GDEs das amostras CMTN do TCGA. Barra preta: concordância dos GDEs em todas as de amostras CMTN do estudo (CMTN\_Estudo); barra cinza escuro: concordância dos GDEs nas amostras dos grupo Hereditário (CMTN Hered\_Estudo); barras cinza claro: concordância dos GDEs nas amostras do grupo Esporádico (CMTN Esp\_Estudo).

Quando essa análise foi feita em relação às amostras CMTN\_TCGA com diferentes idades, observamos uma maior porcentagem de concordância das amostras CMTN\_Estudo com as amostras CMTN\_TCGA com idade superior a 41 anos. Como podemos ver na Figura 16 (Barras pretas), obtivemos 60,81% (1240/2039) de concordância com as amostras CMTN\_TCGA  $\geq 41 \leq 50$  anos e 62,68% (1278/2039) de concordância com as amostras CMTN\_TCGA  $\geq 51$  anos. Já com as amostras CMTN\_TCGA  $< 40$  anos, a concordância foi de 27,07% (552/2039) (Barras pretas da Figura 16).

Quando analisamos essa mesma comparação com as amostras dos grupos CMTN Hereditário\_Estudo e CMTN Esporádico\_Estudo, observamos também uma maior porcentagem de concordância das amostras do grupo Hereditário\_Estudo com as amostras CMTN\_TCGA com idade superior a 41 anos: 65,85% (1294/1965) em relação às amostras CMTN\_TCGA  $\geq 41 \leq 50$  anos e 67,68% (1330/1965) em relação

às amostras CMTN\_TCGA  $\geq 51$  anos. E, com as amostras CMTN\_TCGA  $< 40$  anos, a concordância foi 28,45% (559/1965).

A comparação das amostras do grupo esporádico de estudo com as amostras CMTN\_TCGA com idade superior a 41 anos também mostrou uma maior concordância: 61,35% (1127/1837) em relação às amostras CMTN\_TCGA  $\geq 41 \leq 50$  anos e 63,20% (1161/1837) em relação às amostras CMTN\_TCGA  $\geq 51$  anos, enquanto que, com as amostras CMTN\_TCGA  $< 40$  anos, a concordância foi de 27,22% (500/1837) (Barras cinzas médio e claro da Figura 16).



**Figura 16** - Porcentagem de concordância dos GDEs encontrados nas amostras de CMTN do estudo (CMTN\_Estudo) com os GDEs nas amostras de CMTN do TCGA (CMTN\_TCGA) por grupo de idade: (A) CMTN\_TCGA  $\leq 40$  anos; (B) CMTN\_TCGA  $\geq 41 \leq 50$  anos; (C) CMTN\_TCGA  $\geq 51$  anos (C). Barras pretas: concordância dos GDEs nas amostras de CMTN do estudo (CMTN\_Estudo); barra cinza escuro: concordância dos GDEs nas amostras do grupo Hereditário (CMTN Hered\_Estudo); barras cinza claro: concordância dos GDEs nas amostras do grupo Esporádico (CMTN Esp\_Estudo).

#### 4.5 ANÁLISE FUNCIONAL *IN SILICO* ATRAVÉS DO IPA

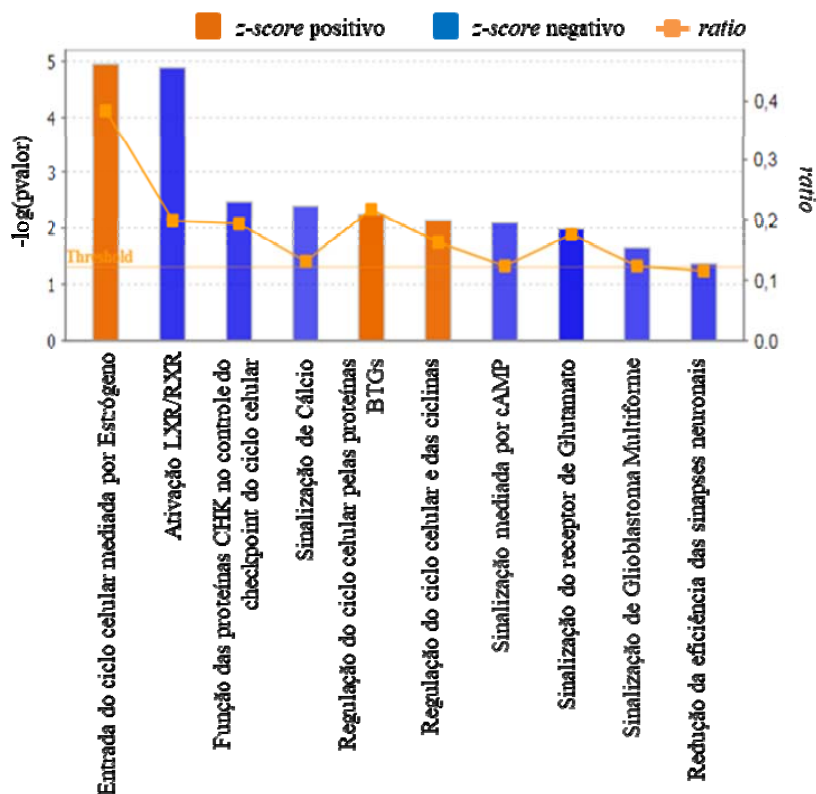
A análise funcional realizada pelo IPA dos GDEs nas amostras tumorais em relação às amostras normais pareadas, com o intuito de melhor entender os processos biológicos envolvidos com a tumorigênese do CMTN, assim como do CMTN Hereditário e do CMTN Esporádico, foi primeiramente descrita para todas as amostras de CMTN e, em seguida, para as amostras de grupo Hereditário e do Esporádico.

O enriquecimento das vias canônicas em todas as amostras de CMTN revelou que a via Entrada do ciclo celular mediada por estrógeno foi a mais significativamente ativada com z-score 2,5, e a via canônica Ativação de LXR/RXR foi a mais significativamente inibida com z-score -2,4 (Quadro 9; Figura 17).

**Quadro 9** - Vias canônicas enriquecidas em todas as amostras de CMTN

Vias canônicas	-log(p valor ajustado)	p valor ajustado	z-score	ratio
Entrada do ciclo celular mediada por estrógeno	5	1,10E-05	2,5	0,39
Regulação do ciclo celular pelas proteínas BTGs	2,2	6,00E-03	2,6	0,22
Regulação do ciclo celular e das ciclinas	2,1	7,40E-03	2,5	0,16
Ativação de LXR/RXR	4,9	1,30E-05	-2,4	0,2
Função das proteínas CHK no controle do checkpoint do ciclo celular	2,5	3,50E-03	-2,3	0,19
Sinalização de Cálcio	2,4	4,20E-03	-2	0,13
Sinalização mediada por cAMP	2,1	8,10E-03	-2,1	0,12
Sinalização do receptor de Glutamato	2	1,00E-02	-2,6	0,18
Sinalização de Glioblastoma Multiforme	1,6	2,30E-02	-2,1	0,12
Redução da eficiência das sinapses neuronais	1,4	4,40E-02	-2,2	0,12



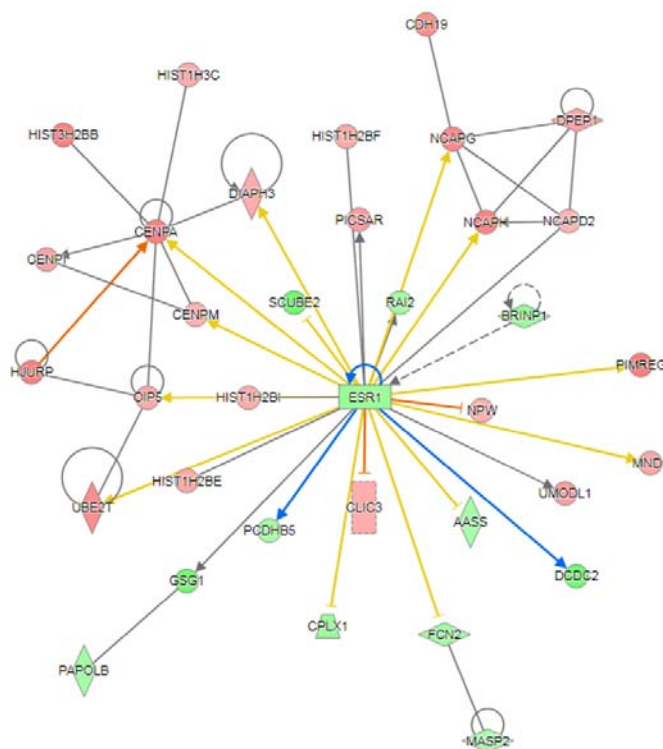


**Figura 17** - Vias canônicas enriquecidas em ordem decrescente da significância estatística em todas as amostras CMTN.

A análise do IPA também mostrou 7 redes de interação biológica cujas funções biológicas podem estar potencialmente envolvidas com a tumorigênese do CMTN. A rede de maior pontuação (*p-score*) e, portanto, de maior valor estatístico significativo, apresentou maior envolvimento das funções de Replicação, recombinação e reparo do DNA, na qual foram identificados 35 genes focais interagindo entre si e que podem estar mediando essas funções biológicas (Quadro 10). Dentre esses genes focais, identificamos que o receptor de estrógeno (*ESR1*) é um gene central nessa rede de interação biológica (Figura 18).

**Quadro 10** - Redes de interação biológica identificadas a partir dos GDEs em todas as amostras CMTN.

Redes	Doenças e funções biológicas	p-score	Genes focais
1	Replicação, recombinação e reparo do DNA	44	35
2	Movimento celular	37	32
3	Sinalização celular	34	31
4	Desenvolvimento	32	30
5	Doença metabólica	32	30
6	Desordem hereditária	30	29
7	Resposta imuno humoral e inflamatória	27	27



**Figura 18** - Rede de interação biológica 1 identificada a partir dos GDEs em todas as amostras CMTN. Os círculos coloridos representam os 35 genes focais envolvidos com as funções biológicas de Replicação, recombinação e reparo do DNA. A cor vermelha e verde são os genes mais e menos expressos nas amostras de CMTN, respectivamente. As setas azuis representam a interação predita inibida, as laranjas representam a interação predita ativada, as amarelas representam a interação não predita ativada e nem inibida, e as setas cinzas representam a não predição da atividade da interação.

Identificamos, também, importantes reguladores *upstream* no conjunto dos GDEs em todas as amostras CMTN. Dentre estes, destacamos o *FOXMI* que foi o gene mais estatisticamente significativo ativado com z-score 3,40. Foi observado, também, que o *FOXMI* interage com 43 genes alvo (*target molecules in dataset*) além de regular outros 27 reguladores *upstream*, ativando-os ou inibindo-os e, assim, regulando diversas cascatas de sinalização (*Mechanistic network*) (Quadro 11). O gene *CDKN2A* foi o mais significativamente inibido com z-score -3,18 e capaz de interagir com 58 genes alvo (*target molecules in dataset*) e de regular outros 17 reguladores *upstream* importantes para ativar ou inibir cascatas de sinalização (*Mechanistic network*) (Quadro 11).

**Quadro 11** - Reguladores *upstream* ativados ou inibidos identificados em todos os GDEs quando todas as amostras de CMTN foram comparadas com os seus respectivos normais

<i>Upstream</i> regulador ativado	Tipo da molécula	p valor ajustado	z-score	<i>Target molecules</i> <i>in dataset</i>	<i>Mechanistic</i> <i>network</i>
<i>FOXMI</i>	regulador transcricional	5,95E-18	3,40	43	564 (27)
<i>EGF</i>	fator de crescimento	2,43E-07	2,45	73	655 (21)
<i>IFNLA</i>	citocina	5,49E-07	2,64	7	202 (8)
<i>STAT1</i>	regulador transcricional	8,36E-06	2,36	52	514 (19)
<i>CCNE1</i>	regulador transcricional	1,37E-05	2,16	10	344 (15)
<i>IFNL1</i>	citocina	1,84E-04	3,93	16	354 (13)
<i>TALI</i>	regulador transcricional	2,42E-04	2,49	33	268 (6)
<i>NRG1</i>	fator de crescimento	4,40E-03	2,49	29	574 (19)
<i>NTRK2</i>	cinase	4,77E-03	2,55	11	506 (21)
<i>SPP1</i>	citocina	6,55E-03	2,29	24	599 (19)
<i>CDKN2A</i>	regulador transcricional	4,17E-11	-3,18	58	394 (17)
<i>let-7</i>	microRNA	4,97E-09	-4,44	41	465 (12)
<i>THRB</i>	receptor nuclear ligante- dependente	7,11E-07	-2,69	38	459 (21)
<i>PPARG</i>	receptor nuclear ligante- dependente	3,78E-06	-4,54	65	469 (22)
<i>LIPE</i>	enzima	2,48E-05	-2,05	24	451 (26)
<i>MYOCD</i>	regulador transcricional	3,09E-04	-2,73	17	121 (5)
<i>LEPR</i>	receptor transmembrana	8,72E-04	-2,30	26	479 (22)

Essa mesma análise funcional realizada através do *IPA* também foi feita separadamente para cada grupo das amostras de CMTN Hereditário e CMTN Esporádico a partir dos GDEs nas amostras tumorais de cada grupo.

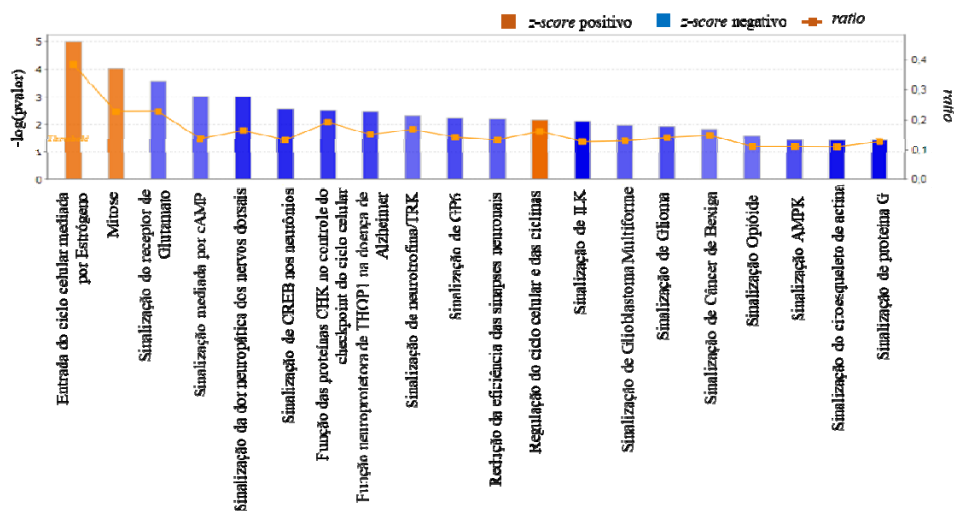
Como podemos observar nos Quadros 12 e 13, os dois grupos apresentaram vias canônicas enriquecidas comuns significativamente ativadas as quais estão envolvidas, de forma geral, com a Regulação do ciclo celular. Entretanto, as vias canônicas significativamente inibidas mostraram-se exclusivas em cada grupo, embora algumas poucas vias foram comuns. Dentre as vias inibidas exclusivas de cada grupo, a via de Sinalização do receptor do glutamato foi a mais estatisticamente significativa com z-score -2,12 no grupo Hereditário (Quadro 12 e Figura 19), enquanto que a via de Ativação de LXR/RXR foi a mais estatisticamente significativa com z-score -2,98 no grupo Esporádico (Quadro 13 e Figura 20).

**Quadro 12** - Vias canônicas enriquecidas nas amostras CMTN Hereditário.

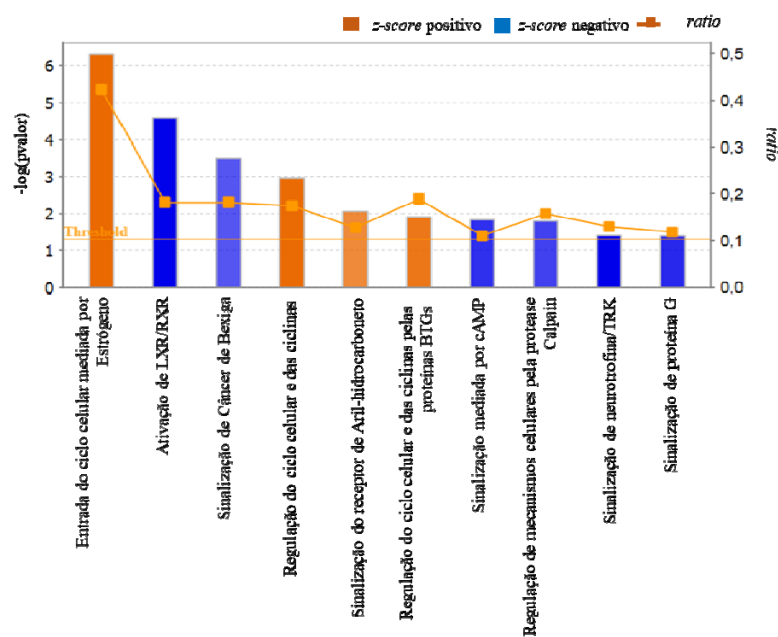
Vias canônicas	<b>-log (p-valor ajustado)</b>	<b>p valor ajustado</b>	<b>z-score</b>	<b>ratio</b>
Entrada da fase S do ciclo celular mediata por estrógeno	5,00	0,00	2,53	0,39
Função mitótica das cinases Polo-Like	4,03	0,00	2,50	0,23
Regulação do ciclo celular	2,18	0,01	3,05	0,16
Sinalização do receptor de Glutamato	3,58	2,63E-04	-2,12	0,23
Sinalização da dor neuropática dos nervos dorsais	3,02	9,55E-04	-2,98	0,17
Sinalização mediada por cAMP	3,02	9,55E-04	-2,19	0,14
Sinalização de CREB nos neurônios	2,57	2,69E-03	-2,98	0,13
Função das proteínas CHK no controle do checkpoint do ciclo celular	2,50	3,16E-03	-2,83	0,19
Função neuroprotetora de THOP1 na doença de Alzheimer	2,45	3,55E-03	-2,67	0,15
Sinalização de neurotrofina/TRK	2,32	4,79E-03	-2,11	0,17
Sinalização de GP6	2,24	5,75E-03	-2,52	0,14
Redução da eficiência das sinapses neuronais	2,21	6,17E-03	-2,29	0,13
Sinalização de ILK	2,13	7,41E-03	-3,55	0,13
Sinalização de Glioblastoma Multiforme	1,98	1,05E-02	-2,32	0,13
Sinalização de Glioma	1,93	1,17E-02	-2,53	0,14
Sinalização de Câncer de Bexiga	1,84	1,45E-02	-2,00	0,15
Sinalização Opióide	1,57	2,69E-02	-2,04	0,11
Sinalização de proteína G	1,43	3,72E-02	-3,46	0,13
Sinalização do citoesqueleto de actina	1,43	3,72E-02	-3,30	0,11
Sinalização AMPK	1,43	3,72E-02	-2,50	0,11

**Quadro 13** - Vias canônicas enriquecidas nas amostras CMTN Esporádico.

Vias canônicas	<b>-log (p-valor ajustado)</b>	<b>p valor ajustado</b>	<b>z-score</b>	<b>ratio</b>
Entrada do ciclo celular mediada por estrógeno	6,31	4,90E-07	2,71	0,42
Regulação do ciclo celular e das ciclinas	2,96	1,10E-03	2,67	0,18
Sinalização do receptor de Aril-hidrocarboneto	2,06	8,71E-03	2,11	0,13
Regulação do ciclo celular e das ciclinas pelas proteínas BTGs	1,91	1,23E-02	2,45	0,19
Ativação de LXR/RXR	4,58	2,63E-05	-2,98	0,18
Sinalização de Câncer de Bexiga	3,50	3,16E-04	-2,00	0,18
Sinalização mediada por cAMP	1,84	1,45E-02	-2,45	0,11
Regulação de mecanismos celulares pela protease Calpain	1,80	1,58E-02	-2,24	0,16
Sinalização de neurotrofina/TRK	1,41	3,89E-02	-3,00	0,13
Sinalização de proteína G	1,39	4,07E-02	-2,53	0,12



**Figura 19** - Enriquecimento das vias canônicas distribuídos em ordem decrescente de significância estatística nas amostras de CMTN Hereditário.



**Figura 20** - Enriquecimento das vias canônicas distribuídos em ordem decrescente de significância estatística nas amostras de CMTN Esporádico.

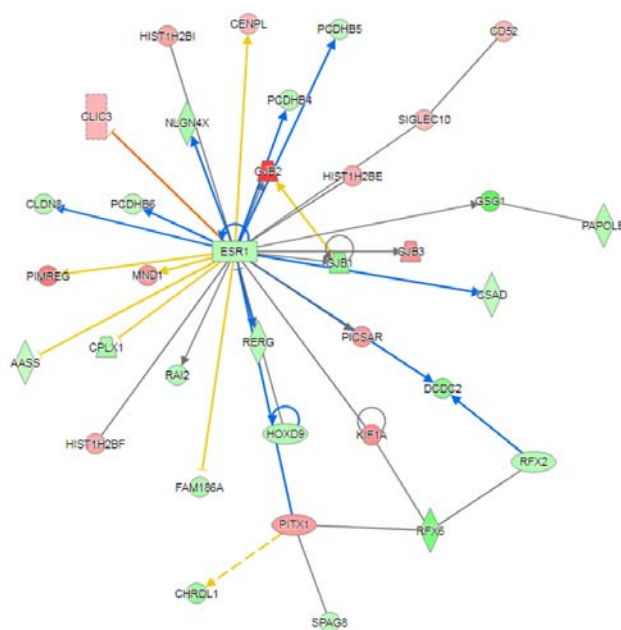
Além disso, podemos observar que as redes de interação biológica, potencialmente envolvidas na tumorigênese do CMTN Hereditário e do CMTN Esporádico, também foram bastante similares com diferenças apenas na ordem de significância (Quadros 14 e 15). Nas amostras CMTN Hereditário, podemos observar que a rede de interação que envolve Desordem hereditária foi a de maior significância estatística (p-score 44) com 35 genes focais (Quadro 14). A Figura 21 mostra a interação dos 35 genes focais onde identificamos o receptor de estrógeno (*ESR1*) como o gene central nessa rede de interação. Nas amostras CMTN Esporádico, observamos que a rede de interação mais significativa com p-score 43 está envolvida com Comportamento e transtornos psicológicos, apresentando 34 genes focais (Quadro 15). E, como podemos observar na Figura 22, nenhum gene específico se mostrou central nessa rede de interação.

**Quadro 14** - Redes de interação biológica identificadas nas amostras CMTN Hereditário

Redes	Doenças e funções biológicas	p-score	Genes focais
1	Desordem hereditária	44	35
2	Replicação, recombinação e reparo do DNA	37	32
3	Câncer, Doença do sistema reprodutivo	37	32
4	Organização e montagem celular, Modificação pós-transcricional	34	31
5	Desenvolvimento	32	30
6	Doença inflamatória, Doença respiratória	32	30
7	Movimento e morfologia celular	31	29
8	Desenvolvimento e função do tecido conjuntivo e do sistema reprodutivo	31	29
9	Metabolismo de carboidratos, Câncer	31	29
11	Desenvolvimento e função do sistema do sistema musculoesquelético e do sistema endócrino	29	28
12	Transtorno psicológicos	27	27
13	Metabolismo de aminoácidos	27	27
14	Metabolismo de droga e de lipídios	25	26
15	Doença hematológica	25	26

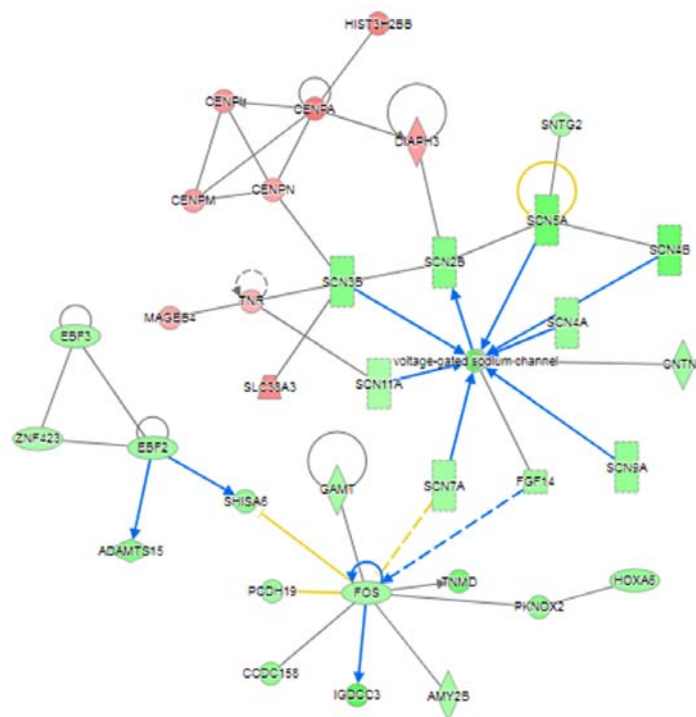
**Quadro 15** - Redes de interação biológica identificadas nas amostras CMTN Esporádico.

Redes	Doenças e funções biológicas	<i>p</i> -score	Genes focais
1	Comportamento, Transtornos psicológicos	43	34
2	Ciclo celular, morfologia celular	40	33
3	Replicação, reparo e recombinação do DNA, Modificação pós-transcricional	40	33
4	Doença hematológica	38	32
5	Transporte de moléculas, Metabolismo de ácido nucléico	36	31
6	Desordem hereditária	36	31
7	Desenvolvimento embrionário	34	30
8	Doenças dermatológicas, Morte celular e sobrevivência	30	28
9	Desenvolvimento e função do sistema musculo-esquelético	30	28
10	Manutenção celular, Transporte de moléculas	28	27
11	Desenvolvimento e função do sistema nervoso	28	27
12	Desenvolvimento e função do sistema cardíaco	26	26
13	Metabolismo de lipídeos, minerais e de vitaminas	26	26
14	Ciclo celular, Câncer	26	26
15	Sinalização célula-célula, Manutenção e função celular	24	25



**Figura 21** - Rede de interação biológica 1 cujos 35 genes focais estão envolvidos com Desordem hereditária nas amostras CMTN Hereditário. Os círculos vermelho e verde representam genes focais hiper e hipoxpressos, respectivamente. As setas azuis representam a interação predita inibida, as laranjas representam a interação predita ativada, as amarelas representam a interação não predita ativada e nem inibida, e as setas cinzas representam que não foi possível prever a atividade da interação.





**Figura 22** - Rede de interação biológica 1 cujos 34 genes focais estão envolvidos com Comportamento e transtornos psicológicos nas amostras CMTN Esporádico. Os círculos vermelho e verde representam genes focais hiper e hipoexpressos, respectivamente. As setas azuis representam a interação predita inibida, as laranjas representam a interação predita ativada, as amarelas representam a interação não predita ativada e nem inibida, e as setas cinzas representam que não foi possível prever a atividade da interação.

Além disso, analisamos importantes reguladores *upstream* para cada grupo das amostras. Observamos alguns reguladores significativamente ativados que foram comuns nos dois grupos, tais como os genes *FOXM1*, *CCNE1*, *NTRK2* e *STAT1*. Identificamos, também, reguladores exclusivos de cada grupo. Dentre estes, identificamos o gene *TNF* como o mais significativamente ativado com z-score 4,06 apenas nas amostras de CMTN Hereditário (Quadro 16), enquanto que o gene *E2F1* foi o mais significativamente ativado com z-score 2,22 apenas nas amostras de CMTN Esporádico (Quadro 17). Em relação aos reguladores inibidos estatisticamente significativos comuns nos dois grupos, identificamos os genes

*CDKN2A* e *THRB*, e o microRNA let-7 (Quadro 16). Além disso, alguns reguladores foram significativamente inibidos exclusivamente em cada grupo. Dentre esses, o microRNA mir-21 foi identificado com maior significância estatística com z-score -4,68 apenas nas amostras CMTN Hereditário (Quadro 16), e o gene *PPARG* foi o de maior significância estatística com z-score -4,24 apenas nas amostras CMTN Esporádico (Quadro 17).

**Quadro 16** - Reguladores *upstream* ativados ou inibidos identificados nas amostras CMTN Hereditário.

Regulador <i>Upstream</i>	Tipo da molécula	p valor ajustado	z-score	Número de genes alvo	Número de genes preditos na regulação causal
TNF	citocina regulador	3,52E-18	4,06	237	424 (14)
FOXM1	transcricional regulador	3,00E-17	3,54	42	514 (24)
TAL1	transcricional regulador	4,93E-08	2,18	42	275 (6)
CCNE1	transcricional	1,28E-05	2,16	10	333 (15)
NTRK2	cinase regulador	2,54E-05	2,35	25	558 (22)
STAT1	transcricional	1,21E-04	2,98	47	389 (16)
IFNL4	citocina regulador	2,70E-04	2,23	5	222 (7)
CDKN2A	transcricional	3,14E-11	-3,63	58	346 (15)
mir-21	microRNA	9,99E-09	-4,68	48	553 (20)
THRB	receptor nuclear ligante-dependente	2,44E-08	-2,85	41	421 (20)
ADIPOQ	outro	4,28E-08	-2,37	33	495 (23)
let-7	microRNA	1,23E-07	-3,35	38	550 (22)
GATA2	regulador transcricional	2,10E-07	-2,71	61	413 (12)
IRS1	fator de crescimento	7,12E-06	-2,14	34	495 (22)
BDNF	citocina	4,00E-05	-2,62	25	507 (19)
CXCL12	citocina	4,39E-03	-2,47	26	482 (21)
PTHLH	outro	8,99E-03	-2,46	12	537 (24)

**Quadro 17** - Reguladores *upstream* ativados ou inibidos identificados nas amostras CMTN Esporádico.

<b>Regulador <i>Upstream</i></b>	<b>Tipo da molécula</b>	<b>p valor ajustado</b>	<b><i>z-score</i></b>	<b>Número de genes da rede</b>	<b>Número de genes preditos na regulação causal</b>
FOXM1	regulador transcricional	7,73E-18	3,95	41	476 (26)
E2F1	regulador transcricional	2,27E-10	2,22	75	318 (14)
ESR1	receptor nuclear ligante-dependente	3,32E-08	2,10	144	578 (21)
CCNE1	regulador transcricional	5,71E-06	2,16	10	308 (15)
S100A8	outro	1,91E-05	2,06	29	342 (11)
IFNL1	citocina	5,82E-05	3,93	6	319 (13)
S100A9	outro	1,33E-04	2,06	26	317 (18)
STAT1	regulador transcricional	2,31E-04	2,29	43	359 (16)
OSM	regulador transcricional	8,05E-04	2,13	58	390 (16)
MYBL2	citocina	1,93E-03	2,10	9	272 (13)
NRG1	fator de crescimento	4,20E-03	2,30	27	528 (19)
NTRK2	cinase	7,07E-03	2,96	10	452 (22)
CDKN2A	regulador transcricional	2,71E-11	-3,16	55	411 (19)
let-7	microRNA	9,83E-10	-4,77	40	448 (12)
THRB	receptor nuclear ligante-dependente	1,47E-06	-2,54	35	409 (21)
PPARG	receptor nuclear ligante-dependente	2,11E-05	-4,24	58	465 (24)
MYOCD	regulador transcricional	2,70E-05	-2,02	18	105 (4)
LIPE	enzima	1,42E-04	-2,24	21	429 (26)
ALDH1A1	enzima	2,06E-03	-2,20	5	366 (12)

## 5 DISCUSSÃO

O presente estudo consistiu do sequenciamento do RNA total (RNA-Seq) através da plataforma NextSeq 500 da Illumina de 37 amostras de tumor de mama de pacientes diagnosticadas com Câncer de Mama Triplo-Negativo (CMTN) em idade jovem ( $\leq 50$  anos), compreendendo 9 casos de pacientes com CMTN Hereditário (portadoras de mutação patogênica germinativa em *BRCA1*) e 28 casos de CMTN Esporádico (não portadoras de mutação germinativa em ambos genes *BRCA1* e *BRCA2*), dos quais 9 casos foram de tumores com promotor de *BRCA1* hipermetilado (*BRCA1*-deficiente) e 19 com o promotor de *BRCA1* não hipermetilado (*BRCA1*-proficiente). De todos os 37 casos, 25 foram pareados com o tecido normal adjacente ao tumor, cujo RNA-Seq permitiu a análise da expressão diferencial e dos processos biológicos envolvidos com a tumorigênese do CMTN no contexto de *BRCA1*.

Nesse trabalho, realizamos o perfil transcricional completo de 62 amostras, satisfatoriamente, com o intuito de investigar as alterações transcricionais e a tumorigênese do CMTN sob a perspectiva da inativação do gene *BRCA1*. Com esse estudo pretendemos contribuir com a caracterização molecular do CMTN, que sabidamente é uma doença bastante heterogênea. Nesse sentido, a tecnologia de sequenciamento de nova geração pode fornecer uma informação transcricional completa que é fundamental para melhor caracterização da doença.

Após o sequenciamento completo das 62 amostras de RNA total no formato *mid-output* 150 ciclos, conseguimos gerar 48.627.204 milhões de sequências, em média, por amostra. Porém, após análise de retirada das sequências do contaminante RNA ribossomal (rRNA), obtivemos uma porcentagem de 77,3% de sequências livre

de rRNA, em média, por amostra, valor abaixo do parâmetro do fabricante de 98% como critério da eficiência da depleção do rRNA. Contudo, mesmo tendo perdido uma parcela considerável das sequências geradas, obtivemos uma boa porcentagem de mapeamento (78,9%), que é um importante indicador da qualidade do RNA-Seq onde espera-se de 70 a 90% de mapeamento para um RNA-Seq acurado (DOBIN et al. 2013). Por isso, o número de sequências filtradas que foram mapeadas nesse estudo não comprometeu a análise de expressão diferencial. Conseguimos obter, satisfatoriamente, 28.911.123 milhões de sequências únicas mapeadas, em média, por amostra, das quais identificamos 18.928 mil genes expressos, em média, com pelo menos 1 sequência única por amostra. Esses valores estão dentro da extensão de cobertura recomendada para análises de expressão diferencial (20 a 30 milhões de sequências únicas mapeadas), de acordo com o fornecedor Illumina e dados relatados em estudos da literatura (ENCODE Project Consortium 2011; HITZEMANN et al. 2013; SIMS et al. 2014).

O CMTN é considerado um subtipo de câncer de mama altamente heterogêneo nos aspectos molecular e clínico (REIS-FILHO e TUTT 2008; LEHMANN et al. 2011; BURSTEIN et al. 2015), tornando a classificação molecular necessária para melhor caracterização de todas as amostras de CMTN. Assim, para todas as 37 amostras de CMTN do nosso estudo, realizamos a classificação molecular nos subtipos moleculares previamente estabelecidos por LEHMANN et al. (2011) através da ferramenta *online TNBCType* (CHEN et al. 2012). A partir do transcriptoma completo de todas as amostras de CMTN obtido pelo RNA-Seq, pudemos observar todos os 7 subtipos moleculares, incluindo o subtipo instável. Identificamos que a maioria das amostras foram classificadas molecularmente nos subtipos imunomodulador (22,8%) e mesenquimal (22,8%), o que também foi

observado pelos estudos clássicos dessa classificação molecular (CHEN et al. 2012; LEHMANN et al. 2016). Em relação ao CMTN no contexto da deficiência de *BRCA1*, observamos algumas diferenças quanto à classificação molecular. Enquanto que, nas amostras *BRCA1*-deficiente, observamos uma predominância do perfil imunomodulador, nas amostras *BRCA1*-proficiente, vimos um maior número de amostras com o perfil mesenquimal.

O CMTN caracterizado pela assinatura gênica imunomodulador, como observamos nas amostras *BRCA1*-deficiente, são tumores altamente enriquecidos com sinalização do sistema imune e que apresentam uma grande quantidade de linfócitos infiltrantes no tumor (TILs), células de um alto valor prognóstico no CMTN (LEHMANN et al. 2016). Essa assinatura imunomoduladora já foi observada enriquecida em câncer de mama RE-negativo (TESCHENDORFF et al. 2007). A presença dos TILs no CMTN é capaz de gerar um perfil de expressão gênica com aumento de expressão dos reguladores do *checkpoint* do sistema imune, tais como PD-1, PD-L1 e CTLA4, sendo fortemente correlacionada com o aumento da sobrevida global e livre de doença e com a resposta patológica completa às quimioterapias neoadjuvante e adjuvante (ADAMS et al. 2014; DENKERT et al. 2015; NANDA et al. 2016; ZHANG et al. 2018). Essa assinatura gênica do sistema imune nos chama atenção pois, recentemente, muitos estudos têm focado em classificar e estratificar as amostras de CMTN com base no perfil imunogenômico com potenciais implicações clínicas para o tratamento CMTN, o que pode suportar a abordagem de imunoterapia para o tratamento desses pacientes (LIU et al. 2018; XIAO et al. 2019). Entretanto, atualmente, as estratégias de imunoterapia exibem efeitos em menos de 20% dos pacientes com câncer, sugerindo que nem todos os pacientes com CMTN poderão responder à imunoterapia. E, de fato, características

moleculares, como perfil mutacional, expressão de moléculas como PD-L1 e deficiência do mecanismo de reparo *mismatch* do DNA, têm sido associadas com o tipo de resposta à imunoterapia (SNYDER et al. 2014; LE et al. 2015; RIZVI et al. 2015; VAN ALLEN et al. 2015; HUGO et al. 2016). Isso evidencia a importância dessa assinatura imune na compreensão da heterogeneidade da doença.

O CMTN caracterizado pela assinatura gênica mesenquimal, conforme vimos nos tumores *BRCA1*-proficiente, é enriquecido com ontologias e perfil de expressão gênica, envolvendo via de sinalização do TGF- $\beta$ , mTOR, Rac1/Rho, Wnt- $\beta$ -catenina, FGFR, PDGFR e VEGF. Considerando isso, são tumores que preferencialmente respondem bem à terapia alvo para essas vias de sinalização, como já encontrado em linhagens celulares de CMTN do subtipo mesenquimal tratadas com o inibidor de PI3K/mTOR (LEHMANN et al. 2011). Além disso, vias de sinalização associadas com a transição epitélio-mesênquima (TEM) estão enriquecidas no subtipo mesenquimal, o que sugere que são tumores mais sensíveis à terapia com Dasatinib (LEHMANN et al. 2011), inibidor de Src, importante gene para a regulação da TEM no processo de invasão das células tumorais (GUARINO 2010).

Embora a classificação molecular de LEHMANN et al. (2011) não seja a única estratégia para definir os subtipos moleculares do CMTN, ela tem demonstrado apresentar um forte impacto na sobrevida global pois o classifica em subtipos homogêneos com efeito preditivo nos resultados de terapias alvo específicas (LEHMANN et al. 2011; MASUDA et al. 2013). No nosso estudo, nós observamos que as amostras de CMTN, classificadas como imunodulador, *basal-like 1* e *basal-like 2*, apresentaram uma tendência de melhor sobrevida em relação aos outros subtipos, embora sem significância estatística. Esse dado corrobora com os estudos

que mostram que esses subtipos apresentam melhores sobrevida global e livre de doença em relação aos demais subtipos. Os tumores *BL1* e *BL2* apresentam enriquecimento de genes de proliferação e de resposta a dano ao DNA o que os favorece para tratamento com agentes antimetabólicos, como os taxanos, e que induzem dano ao DNA, como os agentes de platina, os levando à maiores qCR em relação ao subtipos LRA e M, por exemplo (BAUER et al. 2010; JUUL et al. 2010). E o subtipo IM é enriquecido por sinalização do sistema imune e alta expressão de genes envolvidos com resposta imune o que tem sido fortemente correlacionada com o aumento da sobrevida global e livre de doença nos tumores TN (LEHMANN et al. 2011; NANDA et al. 2016; ZHANG et al. 2018). Além disso, nossos dados, apesar de não terem mostrado significância estatística, sugerem que pacientes com tumores TN *BRCA1*-deficiente sobrevivem melhor, independente do subtipo molecular, em relação aos tumores TN *BRCA1*-proficiente. A ausência de significância estatística dessa análise poderia ser explicada pelo pequeno número de amostras do estudo. Já os tumores *BRCA1*-proficiente, classificados como imunomodulador, demonstraram uma melhor sobrevida global quando comparado com os mesmos tumores classificados nos outros subtipos moleculares, demonstrando que, para esse grupo de amostras, a classificação molecular teve um forte impacto na sobrevida global do paciente. No mais recente estudo do nosso grupo, observamos melhor sobrevida global e livre de doença, de forma significativa, nas pacientes jovens ( $\leq 40$  anos) diagnosticadas com CMTN *BRCA1*-deficiente quando comparadas com aquelas também jovens com CMTN *BRCA1*-proficiente (BRIANESE et al. 2018), embora alguns estudos reportaram nenhuma diferença significativa na sobrevida entre as pacientes com CMTN *BRCA1*-deficiente e com CMTN *BRCA1*-proficiente (BAYRAKTAR et al. 2011; PALUCH-SHIMON et al. 2016).



Em relação aos genes diferencialmente expressos (GDEs) nas amostras de CMTN em relação às normais pareadas, obtivemos uma maior concordância com os dados de expressão diferencial do TCGA para as amostras de pacientes diagnosticadas com CMTN em idade superior aos 40 anos, refletindo a pequena representatividade nos dados do TCGA do CMTN em mulheres com idade jovem. Como o tumor hereditário tende a ocorrer em idades mais jovens, esperávamos ter identificado diferenças nos percentuais de sobreposição dos dados do TCGA entre os grupos Hereditário e o Esporádicos. No entanto, não observamos essas diferenças. As proporções de sobreposição foram similares entre os dois grupos de amostras. De forma geral, nosso estudo poderá contribuir com dados transcricionais de CMTN em pacientes diagnosticadas em idade jovem.

O padrão de expressão diferencial, a partir do transcriptoma completo de todas as amostras de CMTN, revelou significativas vias canônicas ativadas associadas com regulação do ciclo celular, o que também foi observado em outros estudos (CRAIG et al. 2013). Também detectamos vias canônicas inibidas envolvidas com sinalização celular e metabolismo, das quais a de Ativação de LXR/RXR foi a mais estatisticamente significativa. LXR (do inglês, *Liver X Receptor*) e RXR (do inglês, *Retinoid X Receptor*) são receptores nucleares que trabalham juntos regulando o metabolismo lipídico, de colesterol e na modulação de resposta inflamatória (EDWARDS et al. 2002; ZELCER e TONTONOZ 2006; AGONZALEZ e CASTRILLO 2011), e novos estudos já mostraram seu envolvimento em diversas malignidades, além do CMTN (LIN e GUSTAFSSON 2015; CHEN et al. 2016; FANALE et al. 2017). CHENG et al. (2018) reportaram que, em câncer do endométrio, a inibição da via de ativação de LXR/RXR, dentre outras vias relacionadas com o metabolismo, está envolvida com o processo tumorigênico dessa

neoplasia. E, por outro lado, a ativação dessa via, através do uso de agonistas, é capaz de suprimir a proliferação de células tumorais em câncer de próstata *in vitro* e *in vivo* através da parada do ciclo celular (FUKUCHI et al. 2004; CHUU et al. 2007).

Outros importantes processos biológicos envolvidos com a tumorigênese do CMTN que identificamos foram replicação, recombinação e reparo do DNA, movimento celular, sinalização celular, doença metabólica, desordem hereditária e resposta imune e inflamatória. Destaque-se para o gene *ESRI*, o qual observamos ser o gene central nessa rede de interação, regulando o enriquecimento desses processos biológicos no CMTN. Embora o CMTN não apresente a expressão do RE, alguns estudos demonstraram que a sinalização do RE, como através da expressão do Receptor de Estrógeno acoplado à proteína G (*GPER*), está ativa no CMTN e envolvida com o crescimento tumoral assim como viabilidade e motilidade celular (STEIMAN et al. 2013; YU et al. 2014; SULAIMAN et al. 2018), mostrando a importante função da sinalização celular ativada por estrógeno no CMTN.

Além disso, observamos que os genes *FOXMI* e *CDKN2A* apresentaram uma importante função como reguladores transcricionais que regulam cascatas de sinalização importantes para o surgimento do CMTN. O *FOXMI* (do inglês, *Forkhead Box MI*) é um dos mais importantes oncogenes que contribui com a tumorigênese, progressão e resistência à droga em CMTN e que encontramos ser o regulador transcricional mais ativado nas nossas amostras de CMTN (HAMURCU et al. 2016). No mais, CRAIG et al. (2013), estudando o processo tumorigênico do CMTN, também reportaram o aumento de expressão do *FOXMI*, igualmente observado em outros tipos de câncer, como câncer de ovário, de mama, próstata, melanoma, colo retal, gástrico e de pulmão (ALVAREZ-FERNANDEZ et al. 2010; OKADA et al. 2013; KONG et al. 2014; ZHANG et al. 2016; ABDELJAOUED et

al. 2017; ITO et al. 2016; TASSI et al. 2017; LIU et al. 2017). *FOXMI* é um regulador central de muitos processos biológicos com potencial para se tornar um alvo terapêutico para muitos cânceres, inclusive para o CMTN (LIAO et al. 2018). Um outro gene, que identificamos como regulador transcricional e que também apresenta forte impacto para o desenvolvimento de terapia alvo, foi o *CDKN2A* (do inglês, Cyclin Dependent Kinase Inhibitor 2A), um supressor de tumor envolvido com a inibição do crescimento e divisão celular de forma descontrolada e que já foi observado por outros estudos estar deletado em CMTN (BALKO et al. 2014; ZHANG et al. 2014). A ausência de expressão desse gene gera importantes considerações clínicas, pois está associada com a resistência à quimioterapia neoadjuvante em CMTN (ARIMA et al. 2012; BALKO et al. 2014).

Quando os processos biológicos foram também investigados especificamente nos dois grupos das amostras, CMTN Hereditário e CMTN Esporádico, observamos que o enriquecimento das vias canônicas ativadas foi similar nos dois grupos, sendo vias envolvidas com a Regulação do ciclo celular. Por outro lado, as vias canônicas inibidas foram diferencialmente enriquecidas e influenciadas pela presença da mutação germinativa em *BRCA1*. Observamos um maior número de vias canônicas inibidas no grupo Hereditário em relação às vias inibidas no grupo Esporádico, mostrando uma maior instabilidade molecular em importantes processos biológicos envolvidos com a tumorigênese nesses tumores TN. Além disso, identificamos vias exclusivamente inibidas em cada grupo. Dentre estas, a Sinalização do receptor do Glutamato foi a via mais estatisticamente significativa apenas nesse grupo de CMTN Hereditário.

O glutamato é o principal substrato bioenergético para a proliferação das células em condições fisiológicas normais e o principal neurotransmissor excitatório

do sistema nervoso central. Assim, o glutamato tem demonstrado ter um papel não apenas nos processos neurais, como o aprendizado e a memória, mas também nas vias bioenergéticas, biossintéticas e metabólicas oncogênicas e sua sinalização pode estar envolvida no desenvolvimento e progressão de diversas malignidades (NEDERGAARD et al. 2002; WILLARD e KOOCHKEPOUR 2013). Em gliomas, sabe-se que a sinalização do glutamato é responsável pelo crescimento e invasão do tumor (YE e SONTHEIMER 1999; TAKANO et al. 2001) e, em câncer pancreático, foi associado à transformação maligna nas lesões precursoras no pâncreas (HERNER et al. 2011). Em câncer de próstata, alguns estudos reportaram a superexpressão dos receptores de glutamato, e o nível sérico do glutamato foi correlacionado com a agressividade da doença (KOOCHKEPOUR et al. 2012; WILLARD e KOOCHKEPOUR 2013). Em linhagens celulares de CMTN, já foi observado significativas quantidades de glutamato no ambiente extracelular (SEIDLITZ et al. 2009). Um outro estudo também reportou que a sinalização do glutamato em CMTN pode ser um importante alvo no tratamento da doença, pois medeia a regulação do crescimento, progressão e sobrevivência das células tumorais (SPEYER et al. 2012). Além disso, não apenas a ativação da sinalização do glutamato está associada à progressão tumoral. Já foi observado que, em alguns tipos de câncer, a inibição dessa sinalização é capaz de levar à inibição da proliferação e da migração das células tumorais (RZESKI et al. 2001). Neste mesmo trabalho, os autores mostraram que a inibição do receptor do glutamato com consequente inibição da sinalização, como observamos nas amostras de CMTN Hereditário, significativamente inibiu o crescimento de linhagens celulares de CMTN *in vivo*, sugerindo uma potencial estratégia alvo para essa doença e que pode ser melhor abordada por estudos futuros.

Quando analisamos as funções biológicas enriquecidas em redes de interação gênica, interessante mencionar que a Desordem hereditária esteve presente com maior significância estatística no CMTN Hereditário em relação às amostras do CMTN Esporádico, mostrando como este processo biológico pode apresentar um maior envolvimento na tumorigênese desse grupo de CMTN, já que são tumores com mutação patogênica germinativa em *BRCA1*. Além disso, a presença da mutação em *BRCA1* influenciou, significativamente, a expressão de genes de modo que observamos reguladores transcricionais exclusivos em cada grupo das amostras. Dentre estes, o *TNF* foi o regulador mais significativamente ativado apenas no grupo Hereditário. O *TNF* (do inglês, *Tumor Necrosis Factor*) codifica uma citocina pró-inflamatória envolvida em importantes efeitos celulares como apoptose, angiogênese e ativação do sistema imune (WAJANT 2009), o que mostra a importância do papel do processo inflamatório nesses tumores. Conforme visto por KIDANE et al. (2014), tumores com *BRCA1*-mutado apresentam um processo de inflamação acentuado, o que pode resultar em uma instabilidade genômica devido às espécies reativas de oxigênio (ROS, do inglês, *Reactive Oxygen Species*) presentes nesses tumores. Nesse sentido, o reparo ao DNA apresenta uma importante função em controlar o dano causado durante o processo inflamatório. Porém, em células onde o sistema de reparo é ineficiente, como tumores com *BRCA1*-mutado, esse dano é acumulado favorecendo a tumorigênese (KIDANE et al. 2014). Porém, alguns estudos já mostraram que a deleção de genes de reparo, como *BRCA1*, pode resultar em baixa expressão de citocinas pró-inflamatórias, como TNF-alfa, diminuindo os danos causados na inflamação (ZIELINSKI et al. 2003; MABLEY et al. 2005), diferentemente do nosso dado onde detectamos que o TNF está expresso ativamente nos tumores *BRCA1*-mutado. O papel do TNF, no processo inflamatório dos tumores

com recombinação homóloga deficiente, tem um importante papel clínico pois são tumores que estão sendo avaliados para o tratamento com inibidores de PARP, os quais levam à letalidade sintética e, conseqüentemente, ao aumento da produção de citocinas como o TNF que pode causar a morte celular (HUANG et al. 2015). Além disso, a ativação do TNF induz a permeabilidade vascular do tumor, facilitando a absorção de fármacos quimioterápicos, como a doxorrubicina usada no tratamento para CMTN (EGGERMONT et al. 2003). Além do TNF ser um importante regulador ativado que observamos nos tumores TN associado com *BRCA1*, também identificamos a expressão diminuída de genes não codificadores de proteínas, como do microRNA mir21, presente apenas nas amostras do grupo Hereditário, sugerindo o envolvimento dessa classe de RNAs no processo tumorigênico do CMTN com *BRCA1*-mutado, como também já observado por outro estudo (TANIC et al. 2011).

A análise dos processos biológicos envolvidos com a tumorigênese do CMTN Esporádico revelou um enriquecimento das vias canônicas ativadas semelhante com o CMTN Hereditário, sendo as mesmas envolvidas com a Regulação do ciclo celular. Por outro lado, a maioria das vias canônicas inibidas foram exclusivas de cada grupo, como já mencionado quando analisamos o CMTN Hereditário.

Nas amostras de CMTN Esporádico, a via de Ativação de LXR/RXR foi a mais estatisticamente inibida, exclusivamente, nesse grupo, como detectamos quando a análise foi feita com toda a casuística já que a mesma está enriquecida com esse grupo amostral. Corroborando com esse dado, o gene *PPARG*, envolvido com a ativação dessa via, foi identificado como o regulador transcricional mais estatisticamente inibido dentre os reguladores transcrpcionais exclusivos nesse grupo de amostra. O *PPARG* (do inglês, *Peroxisome Proliferator Activated Receptor Gamma*) codifica uma família de receptores nucleares (PPAR, do inglês *Peroxisome*

*Proliferator-Activated Receptor*), capaz de formar heterodímeros com os receptores retinóides RXRs que, por sua vez, regulam a transcrição de diversos genes e que também já foi mostrado estar inibido em câncer de próstata, cuja inibição está envolvida com a tumorigênese da doença (LIN et al. 2015). A família dos PPARs é também definida como moduladores seletivos que controlam a proliferação e diferenciação celular, funções biológicas tão importantes para o processo tumorigênico (CHINETTI et al. 2001; LORINCZ e SUKUMAR 2006). O *PPARG* é um regulador crítico na regulação do metabolismo de glicose e lipídeos, diferenciação de adipócitos e da sinalização intracelular de insulina (TANG et al. 2015). Estudos têm explorado a hipótese que a mutação em *PPARG* pode influenciar na tumorigênese e progressão tumoral em diversas malignidades (SABATINO et al. 2012; TANG et al. 2015; WANG et al. 2015), o que indica que a expressão inibida desse gene pode estar participando da tumorigênese do CMTN sem associação com *BRCA1*, como observamos neste trabalho. Em câncer de mama, a função dos ligantes de PPAR, como potenciais agentes antitumorais, já foi reportada (ELSTNER et al. 1998; BONOFIGLIO et al. 2006) além de ser investigada como efetivo alvo terapêutico nessa neoplasia (CROWE e CHANDRARATNA 2004). E, em CMTN, a via PPAR/RXR já foi observada como sendo importante na tumorigênese da doença, regulando o crescimento e sobrevivência celular (NARRANDES et al. 2018). No mais, o gene *PPARG* já foi demonstrado ser um biomarcador de resistência à fármacos em pacientes com CMTN que não respondem ao tratamento quimioterápico (WU et al. 2015).

Nesse mesmo grupo de CMTN Esporádico, também identificamos reguladores transcricionais ativados exclusivamente nesse grupo de amostras. Dentre estes, detectamos o gene *E2F1* como o regulador mais estatisticamente ativado. O

*E2F1* (do inglês, *E2F Transcription Factor*) apresenta função central na regulação do ciclo celular, como observamos no enriquecimento de vias envolvidas com a regulação do ciclo celular nesse grupo de amostra, sugerindo o envolvimento desse gene na tumorigênese do CMTN Esporádico, como também foi reportado por XIONG et al. (2018).

Como discutido anteriormente, uma vez que o CMTN é uma doença que apresenta uma grande diversidade transcricional refletindo na sua heterogeneidade molecular, abordagens de sequenciamento de nova geração para quantificar o transcriptoma completo, como o RNA-Seq, são de grande importância como ferramenta para melhor explorar e obter um maior entendimento da biologia desse tipo de tumor. Além disso, essas abordagens podem contribuir na elucidação da complexidade funcional de tumores TN, além de identificar uma diversidade de potenciais alvos terapêuticos ainda não estudados, principalmente no contexto da deficiência em *BRCA1*.



## 6 CONCLUSÃO

A partir dos dados obtidos com o desenvolvimento desse estudo, podemos concluir que:

- As amostras de CMTN foram classificadas nos 7 subtipos moleculares, sendo a maioria classificada nos perfis imunomodulador e mesenquimal. As amostras dos subtipos imunomodulador e *basal-like* 1/2 apresentaram uma tendência de melhor sobrevida global em relação às amostras dos outros subtipos moleculares. Em relação às amostras *BRCA1*-deficiente, observamos uma predominância do subtipo imunomodulador, enquanto que nas amostras *BRCA1*-proficiente, vimos que a maioria foi mesenquimal. Além disso, no grupo *BRCA1*-proficiente, detectamos que aquelas imunomoduladoras apresentaram melhor sobrevida global quando comparadas com as amostras classificadas nos outros subtipos. O mesmo não foi observado no grupo *BRCA1*-deficiente;
- O padrão de expressão global gerou 2039 genes diferencialmente expressos (GDEs) nas amostras de CMTN em relação às amostras normais pareadas, com 67,3% menos expressos no tumor. Houve uma maior predominância dos GDEs menos expressos no tumor no grupo dos genes não codificadores de proteínas do que no grupo dos genes codificadores de proteínas. Esses mesmos padrões foram observados na expressão diferencial dos grupos amostrais Hereditário e Esporádico;

- A análise funcional *in silico* revelou importantes processos biológicos envolvidos com a tumorigênese do CMTN, como ciclo celular, replicação, recombinação e reparo do DNA, sinalização celular, metabolismo e resposta imune e inflamatória. Além disso, detectamos os reguladores transcricionais *FOXMI* e o *CDKN2A*. Em relação aos processos biológicos nos grupos de CMTN Hereditário e CMTN Esporádico, a análise funcional revelou nenhuma diferença das vias canônicas ativadas. Porém, em relação à inibição das vias canônicas, observamos que a via de Sinalização do Receptor de Glutamato foi a mais estatisticamente significativa e exclusiva no grupo Hereditário, enquanto que a via de Ativação de LXR/RXR foi a mais estatisticamente significativa e exclusiva no grupo Esporádico. Observamos também importantes reguladores transcricionais únicos de cada grupo: *TNF* e *mir-21* foram os genes mais significativamente alterados presentes apenas no Hereditário; e os genes *E2F1* e *PPARG* foram os genes mais significativamente alterados apenas no Esporádico.

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

Abdeljaoued S, Bettaieb I, Nasri M, et al. Overexpression of FOXM1 is a potential prognostic marker in male breast cancer. **Oncol Res Treat** 2017; 40:167-72.

Adams S, Gray RJ, Demaria S, et al. Prognostic value of tumor-infiltrating lymphocytes in triple-negative breast cancers from two phase III randomized adjuvant breast cancer trials: ECOG 2197 and ECOG 1199. **J Clin Oncol** 2014; 32:2959-66.

Afghahi A, Telli ML, Kurian AW. Genetics of triple-negative breast cancer: Implications for patient care. **Curr Probl Cancer** 2016; 40:130-40.

A-Gonzalez N, Castrillo A. Liver X receptors as regulators of macrophage inflammatory and metabolic pathways. **Biochim Biophys Acta** 201; 1812:982-94.

Alvarez-Fernandez M, Halim VA, Krenning L, et al. Recovery from a DNA-damage-induced G2 arrest requires Cdk-dependent activation of FoxM1. **EMBO Rep** 2010; 11:452-8.

[ACS] American Cancer Society. **Cancer facts & figures 2017**. Atlanta 2017. Available from: <URL:<https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2017/cancer-facts-and-figures-2017.pdf>>. [2018 nov 12].

Anders S, Huber W. Differential expression analysis for sequence count data. **Genome Biol** 2010; 11:R106.

Anders S, Pyl PT, Huber W. HTSeq - a Python framework to work with high-throughput sequencing data. **Bioinformatics** 2015; 31:166-9.

Andre F, Job B, Dessen P, et al. Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. **Clin Cancer Res** 2009; 15:441-51.

Andreopoulou E, Kelly CM, McDaid HM. Therapeutic advances and new directions for triple-negative breast cancer. **Breast Care (Basel)** 2017; 12:21-8.

Aken BL, Ayling S, Barrell D, et al. The ensembl gene annotation system. **Database (Oxford)** 2016; 2016:baw093.

Antoniou A, Pharoah PDP, Narod S, et al. Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case Series unselected for family history: a combined analysis of 22 studies. **Am J Hum Genet** 2003; 72:1117-30.

Antoniou AC, Casadei S, Heikkinen T, et al. Breast-cancer risk in families with mutations in PALB2. **N Engl J Med** 2014; 371:497-506.

Arima Y, Hayashi N, Hayashi H, et al. Loss of p16 expression is associated with the stem cell characteristics of surface markers and therapeutic resistance in estrogen receptor-negative breast cancer. **Int J Cancer** 2012; 130:2568-79.

Ashworth A. A synthetic lethal therapeutic approach: poly(ADP) ribose polymerase inhibitors for the treatment of cancers deficient in DNA double-strand break repair. **J Clin Oncol** 2008; 26:3785-90.

Ayca G, Traina TA. Triple-negative breast cancer: role of the androgen receptor. **Cancer J** 2010; 16:62-5.

Balko JM, Giltane JM, Wang K, et al. Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. **Cancer Discov** 2014; 4:232-45.

Balmaña J, Tung NM, Isakoff SJ, et al. Phase I trial of olaparib in combination with cisplatin for the treatment of patients with advanced breast, ovarian and other solid tumors. **Ann Oncol** 2014; 25:1656-63.

Barber LJ, Sandhu S, Chen L, et al. Secondary mutations in BRCA2 associated with clinical resistance to a PARP inhibitor. **J Pathol** 2013; 229:422-9.

Baselga J, Gómez P, Greil R, et al. Randomized phase II study of the anti-epidermal growth factor receptor monoclonal antibody cetuximab with cisplatin versus cisplatin alone in patients with metastatic triple-negative breast cancer. **J Clin Oncol** 2013; 31:2586-92.

Bauer JA, Chakravarthy AB, Rosenbluth JM, et al. Identification of markers of taxane sensitivity using proteomic and genomic analyses of breast tumors from patients receiving neoadjuvant paclitaxel and radiation. **Clin Cancer Res** 2010; 16:681-90.

Bauer KR, Brown M, Cress RD, Parise CA, Caggiano V. Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple negative phenotype: a population-based study from the California Cancer Registry. **Cancer** 2007; 109:1721-8.

Bayraktar S, Gutierrez-Barrera AM, Liu D, et al. Outcome of triple-negative breast cancer in patients with or without deleterious BRCA mutations. **Breast Cancer Res Treat** 2011; 130:145-53.

Bear HD, Tang G, Rastogi M, et al. The effect on pCR of bevacizumab and/or antimetabolites added to standard neoadjuvant chemotherapy: NSABP protocol B-40. **J Clin Oncol** 2011; 29(suppl):A1005. [Presented at 2011 Annual Meeting, June 3-7, Chicago, Illinois, USA]

Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **J Royal Stat Soc Series B (Methodological)** 1995; 57:289-300.

Bernhardt SM, Dasari P, Walsh D, Townsend AR, Price TJ, Ingman WV. Hormonal modulation of breast cancer gene expression: implications for intrinsic subtyping in premenopausal women. **Front Oncol** 2016; 6:241.

Bertucci F, Finetti P, Cervera N, et al. How basal are triple-negative breast cancers? **Int J Cancer** 2008; 123:236-40.

Bianchini G, Qi Y, Alvarez RH, et al. Molecular anatomy of breast cancer stroma and its prognostic value in estrogen receptor-positive and -negative cancers. **J Clin Oncol** 2010; 28:4316-23.

Bonofiglio D, Aquila S, Catalano S, et al. Peroxisome proliferator-activated receptor-gamma activates p53 gene promoter binding to the nuclear factor-kappaB sequence in human MCF7 breast cancer cells. **Mol Endocrinol** 2006; 20:3083-92.

Bonsang-Kitzis H, Sadacca B, Hamy-Petit AS, et al. Biological network-driven gene selection identifies a stromal immune module as a key determinant of triple-negative breast carcinoma prognosis. **Oncoimmunology** 2016; 5:e1061176.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **Ca Cancer J Clin** 2018; 68:394-42.

Brekelmans CT, Tilanus-Linthorst MM, Seynaeve C, et al. Tumour characteristics, survival and prognostic factors of hereditary breast cancer from BRCA2-, BRCA1- and non-BRCA1/2 families as compared to sporadic breast cancer cases. **Eur J Cancer** 2007; 43:867-76.

Brianese RC, Nakamura KDM, Almeida FGDSR, et al. BRCA1 deficiency is a recurrent event in early-onset triple-negative breast cancer: a comprehensive analysis of germline mutations and somatic promoter methylation. **Breast Cancer Res Treat** 2018; 167:803-14.

Brockhoff G, Seitz S, Weber F, et al. The presence of PD-1 positive tumor infiltrating lymphocytes in triple negative breast cancers is associated with a favorable outcome of disease. **Oncotarget** 2017; 9:6201-12.

Bronner IF, Quail MA, Turner DJ, Swerdlow H. Improved protocols for illumina sequencing. **Curr Protoc Hum Genet** 2014; 80:18.2.1-42.

Bryant HE, Helleday T. Inhibition of poly (ADP-ribose) polymerase activates ATM which is required for subsequent homologous recombination repair. **Nucleic Acids Res** 2006; 34:1685-91.

Burstein MD, Tsimelzon A, Poage GM, et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. **Clin Cancer Res** 2015; 2:1688-98.

Buys SS, Sandbach JF, Gammon A, et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. **Cancer** 2017; 123:1721-30.

Byrski T, Gronwald J, Huzarski T, et al. Pathologic complete response rates in young women with BRCA1-positive breast cancers after neoadjuvant chemotherapy. **J Clin Oncol** 2010; 28:375-9.

Byrski T, Dent R, Blecharz P, et al. Results of a phase II open-label, non-randomized trial of cisplatin chemotherapy in patients with BRCA1-positive metastatic breast cancer. **Breast Cancer Res** 2012; 14:R110.

Byrski T, Huzarski T, Dent R, et al. Pathologic complete response to neoadjuvant cisplatin in BRCA1-positive breast cancer patients. **Breast Cancer Res Treat** 2014; 147:401-5.

Caestecker KW, Van de Walle GR. The role of BRCA1 in DNA double-strand repair: past and present. **Exp Cell Res** 2013; 319:575-87.

Caramelo O, Silva C, Caramelo F, Frutuoso C, Almeida-Santos T. The effect of neoadjuvant platinum-based chemotherapy in BRCA mutated triple negative breast cancers -systematic review and meta-analysis. **Hered Cancer Clin Pract** 2019; 17:11.

Carraro DM, Koike Folgueira MA, et al. Comprehensive analysis of BRCA1, BRCA2 and TP53 germline mutation and tumor characterization: a portrait of early-onset breast cancer in Brazil. **PLoS One** 2013; 8:e57581.

Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. **JAMA** 2006; 295:2492-502.

Carey LA, Dees EC, Sawyer L, et al. The triple negative paradox: primary tumor chemosensitivity of breast cancer subtypes. **Clin Cancer Res** 2007; 13:2329-34.

Carey LA, Rugo HS, Marcom PK, et al. TBCRC 001: randomized phase II study of cetuximab in combination with carboplatin in stage IV triple-negative breast cancer. **J Clin Oncol** 2012; 30:2615-23.

Chen S, Parmigiani G. Meta-analysis of BRCA1 and BRCA2 penetrance. **J Clin Oncol** 2007; 25:1329-33.

Chen X, Li J, Gray WH, et al. TNBCtype: a subtyping tool for triplenegative breast cancer. **Cancer Inform** 2012; 11:147-56.



Chen X, Wu J, Huang H, et al. Comparative profiling of triple-negative breast carcinomas tissue glycoproteome by sequential purification of glycoproteins and stable isotope labeling. **Cell Physiol Biochem** 2016; 38:110-21.

Cheng Y, Lv Q, Xie B, et al. Estrogen and high-fat diet induced alterations in C57BL/6 mice endometrial transcriptome profile. **Endocr Connect** 2018; 7:36-46.

Chinetti G, Lestavel S, Bocher V, et al. PPAR- $\alpha$  and PPAR- $\gamma$  activators induce cholesterol removal from human macrophage foam cells through stimulation of the ABCA1 pathway. **Nat Med** 2001; 7:53-8.

Chuu CP1, Kokontis JM, Hiipakka RA, Liao S. Modulation of liver X receptor signaling as novel therapy for prostate cancer. **J Biomed Sci** 2007; 14:543-53.

Clifton K, Gutierrez-Barrera A, Ma J, et al. Adjuvant versus neoadjuvant chemotherapy in triple-negative breast cancer patients with BRCA mutations. **Breast Cancer Res Treat** 2018; 170:101.

Couch FJ, Hart SN, Sharma P, et al. Inherited mutations in 17 breast cancer susceptibility genes among a large triple negative breast cancer cohort unselected for family history of breast cancer. **J Clin Oncol** 2015; 33:304-11.

Craig DW, O'Shaughnessy JA, Kiefer JA, et al. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. **Mol Cancer Ther** 2013; 12:104-16.

Crowe DL, Chandraratna RA. A retinoid X receptor (RXR)-selective retinoid reveals that RXR- $\alpha$  is potentially a therapeutic target in breast cancer cell lines, and that it potentiates antiproliferative and apoptotic responses to peroxisome proliferator-activated receptor ligands. **Breast Cancer Res** 2004; 6:R546.

Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. **Nature** 2012; 486:346-52.

Dedes KJ, Wilkerson PM, Wetterskog D, Weigelt B, Ashworth A, Reis-Filho JS. Synthetic lethality of PARP inhibition in cancers lacking BRCA1 and BRCA2 mutations. **Cell Cycle** 2011; 10:1192-9.

Denkert C, von Minckwitz G, Brase JC, et al. Tumor-infiltrating lymphocytes and response to neoadjuvant chemotherapy with or without carboplatin in human epidermal growth factor receptor 2-positive and triple-negative primary breast cancers. **J Clin Oncol** 2015; 33:983-91.

Dent R, Trudeau M, Pritchard KI, et al. Triple-negative breast cancer: clinical features and patterns of recurrence. **Clin Cancer Res** 2007; 13:4429-34.

Dent R, Hanna WM, Trudeau M, et al. Pattern of metastatic spread in triple-negative breast cancer. **Breast Cancer Res Treat** 2009; 115:423-8.

Dent RA, Lindeman GJ, Clemons M, et al. Phase I trial of the oral PARP inhibitor olaparib in combination with paclitaxel for first- or second-line treatment of patients with metastatic triple-negative breast cancer. **Breast Cancer Res** 2013; 15:1.

Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. **Nature** 2010; 464:999-1005.

Ding X, Zhu L, Ji T, et al. Long intergenic non-coding RNAs (LincRNAs) identified by RNA-seq in breast cancer. **PloS One** 2014; 9:e103270.

Dhillon KK, Swisher EM, Taniguchi T. Secondary mutations of BRCA1/2 and drug resistance. **Cancer Sci** 2011; 102:663-9.

Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. **Bioinformatics** 2013; 29:15-21.

Edwards PA, Kennedy MA, Mak PA. LXRs: oxysterol-activated nuclear receptors that regulate genes controlling lipid homeostasis. **Vascul Pharmacol** 2002; 38:249-56.

Edwards SL, Brough R, Lord CJ, et al. Resistance to therapy caused by intragenic deletion in BRCA2. **Nature** 2008; 451:1111-5.

Eggermont AM, de Wilt JH, ten Hagen TL. Current uses of isolated limb perfusion in the clinic and a model system for new strategies. **Lancet Oncol** 2003; 4:429-37.

Elstner E, Muller C, Koshizuka K, et al. Ligands for peroxisome proliferator-activated receptor gamma and retinoic acid receptor inhibit growth and induce apoptosis of human breast cancer cells in vitro and in BXD mice. **Proc Natl Acad Sci U S A** 1998; 95:8806-11.

ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. **Nature** 2012; 489:57-74.

Esserman LJ, Berry DA, Cheang MCU, et al. Chemotherapy response and recurrence-free survival in neoadjuvant breast cancer depends on biomarker profiles: results from the ISPY 1 TRIAL (CALGB 150007/150012; ACRIN 6657). **Breast Cancer Res Treat** 2012; 132:1049-62.

Esteller M, Silva JM, Dominguez G, et al. Promoter hypermethylation and BRCA1 inactivation in sporadic breast and ovarian tumors. **J Natl Cancer Inst** 2000; 92:564-9.

Farmer H, McCabe N, Lord CJ, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. **Nature** 2005; 434:917-21.

Fackenthal JD, Olopade OI. Breast cancer risk associated with BRCA1 and BRCA2 in diverse populations. **Nat Rev Cancer** 2007; 7:937-48.

Fanale D, Amodeo V, Caruso S. The interplay between metabolism, PPAR signaling pathway, and cancer. **PPAR Rese** 2017; 2017:1830626.

Feld E, Horn L. Targeting PD-L1 for non-small-cell lung cancer. **Immunotherapy** 2016; 8:747-58.

Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. **N Engl J Med** 2009; 361:123-34.

Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. **N Engl J Med** 2010; 363:1938-48.

Fukuchi J, Kokontis JM, Hiipakka RA, Chuu CP, Liao S. Antiproliferative effect of liver X receptor agonists on LNCaP human prostate cancer cells. **Cancer Res** 2004; 64:7686-9.

Gallagher DJ, Gaudet MM, Pal P, et al. Germline BRCA mutations denote a clinicopathologic subset of prostate cancer. **Clin Cancer Res** 2010; 16:2115-21.

Geenen JJJ, Linn SC, Beijnen JH Schellens JHM. PARP inhibitors in the treatment of triple-negative breast cancer. **Clin Pharmacokinet** 2018; 57:427-37.

Gerber B, Eidtman H, Rezai M, et al. Neoadjuvant bevacizumab and anthracycline-taxane-based chemotherapy in 686 triple-negative primary breast cancers: secondary endpoint analysis of the GeparQuinto study (GBG 44). **Ann Oncol** 2013; 24:2978-84.

Gelmon KA, Tischkowitz M, Mackay H, et al. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. **Lancet Oncol** 2011; 12:852-61.

Geyer FC, Pareja F, Weigelt B, Rakha E, Ellis IO, Schnitt SJ, Reis-Filho JS. The spectrum of triple-negative breast disease: high-and low-grade lesions. **Am J Pathol** 2017; 187:2139-51.

Glück S, Ross JS, Royce M, et al. TP53 genomics predict higher clinical and pathologic tumor response in operable early-stage breast cancer treated with docetaxelcapecitabine ± trastuzumab. **Breast Cancer Res Treat** 2012; 132:781-91.

Goldhirsch A, Wood WC, Coates AS, et al. Strategies for subtypes - dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. **Ann Oncol** 2011; 22:1736-47.

Gonzalez-Angulo AM, Timms KM, Liu S, et al. Incidence and outcome of BRCA mutations in unselected patients with triple receptor-negative breast cancer. **Clin Cancer Res** 2011; 17:1082-9.

Guarino M. Src signaling in cancer invasion. **J Cell Physiol** 2010; 223:14-26.

Ha G, Roth A, Lai D, et al. Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. **Genome Res** 2012; 22:1995-2007.

Habib JG, O'Shaughnessy J. The hedgehog pathway in triple-negative breast cancer. **Cancer Med** 2016; 5:2989-3006.

Hatzis C, Symmans WF2, Zhang Y, et al. Relationship between complete pathologic response to neoadjuvant chemotherapy and survival in triple-negative breast cancer. **Clin Cancer Res** 2016; 22:26-33.

Hamurcu Z, Ashour A, Kahraman N, Ozpolat B. FOXM1 regulates expression of eukaryotic elongation factor 2 kinase and promotes proliferation, invasion and tumorigenesis of human triple negative breast cancer cells. **Oncotarget** 2016; 7:16619-35.

Han W, Jung EM, Cho J, et al. DNA copy number alterations and expression of relevant genes in triple-negative breast cancer. **Genes Chromosomes Cancer** 2008; 47:490-9.

Hayes DF, Thor AD, Dressler LG, et al. HER2 and response to paclitaxel in node-positive breast cancer. **N Engl J Med** 2007; 357:1496-506.

Herner A, Sauliunaite D, Michalski CW, et al. Glutamate increases pancreatic cancer cell invasion and migration via AMPA receptor activation and Kras-MAPK signaling. **Int J Cancer** 2011; 129:2349-59.

Herschkowitz J, Simin K, Weigman V, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. **Genome Biol** 2007; 8:R76.

Hersey P, Gowrishankar K. Pembrolizumab joins the anti-PD-1 armamentarium in the treatment of melanoma. **Future Oncol** 2015; 11:133-40.

Hitzemann R, Bottomly D, Darakjian P, et al. Genes, behavior and next-generation RNA sequencing. **Genes Brain Behav** 2013; 12:1-12.

Huang F, Kushner YB, Langleben A, Foulkes WD. Eleven years disease-free: role of chemotherapy in metastatic BRCA2-related breast cancer. **Nat Rev Clin Oncol** 2009; 6:488-92.

Huang R, Jaritz M, Guenzl P, et al. An RNA-Seq strategy to detect the complete coding and non-coding transcriptome including full-length imprinted macro ncRNAs. **PLoS One** 2011; 6:e27288.

Huang J, Wang L, Cong Z, et al. The PARP1 inhibitor BMN 673 exhibits immunoregulatory effects in a Brca1(-/-) murine model of ovarian cancer. **Biochem Biophys Res Commun** 2015; 463:551-6.

Hu XC, Zhang J, Xu BH, et al. Cisplatin plus gemcitabine versus paclitaxel plus gemcitabine as first-line therapy for metastatic triple-negative breast cancer (CBCSG006): a randomised, open-label, multicentre, phase 3 trial. **Lancet Oncol** 2015; 16:436-46.

Huen MS, Sy SM, Chen J. BRCA1 and its toolbox for the maintenance of genome integrity. **Nat Rev Mol Cell Biol** 2010; 11:138-48.

Hugo W, Zaretsky JM, Sun L, et al. Genomic and transcriptomic features of response to anti-PD1 therapy in metastatic melanoma. **Cell** 2016; 165:35-44.

Ilott NE, Ponting CP. Predicting long non-coding RNAs using RNA sequencing. **Methods** 2013; 63:50-9.

Ito T, Kohashi K, Yamada Y, et al. Prognostic significance of forkhead box M1 (FoxM1) expression and antitumour effect of FoxM1 inhibition in melanoma. **Histopathology** 2016; 69:63-71.

Jiang T, Shi W, Wali VB, et al. Predictors of chemosensitivity in triple negative breast cancer: an integrated genomic analysis. **PLoS Med** 2016; 13:e1002193.

Joosse SA, Brandwijk KIM, Mulder L, et al. Genomic signature of BRCA1 deficiency in sporadic basal like breast tumors. **Genes Chromosom Cancer** 2011; 50:71-81.

Judes G, Rifai K, Daures M, et al. High-throughput <<Omics>> technologies: new tools for the study of triple-negative breast cancer. **Cancer Lett** 2016; 382:77-85.

Juul N, Szallasi Z, Eklund AC, et al. Assessment of an RNA interference screen-derived mitotic and ceramide pathway metagene as a predictor of response to neoadjuvant paclitaxel for primary triple-negative breast cancer: a retrospective analysis of five clinical trials. **Lancet Oncol** 2010; 11:358-65.

Karn T, Pusztai L, Rody A, Holtrich U, Becker S. The influence of host factors on the prognosis of breast cancer: stroma and immune cell components as cancer biomarkers. **Curr Cancer Drug Targets** 2015; 15:652-64.

Kidane D, Chae WJ, Czocho J, et al. Interplay between DNA repair and inflammation, and the link to cancer. **Crit Rev Biochem Mol Biol** 2014; 49:116-39.

Koboldt DC, Fulton RS, McLellan MD, et al. Comprehensive molecular portraits of human breast tumours. **Nature** 2012; 490:61-70.

Koduru SV, Tiwari AK, Leberfinger A, et al. A comprehensive ngs data analysis of differentially regulated miRNAs, piRNAs, lncRNAs and sn/snoRNAs in triple negative breast cancer. **J Cancer** 2017; 8:578-96.

Kong F-F, Qu Z-Q, Yuan H-H, et al. Overexpression of FOXM1 is associated with EMT and is a predictor of poor prognosis in non-small cell lung cancer. **Oncol Rep** 2014; 31:2660-8.



Koochekpour S, Majumdar S, Azabdaftari G, et al. Serum glutamate levels correlate with Gleason score and glutamate blockade decreases proliferation, migration, and invasion and induces apoptosis in prostate cancer cells. **Clin Cancer Res** 2012; 18:5888-901.

Kukurba KR, Montgomery SB. RNA sequencing and analysis. **Cold Spring Harb Protoc** 2015; 2015:951-69.

Lakhani SR, Gusterson BA, Jacquemier J, et al. The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. **Clin Cancer Res** 2000; 6:782-9.

Lee JM, Hays JL, Annunziata CM, et al. Phase I/Ib study of olaparib and carboplatin in BRCA1 or BRCA2 mutation associated breast or ovarian cancer with biomarker analyses. **J Natl Cancer Inst** 2014; 106:1-11.

Lee J, Kefford R, Carlino M. PD-1 and PD-L1 inhibitors in melanoma treatment: past success, present application and future challenges. **Immunotherapy** 2016; 8:733-46.

Le DT, Uram JN, Wang H, et al. PD-1 blockade in tumors with mismatch-repair deficiency. **N Engl J Med** 2015; 372:2509-20.

Levin JZ, Yassour M, Adiconis X, et al. Comprehensive comparative analysis of strand-specific RNA sequencing methods. **Nat Methods** 2010; 7:709-15.

Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics** 2009; 25:2078-9.

Li X, Yang J, Peng L, et al. Triple-negative breast cancer has worse overall survival and cause-specific survival than non-triple-negative breast cancer. **Breast Cancer Res Treat** 2017; 161:279-87.

Lin CY, Gustafsson JA. Targeting liver X receptors in cancer therapeutics. **Nat Rev Cancer** 2015; 15:216-24.

Lin SJ, Yang DR, Wang N, et al. TR4 nuclear receptor enhances prostate cancer initiation via altering the stem cell population and EMT signals in the PPAR $\gamma$ -deleted prostate cells. **Oncoscience** 2015; 2:142-50.

Lehmann BD, Bauer JA, Chen X, et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. **J Clin Invest** 2011; 121:2750-67.

Lehmann BD, Pietenpol JA. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. **J Pathol** 2014; 232:142-10.

Lehmann BD, Jovanović B, Chen X, et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. **PLoS One** 2016; 11:e0157368.

Ledermann J, Harter P, Gourley C, et al. Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer: a preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. **Lancet Oncol** 2014; 15:852-61.

Liedtke C, Mazouni C, Hess KR, et al. Response to neoadjuvant therapy and long-term survival in patients with triple-negative breast cancer. **J Clin Oncol** 2008; 26:1275-81.

Lips EH, Mulder L, Oonk A, et al. Triple-negative breast cancer: BRCAness and concordance of clinical features with BRCA1-mutation carriers. **Br J Cancer** 2013; 108:2172-7.

Liao GB, Li XZ, Zeng S, et al. Regulation of the master regulator FOXM1 in cancer. **Cell Commun Signal** 2018; 16:57.

Liu H, Murphy CJ, Karreth FA, et al. Identifying and targeting sporadic oncogenic genetic aberrations in mouse models of triple-negative breast cancer. **Cancer Discov** 2018; 8:354-69.

Liu YR, Jiang YZ, Xu XE, et al. Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. **Breast Cancer Res** 2016; 18:33.

Liu Y, Liu Y, Yuan B, et al. FOXM1 promotes the progression of prostate cancer by regulating PSA gene transcription. **Oncotarget** 2017; 8:17027-37.

Liu Z, Li M, Jiang Z, Wang X. A comprehensive immunologic portrait of triple-negative. **Breast Cancer Transl Oncol** 2018; 11:311-29.

Locatelli M, Curigliano G. Notch inhibitors and their role in the treatment of triple negative breast cancer: promises and failures. **Curr Opin Oncol** 2017; 29:411-27.

Loi S, Michiels S, Salgado R, et al. Tumor infiltrating lymphocytes are prognostic in triple negative breast cancer and predictive for trastuzumab benefit in early breast cancer: results from the FinHER trial. **Ann Oncol** 2014; 25:1544-50.

Lord CJ, Ashworth A. BRCAness revisited. **Nat Rev Cancer** 2016; 16:110-20.

Lorincz A, Sukumar S. Molecular links between obesity and breast cancer. **Endocr Relat Cancer** 2006; 13:279-92.

Mabley JG, Pacher P, Deb A, Wallace R, Elder RH, Szabó C. Potential role for 8-oxoguanine DNA glycosylase in regulating inflammation. **FASEB J** 2005; 19:290-2.

Maxwell KN, Domcheck SM. Familial breast cancer risk. **Curr Breast Cancer Rep** 2013; 5:170-82.

Masuda H, Baggerly KA, Wang Y, et al. Differential response to neoadjuvant chemotherapy among 7 triple-negative breast cancer molecular subtypes. **Clin Cancer Res** 2013; 19:5533-40.

Massihnia D, Galvano A, Fanale D, et al. Triple negative breast cancer: shedding light onto the role of pi3k/akt/mtor pathway. **Oncotarget** 2016; 7:60712.

Mavaddat N, Peock S, Frost D, et al. Cancer risks for BRCA1 and BRCA2 mutation carriers: results from prospective analysis of EMBRACE. **J Natl Cancer Inst** 2013; 105:812-22.

Mersch J, Jackson MA, Park M, et al. Cancers associated with BRCA1 and BRCA2 mutations other than breast and ovarian. **Cancer** 2015; 121:269-75.

Ministério da Saúde. Instituto Nacional de Câncer José de Alencar Gomes da Silva. **Estimativa 2016/incidência de câncer no Brasil**. Rio de Janeiro: INCA; 2016.

Ministério da Saúde. Instituto Nacional de Câncer José de Alencar Gomes da Silva. **Estimativa 2018/incidência do Câncer no Brasil**. Rio de Janeiro: INCA; 2018.

Ministério da Saúde. Instituto Nacional de Câncer José de Alencar Gomes da Silva. **Câncer de mama: fatores de risco**. Última modificação: 05/02/2019. Disponível em: <URL:<https://www.inca.gov.br/tipos-de-cancer/cancer-de-mama>>. [2019 jan 10].

Miki Y, Swensen J, Shattuck-Eidens D, et al. strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. **Science** 1994; 266:66-71.

Mittendorf EA, Philips AV, Meric-Bernstam F, et al. PD-L1 expression in triple-negative breast cancer. **Cancer Immunol Res** 2014; 2:361-70.

Mylavarapu S, Das A, Roy M. Role of BRCA mutations in the modulation of response to platinum therapy. **Front Oncol** 2018; 8:16.

Narrandes S, Huang S, Murphy L, Xu W. The exploration of contrasting pathways in Triple Negative Breast Cancer (TNBC). **BMC Cancer** 2018; 18:22.

Nakai K, Hung MC, Yamaguchi H. A perspective on anti-EGFR therapies targeting triple-negative breast cancer. **Am J Cancer Res** 2016; 6:1609-23.

Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole genome sequences. **Nature** 2016; 534:47-54.

Nanda R, Chow LQM, Dees EC, et al. Pembrolizumab in patients with advanced triple-negative breast cancer: phase Ib KEYNOTE-012 study. **J Clin Oncol** 2016; 34:2460-7.

Nedergaard M, Takano T, Hansen AJ. Beyond the role of glutamate as a neurotransmitter. **Nat Rev Neurosci** 2002; 3:748-55.

Norquist B, Wurz KA, Pennil CC, et al. Secondary somatic mutations restoring BRCA1/2 predict chemotherapy resistance in hereditary ovarian carcinomas. **J Clin Oncol** 2011; 29:3008-15.

Oakman C, Viale G, Di Leo A. Management of triple negative breast cancer. **Breast** 2010; 19: 312-21.

Okada K, Fujiwara Y, Takahashi T, et al. Overexpression of forkhead box M1 transcription factor (FOXM1) is a potential prognostic marker and enhances chemoresistance for docetaxel in gastric cancer. **Ann Surg Oncol** 2013; 20:1035-43.

O'Leary NA, Wright MW, Brister JR, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. **Nucleic Acids Res** 2016; 44:D733-45.

Olivieri EH, Franco Lde A, Pereira RG, Mota LD, Campos AH, Carraro DM. Biobanking practice: RNA storage at low concentration affects integrity. **Biopreserv Biobank** 2014; 12:46-52.

Pahuja S, Beumer JH, Appleman LJ, et al. Outcome of BRCA 1/2-mutated (BRCA+) and triple-negative, BRCA wild type (BRCA-wt) breast cancer patients in a phase I study of single-agent veliparib (V). **J Clin Oncol** 2014; 32(suppl 26):A135. [Presented at 2014 Annual Meeting, May 30 to June 2, Chicago, Illinois, USA]

Palazzo AF, Lee ES. Non-coding RNA: what is functional and what is junk? **Front Genet** 2015; 6:2.

Papadimitriou M, Theodosopoulos T, Baltas D, Papadimitriou CA. Therapeutic targets for metastatic triple-negative breast cancer: focus on PARP inhibition. **Global J Breast Cancer Res** 2017; 5:12-21.

Park HS, Jang MH, Kim EJ, Park SY. High EGFR gene copy number predicts poor outcome in triple-negative breast cancer. **Mod Pathol** 2014; 27:1212-22.

Park JH, Lee HJ, Ahn J, et al. Prognostic significance of tumor-infiltrating lymphocytes (TILs) in patients with early-stage triple-negative breast cancer (TNBC) treated with curative resection alone. **Ann Oncol** 2017; 28(suppl\_10):x16-24.

Parkin DM. Global cancer statistics in the year 2000. **Lancet Oncol** 2001; 2:533-43.

Paluch-Shimon S, Friedman E, Berger R, et al. Neo-adjuvant doxorubicin and cyclophosphamide followed by paclitaxel in triple-negative breast cancer among BRCA1 mutation carriers and non-carriers. **Breast Cancer Res Treat** 2016; 157:157-65.

Perou CM, Sùrlie T, Eisen MB, et al. Molecular portraits of human breast tumors. **Nature** 2000; 533:747-52.

Petrucelli N, Daly MB, Feldman GL. Hereditary breast and ovarian cancer due to mutations in BRCA1 and BRCA2. **Genet Med** 2010; 12:245-59.

Pharoah PDP, Antoniou A, Bobrow M, Zimmern RL, Easton DF, Ponder BAJ. Polygenic susceptibility to breast cancer and implications for prevention. **Nat Genet** 2002; 31:33-6.

Pohl SG, Brook N, Agostino M, Arfuso F, Kumar AP, Dharmarajan A. Wnt signaling in triple-negative breast cancer. **Oncogenesis** 2017; 6:e310.

Prat A, Parker JS, Karginova O, et al. Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. **Breast Cancer Res** 2010; 12:R68.

Pellegrino B, Bella M, Michiara M, et al. Triple negative status and BRCA mutations in contralateral breast cancer: a population-based study. **Acta Biomed** 2016; 87:54-63.

Rampurwala M, Wisinski KB, O'Regan R. Role of the androgen receptor in triplenegative breast cancer. **Clin Adv Hematol Oncol** 2016; 14:186-93.

Reis-Filho JS, Tutt AN. Triple negative tumours: a critical review. **Histopathology** 2008; 52:108-18.

Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology: mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. **Science** 2015; 348:124-8.

Robertson L, Hanson H, Seal S, et al. BRCA1 testing should be offered to individuals with triple-negative breast cancer diagnosed below 50 years. **Br J Cancer** 2012; 106:1234-8.

Robson M, Im SA, Senkus E, et al. Olaparib for metastatic breast cancer in patients with a germline BRCA mutation. **N Engl J Med** 2017; 377:523-33.

Rosen EM, Fan S, Pestell RG, Goldberg ID. BRCA1 gene in breast cancer. **J Cell Physiol** 2003; 196:19-41.

Rottenberg S, Nygren AO, Pajic M, et al. Selective induction of chemotherapy resistance of mammary tumors in a conditional mouse model for hereditary breast cancer. **Proc Natl Acad Sci USA** 2007; 104:12117-22.

Roy R, Chun J, Powell SN. BRCA1 and BRCA2: different roles in a common pathway of genome protection. **Nat Rev Cancer** 2012; 12:68-78.

Rummel S, Varner E, Shriver CD, Ellsworth RE. Evaluation of BRCA1 mutations in an unselected patient population with triple-negative breast cancer. **Breast Cancer Res Treat** 2013; 137:119-25.

Rzeski W, Turski L, Ikonomidou C. Glutamate antagonists limit tumor growth. **Proc Natl Acad Sci U S A** 2001; 98:6372-7.

Sabatino L, Fucci A, Pancione M. UHRF1 coordinates peroxisome proliferator activated receptor gamma (PPARG) epigenetic silencing and mediates colorectal cancer progression. **Oncogene** 2012; 31:5061-72.



Sakai W, Swisher EM, Jacquemont C, et al. Functional restoration of BRCA2 protein by secondary BRCA2 mutations in BRCA2-mutated ovarian carcinoma. **Cancer Res** 2009; 69:6381-6.

Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, et al. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. **BMC Mol Biol** 2006; 7:3-16.

Seidlitz EP, Sharma MK, Saikali Z, Ghert M, Singh G. Cancer cell lines release glutamate into the extracellular environment. **Clin Exp Metastasis** 2009; 26:781-7.

Severson TM, Peeters J, Majewski I, et al. BRCA1-like signature in triple negative breast cancer: molecular and clinical characterization reveals subgroups with therapeutic potential. **Mol Oncol** 2015; 9:1528-38.

Sikov WM, Berry DA, Perou CM, et al. Impact of the addition of carboplatin (Cb) and/or bevacizumab (B) to neoadjuvant weekly paclitaxel (P) followed by dose-dense AC on pathologic complete response (pCR) rates in triple-negative breast cancer (TNBC): CALGB 40603 (Alliance). **J Clin Oncol** 2015; 33:13-21.

Silver DP, Richardson AL, Eklund AC, et al. Efficacy of neoadjuvant cisplatin in triple-negative breast cancer. **J Clin Oncol** 2010; 28:1145-53.

Sikandar B, Qureshi MA, Naseem S, Khan S, Mirza T. Increased tumour infiltration of CD4+ and CD8+ T-lymphocytes in patients with triple negative breast cancer suggests susceptibility to immune therapy. **Asian Pac J Cancer Prev** 2017; 18:1827-32.

Sims D, Sudbery I, Illott NE, Heger A, Ponting CP. Sequencing depth and coverage: key considerations in genomic analyses. **Nat Rev Genet** 2014; 15:121-32.

Shah SP, Roth A, Goya R, et al. The clonal and mutational evolution spectrum of primary triple-negative breast cancers. **Nature** 2012; 486:395-9.

Sharma P, Klemp JR, Kimler BF, et al. Germline BRCA mutation evaluation in a prospective triple-negative breast cancer registry: implications for hereditary breast and/or ovarian cancer syndrome testing. **Breast Cancer Res Treat** 2014; 145:707-14.

Silva FC, Lisboa BC, Figueiredo MC, et al. Hereditary breast and ovarian cancer: assessment of point mutations and copy number variations in Brazilian patients. **BMC Med Genet** 2014; 15:55.

Snyder A, Makarov V, Merghoub T, et al. Genetic basis for clinical response to CTLA-4 blockade in melanoma. **N Engl J Med** 2014; 371:2189-99.

Solinas C, Gombos A, Latifyan S, Piccart-Gebhart M, Kok M, Buisseret L. Targeting immune checkpoints in breast cancer: an update of early results. **ESMO Open** 2017; 2:e000255.

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

Sørli T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. **Proc Natl Acad Sci USA** 2003; 100:8418-23.

Speyer CL, Smith JS, Banda M, DeVries JA, Mekani T, Gorski DH. Metabotropic glutamate receptor-1: a potential therapeutic target for the treatment of breast cancer. **Breast Cancer Res Treat** 2012; 132:565-73.

Stanton SE, Adams S, Disis ML. Variation in the incidence and magnitude of tumor-infiltrating lymphocytes in breast cancer subtypes: a systematic review. **JAMA Oncol** 2016; 2:1354-60.

Steiman J, Peralta EA, Louis S, Kamel O. Biology of the estrogen receptor, GPR30, in triple negative breast cancer. **Am J Surg** 2013; 206:698-703.

Stephens PJ, Tarpey PS, Davies H, et al. The landscape of cancer genes and mutational processes in breast cancer. **Nature** 2012; 486:400-4.

Su J, Blumenthal GM, Jiang X, He K, Keegan P, Pazdur R. FDA approval summary: pembrolizumab for the treatment of patients with metastatic non-small cell lung cancer whose tumors express programmed death-ligand 1. **Oncologist** 2016; 21:643-50.

Steward BW, Wild CP. **World Cancer Report 2014**. Lyon: International Agency for Research on Cancer; 2014.

Sturtz LA, Melley J, Mamula K, et al. Outcome disparities in African American women with triple negative breast cancer: a comparison of epidemiological and molecular factors between African American and Caucasian women with triple negative breast cancer. **BMC Cancer** 2014; 14:62.

Sulaiman A, McGarry S, Lam KM, et al. Co-inhibition of mTORC1, HDAC and ESR1 $\alpha$  retards the growth of triple-negative breast cancer and suppresses cancer stem cells. **Cell Death Dis** 2018; 9:815.

Swisher EM, Sakai W, Karlan BY, Wurz K, Urban N, Taniguchi T. Secondary BRCA1 mutations in BRCA1-mutated ovarian carcinomas with platinum resistance. **Cancer Res** 2008; 68:2581-6.

Székely B, Silber AL, Pusztai L, Székely B, Silber AL, Pusztai L. New therapeutic strategies for triple-negative breast cancer. **Oncology** (Williston Park) 2017; 31:130-7.

Tanic M, Zajac M, Gómez-López G, Benítez J, Martínez-Delgado B. Integration of BRCA1-mediated miRNA and mRNA profiles reveals microRNA regulation of TRAF2 and NFκB pathway. **Breast Cancer Res Treat** 2012; 134:41-51.

Tang W, Chen Y, Wang Y, Gu H, Chen S, Kang M. Peroxisome proliferator-activated receptor gamma (PPARG) polymorphisms and breast cancer susceptibility: a meta-analysis. **Int J Clin Exp Med** 2015; 8:12226-38.

Takano T, Lin JH, Arcuino G, Gao Q, Yang J, Nedergaard M. Glutamate release promotes growth of malignant gliomas. **Nat Med** 2001; 7:1010-5.

Tassi RA, Todeschini P, Siegel ER, et al. FOXM1 expression is significantly associated with chemotherapy resistance and adverse prognosis in non-serous epithelial ovarian cancer patients. **J Exp Clin Cancer Res** 2017; 36:63.

Tassone P, Di Martino MT, Ventura M, et al. Loss of BRCA1 function increases the antitumor activity of cisplatin against human breast cancer xenografts in vivo. **Cancer Biol Ther** 2009; 8:648-53.

Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C. An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. **Genome Biol** 2007; 8:R157.

The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. **Nature** 2012; 490:61-70.

Thompson D, Easton DF. Breast cancer linkage consortium. cancer incidence in BRCA1 mutation carriers. **J Natl Cancer Inst** 2002; 94:1358-65.

Tung NM, Winer EP. Tumor-infiltrating lymphocytes and response to platinum in triple-negative breast cancer. **J Clin Oncol** 2015; 33:969-71.

Turner N, Tutt A, Ashworth A. Hallmarks of “BRCAness” in sporadic cancers. **Nat Rev Cancer** 2004; 4:814-9.

Turner N, Lambros MB, Horlings HM, Pearson A, Sharpe R, Natrajan R, et al. Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. **Oncogene** 2010; 29:2013-23.

Tutt A, Robson M, Garber JE, Domchek S. Phase II trial of the oral PARP inhibitor olaparib in BRCA-deficient advanced breast cancer. **J Clin Oncol** 2009; 27:CRA501.

Tutt A, Robson M, Garber JE, et al. Oral poly (ADP-ribose) polymerase inhibitor olaparib in patients with BRCA1 or BRCA2 mutations and advanced breast cancer: a proof-of-concept trial. **Lancet** 2010; 376:235-44.

Tutt A, Tovey H, Cheang MCU, et al. Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT trial. **Nat Med** 2018; 24:628-37.

van Allen EM, Miao D, Schilling B, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. **Science** 2015; 350:207-11.

Venkitaraman AR. Cancer suppression by the chromosome custodians, BRCA1 and BRCA2. **Science** 2014; 343:1470-5.

von Minckwitz G, Martin M. Neoadjuvant treatments for triple-negative breast cancer (TNBC). **Ann Oncol** 2012; 23 Suppl 6:vi35-9.

von Minckwitz G, Untch M, Blohmer JU, et al. Definition and impact of pathologic complete response on prognosis after neoadjuvant chemotherapy in various intrinsic breast cancer subtypes. **J Clin Oncol** 2012; 30:1796-804.

Von Minckwitz G, Schneeweiss A, Salat C. A randomized phase II trial investigating the addition of carboplatin to neoadjuvant therapy for triple-negative and HER2-positive early breast cancer (GeparSixto). **J Clin Oncol** 2013; 31(suppl 15):A1004. [Presented at 2013 Annual Meeting, May 31 to June 4, Chicago, Illinois, USA]

von Minckwitz G, Schneeweiss A, Loibl S, et al. Neoadjuvant carboplatin in patients with triple-negative and HER2-positive early breast cancer (GeparSixto; GBG 66): a randomised phase 2 trial. **Lancet Oncol** 2014; 15:747-56.

Wajant H. The role of TNF in cancer. **Results Probl Cell Differ** 2009; 49:1-15.

Walsh CS. Two decades beyond BRCA1/2: homologous recombination, hereditary cancer risk and a target for ovarian cancer therapy. **Gynecol Oncol** 2015; 137:343-50.

Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. **Nat Rev Genet** 2009; 10:57-63.

Wang Y, Chen Y, Jiang H, Peroxisome proliferator-activated receptor gamma (PPARG) rs1801282 C>G polymorphism is associated with cancer susceptibility in asians: an updated meta-analysis. **Int J Clin Exp Med** 2015; 8:12661-73.

West NR, Milne K, Truong PT, Macpherson N, Nelson BH, Watson PH. Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. **Breast Cancer Res** 2011; 13:R126.

Weisman PS, Ng CK, Reis-Filho JS, Wen HY. Genetic alterations of triple negative breast cancer by targeted next-generation sequencing and correlation with tumor morphology. **Mod Pathol** 2016; 29:476-88.

Wilhelm BT, Landry JR. RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing. **Methods** 2009; 48:249-57.

Willard SS, Koochekpour S. Glutamate, glutamate receptors, and downstream signaling pathways. **Int J Biol Sci** 2013; 9:948-59.

Wooster R, Neuhausen SL, Mangion J, et al. Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. **Science** 1994; 265:2088-90.

Wu T, Wang X, Li J, et al. Identification of personalized chemoresistance genes in subtypes of basal-like breast cancer based on functional differences using pathway analysis. **PLoS One** 2015; 10:e0131183.

Wushou A, Jiang YZ, Hou J, Liu YR, Guo XM, Shao ZM. Development of triple-negative breast cancer radiosensitive gene signature and validation based on transcriptome analysis. **Breast Cancer Res Treat** 2015; 154:57-62.

Xiong Z, Ye L, Zhenyu H, et al. ANP32E induces tumorigenesis of triple-negative breast cancer cells by upregulating E2F1. **Mol Oncol** 2018; 12:896-912.

Xu H, Eirew P, Mullaly SC, Aparicio S. The omics of triple-negative breast cancers. **Clin Chem** 2014; 60:122-33.

Ye ZC, Sontheimer H. Glioma cells release excitotoxic concentrations of glutamate. **Cancer Res** 1999; 59:4383-91.

Yu KD, Shao ZM. Initiation, evolution, phenotype and outcome of BRCA1 and BRCA2 mutation-associated breast cancer. **Nat Rev Cancer** 2012; 12:372-3.

Yu T, Liu M, Luo H. et al. GPER mediates enhanced cell viability and motility via non-genomic signaling induced by 17beta-estradiol in triple-negative breast cancer cells. **J Steroid Biochem Mol Biol** 2014; 143:392-403.

Yoshida K, Miki Y. Role of BRCA1 and BRCA2 as regulators of DNA repair, transcription, and cell cycle in response to DNA damage. **Cancer Sci** 2004; 95:866-71.

Xiao Y, Ma D, Zhao S, Suo C, et al. Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. **Clin Cancer Res** 2019 Mar 5. [Epub ahead of print]

Zelcer N, Tontonoz P. Liver X receptors as integrators of metabolic and inflammatory signaling. **J Clin Invest** 2006; 116:607-14.

Zhang H, Zhong H, Li L, Ji W, Zhang X. Overexpressed transcription factor FOXM1 contributes to the progression of colorectal cancer. **Mol Med Rep** 2016; 13:2696-700.

Zhang L, Wang XI, Zhang S. Tumor infiltrating lymphocyte volume is a better predictor of neoadjuvant therapy response and overall survival in triple negative invasive breast cancer. **Hum Pathol** 2018; 80:47-54.

Zhang S, Shao Y, Hou G, et al. QM-FISH analysis of the genes involved in the G1/S checkpoint signaling pathway in triple-negative breast cancer. **Tumour Biol** 2014; 35:1847-5.

Zhu Q, Pao GM, Huynh AM, et al. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. **Nature** 2011; 477:179-84.



Zielinski CC, Budinsky AC, Wagner TM, et al. Defect of tumour necrosis factor-alpha (TNF-alpha) production and TNF-alpha-induced ICAM-1-expression in BRCA1 mutations carriers. **Breast Cancer Res Treat** 2003; 81:99-105.

Anexo 1 - Carta de aprovação do Comitê de Ética em Pesquisa-CEP



A.C. Camargo  
Cancer Center

Comitê de Ética em  
Pesquisa - CEP

São Paulo, 07 de outubro de 2014.

A  
Dra. Dirce Maria Carraro.


Aluna: Kivvi Duarte de Mello Nakamura (Doutorado).

Ref.: Projeto de Pesquisa nº. 1746/13C  
"Aspectos moleculares envolvidos no surgimento do Tumor Triplo-Negativo de mama em pacientes portadores ou não de mutação germinativa em BRCA1".

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – Hospital do Câncer - A.C. Camargo/SP, em sua última reunião de 23/09/2014, **tomaram conhecimento e aprovaram** o seguinte documento:

- Solicitação de dispensa da submissão da documentação obrigatória e análise ética do projeto acima mencionado por se tratar de um projeto afiliado ao temático intitulado "Aspectos moleculares envolvidos no desenvolvimento e progressão no carcinoma *in situ* e do papel da mutação em BRCA1 no Tumor Triplo-Negativo", registrado neste CEP sob nº 1746/13. O projeto afiliado em referência será a tese de Doutorado da aluna *Kivvi Duarte de Mello Nakamura*.
- Projeto de Doutorado – datado de 2014.

Atenciosamente,



Prof. Dr. Luiz Paulo Kowalski  
Coordenador do Comitê de Ética em Pesquisa

1/1