**FACULDADE DE MEDICINA
UNIVERSIDADE DE SÃO PAULO**

MAIRA ANDREA VALOYES VALOYES

**Avaliação do papel funcional e do potencial valor prognóstico dos
membros da família PHLDA (*Pleckstrin homology-like domain A*)
utilizando data mining**

SÃO PAULO

2018

MAIRA ANDREA VALOYES VALOYES

**Avaliação do papel funcional e do potencial valor prognóstico dos membros da
família *PHLDA* (*Pleckstrin homology-like domain A*) utilizando *data mining***

Dissertação apresentada à Faculdade de
Medicina da Universidade de São Paulo
para obtenção do título de Mestre em
Ciências

Programa de Oncologia

Orientadora: Profa. Dra. Maria Aparecida
Nagai

SÃO PAULO

2018

FICHA CATALOGRÁFICA

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Valoyes, Maira Andrea Valoyes

Avaliação do papel funcional e do potencial valor
prognóstico dos membros da família PHLDA (Pleckstrin
homology-like domain A) utilizando data mining /
Maira Andrea Valoyes Valoyes. -- São Paulo, 2018.

Dissertação (mestrado) -- Faculdade de Medicina da
Universidade de São Paulo.

Programa de Oncologia.

Orientadora: Maria Aparecida Nagai.

Descritores: 1.Neoplasias mamárias 2.Mineração de
dados 3.Expressão genica 4.Metilação 5.Bases de
dados

USP/FM/DBD-104/18

Responsável: Kátia Maria Bruno Ferreira - CRB-8/6008

DEDICATORIA

Dedico esse trabalho,

Inicialmente a **Deus** por ser minha força para continuar, por sua orientação
e por seu amor infinito.

A minha formosa **Família**, meus pais por seu apoio e seu amor
incondicional, as minhas irmãs **Gina, Angie e Wendy** por ser meu motor, e
me encorajar para conseguir minhas metas.

A todas as pessoas que direta ou indiretamente contribuíram para que eu
chegasse até aqui.

AGRADECIMENTOS

Agradeço principalmente a **DEUS**, meu TUDO, meu motor, meu guia, meu principal orientador neste caminho, sem ele absolutamente NADA disso seria possível.

Aos meus pais **Felipe e Andrea**, pela compreensão, pelo amor, pela imensa paciência, pelo apoio tanto econômico como emocional, vocês foram minha principal motivação.

Às minhas irmãs **Gina, Angie e Wendy** por estarem ali sempre para mim, por ser um exemplo, também foram minha grande motivação não tenho como agradecer-lhes tudo o que vocês fizeram por mim em todo este tempo.

À **Dra. Maria Aparecida Nagai**, pela oportunidade, ajuda, orientação e por me apresentar o universo da ciência.

Ao **Professor Alexandre** por me adotar como uma aluna do seu grupo, pela paciência, dedicação, pelos conhecimentos compartilhados e que foram chaves para concluir este trabalho.

À **Denilson e Mauro** pela ajuda nas análises, por me auxiliarem quando mais precisei generosamente.

À **Betty, Carol e Flavia** por auxiliarem-me quando precisei, pelos conselhos, aprendizado e acompanhamento neste processo.

Aos meus colegas de laboratório **Natália, Naieli, Lourival, Maria, Monica, Aline, Erika, Renan e Andrés** pelo companheirismo, por fazerem das jornadas no laboratório mais divertidas, vocês são pessoas maravilhosas agradeço a Deus por tê-los conhecido.

À **David** por estar sempre ao meu lado, por dar-me forças para continuar, por todo seu apoio e compreensão quando mais precisei.

À **Liz** por ser minha irmã neste tempo, por sua ajuda, os sorrisos, as conversas, as histórias repetidas, as alegrias e as palavras de conforto.

Aos **alpinitos** minha segunda família neste país, meninos foi lindo conviver todo este tempo com vocês, obrigada porquê com cada um de vocês aprendi muitas coisas.

À **pós-graduação** pela oportunidade de cumprir esse meu sonho que era fazer mestrado nesta excelente instituição, ao ICESP minha escola, minha casa, meu segundo lar, obrigada por providenciar tudo que precisei para fazer deste projeto uma realidade.

À **Dona Iara e Debora** pela ajuda em tudo, pela disposição para responder a cada pergunta.

À todos **os pacientes** com câncer por me motivar a fazer parte desta incessante luta contra esta doença.

Ao **Brasil** inteiro por ter-me acolhido como uma Brasileira, esta experiência foi única para mim, me sinto feliz de ter conhecido tantas pessoas especiais. Neste país cresci como pessoa, como mulher a nível espiritual e emocional, me levo grandes lembranças daqui.

Às agências de fomento **FAPESP e CNPq/CAPES** pelo apoio financeiro.

“Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito. Não sou o que deveria ser, mas Graças a Deus, não sou o que era antes”.

Martin Luther King

RESUMO

Valoyes MAV. *Avaliação do papel funcional e do potencial valor prognóstico dos membros da família PHLDA (Pleckstrin homology-like domain A) utilizando data mining* [dissertação]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2018.

O câncer de mama é uma doença complexa que envolve alterações genéticas e epigenéticas junto com fatores ambientais. Dentre os tipos de câncer o de mama é o de maior incidência e mortalidade na população feminina, portanto a caracterização desta doença tem sido um desafio contínuo para a comunidade científica. Numerosos biomarcadores moleculares tem sido associados com o câncer de mama e seus subtipos; no entanto, existe uma boa quantidade que tem sido pouco explorado. Este é o caso dos genes da família PHLDA (*pleckstrin homology like domain family A*), identificados repetidamente em vários estudos de perfil transcricional em câncer de mama. Esta família compreende três genes: o gene *PHLDA1* expresso em diferentes tecidos e com papel na regulação da apoptose, o gene *PHLDA2* localizado em uma importante região de localização de genes supressores de tumor que sofrem regulação por *imprinting*, e o gene *PHLDA3* envolvido na apoptose mediada por *p53*. Diversos estudos têm mostrado que os genes da família PHLDA estão frequentemente alterados em diferentes tipos de tumores, incluindo o câncer de mama, no entanto, o papel desses genes no desenvolvimento e progressão do câncer de mama ainda não está bem estabelecido. Na atualidade há uma enorme quantidade de dados experimentais integrados e disponibilizados em bancos de dados públicos para análise de dados genéticos e epigenéticos em diferentes tipos de tumores. Muitos dos estudos depositados nesses bancos tem facilitado a identificação de novos subtipos intrínsecos de câncer de mama, a predição de sobrevida e a resposta a drogas. Nosso objetivo é fazer uma busca em bases de dados públicos, para avaliar o padrão de expressão da família PHLDA em tumores e em linhagens celulares de mama, do perfil de metilação, mutação e do potencial valor prognóstico da expressão desses genes em câncer de mama. Para isso foram utilizadas as plataformas *TCGA*, *GEO* e *GOBO* para a busca de dados de expressão de mRNA, *TCGA* e *GEO* para extração de dados de metilação de DNA, *TCGA* e *COSMIC* para análise de dados de mutação e CNA, *miRTarBASE* e *GEO* para dados de microRNA em câncer de mama, *Netdecoder* para a construção das redes e *KMplotter* para a busca de dados de sobrevida. Os resultados mostraram que *PHLDA1* se encontra com baixa expressão em tumor e em linhagens celulares de mama e mais expresso nos tumores ER negativos. Os tumores ER negativos também mostraram um baixo nível de metilação quando comparados com amostra normal. Adicionalmente nos encontramos que *PHLDA1* é alvo do *microRNA* miR-181a-5p cujos altos níveis de expressão em tumores foram associados com baixa sobrevida. *PHLDA2* foi mais expresso nos tumores do que em amostra normal de mama, principalmente nos subtipos her2+; estes níveis elevados de expressão foram associados com baixa sobrevida em quase todos os subtipos moleculares; por outro lado, nenhuma diferença foi encontrada no perfil de metilação por receptor de estrógeno. O *microRNA* miR-193b-3p foi identificado como regulador de *PHLDA2*. O gene *PHLDA3* foi encontrado menos expresso nos tumores ER negativos, e esses tumores também se encontravam com hipermetilação. Pacientes com todos os subtipos moleculares

apresentaram um aumento de sobrevida livre de recorrência quando os níveis de *PHLDA3* foram altos. Nenhum membro da família apresentou mutações nos dados analisados, enquanto que alterações no número de cópia foram encontradas nos três genes. Os dados obtidos até o momento mostram que a expressão dos membros da família *PHLDA* é alterada em câncer de mama e tem impacto na sobrevida dos pacientes. Processos como metilação do DNA, alterações no número de cópia e a participação de microRNAs podem ser os mecanismos implicados na desregulação desses genes.

Descritores: neoplasias mamarias; mineração de dados; expressão genica; metilação; bases de dados

ABSTRACT

Valoyes MAV. *Evaluation of the functional role and the potential prognostic value of the members of the PHLDA family (pleckstrin homology like domain family A) using data mining* [thesis] São Paulo: “Faculdade de Medicina, Universidade de São Paulo”; 2018.

Breast cancer is a complex disease involving genetic and epigenetic alterations together with environmental factors. Among the cancer types, breast cancer is the most incident and deadliest in women population worldwide in both developed and developing countries, therefore, characterization of this disease has been a continuous challenge for the scientific community. A large quantity of molecular biomarkers has been associated with breast cancer development and their subtypes. However, there is a good amount that has been little explored. This is the case of the genes of the PHLDA family (pleckstrin homology-like domain family A) previously identified in a series of transcriptional profiling studies and recognized for their role in apoptosis. This family comprises three members; PHLDA1 is expressed in different tissues and has an important role in apoptosis regulation; PHLDA2 is located in a region harboring important tumor suppressor genes and is regulated by imprinting process. PHLDA3 gene is involved in p53-mediated apoptosis. Several studies have shown that the genes of the PHLDA family are frequently altered in different types of tumors including breast cancer. However, the role of these genes in breast cancer progression and development is not well established yet. Currently, there is a vast amount of genomic, transcriptomic, proteomic and epigenetic data generated by new high-performance technologies available in public databases. Many of the studies deposited in these banks have facilitated the identification of new intrinsic subtypes of breast cancer, prediction of survival and drug responses. Regarding the potential role of PHLDA family as biomarkers and the limited information in breast cancer, we pretend in this study to take advantage of these platforms downloading information about expression, methylation, of PHLDA family and correlating it with prognosis in breast cancer. We used TCGA, GEO and GOBO platforms to look for mRNA expression data, TCGA and GEO to extract methylation data, TCGA and COSMIC to analyze mutation and CNA, miRTarBase and GEO for microRNA analysis, KMplotter to assess prognosis and NetDecoder to construct the networks. The results showed that PHLDA1 is downregulated in tumors and breast cell line compared with breast tissue and it was more expressed in ER negative tumors. These tumors also showed a low level of methylation when compared to normal tissue. Additionally, the mining of miRNA revealed that PHLDA1 is target of two microRNAs, miR-181a-5p, whose high levels of expression in tumors were associated with low survival. On the other hand, the

transcripts of PHLDA2 were more expressed in tumors than in normal sample, mainly in HER2+ subtype, and PHLDA2 high expression was associated with poor outcome in almost all the subtypes. PHLDA2 is methylated only on ER tumors and was found to be target of the microRNA has-miR-193b-3p. PHLDA3 was less expressed in ER-negative tumors, and these tumors also exhibited DNA hypermethylation. Patients with tumors expressing high levels of PHLDA3 showed better recurrence-free survival than patients with low levels of PHLDA3. No family member had mutations in the analysed data, while copy number changes were found in the three genes. The data obtained so far show that the expression of PHLDA1 family members is altered in breast cancer and has an impact on the survival of patients. Processes such as DNA methylation, copy number changes and the involvement of microRNAs may be the mechanisms involved in the deregulation of these genes.

Descriptors: breast neoplasm; data mining; gene expression; methylation; databases

LISTA DE FIGURAS

Figura 1. Modelo hipotético de iniciação e progressão do câncer de mama.	24
Figura 2. Estrutura da proteína dos três membros da família PHLDA.	34
Figura 3. Análise de correlação da expressão de mRNA e status de metilação entre os membros da família PHLDA em câncer de mama.	51
Figura 4. Análise comparativa de expressão de mRNA, proteína e status de metilação da família PHLDA no banco de dados TCGA.	52
Figura 5. Expressão de mRNA e perfil de metilação do gene PHLDA1 em amostra normal comparada com tumor de mama	53
Figura 6. Expressão de mRNA de PHLDA1 em subtipos moleculares de câncer de mama.	55
Figura 7. Expressão de mRNA e perfil de metilação de PHLDA1 em tumores de mama estratificados pelo receptor de estrógeno.....	56
Figura 8. Perfil de expressão de mRNA de PHLDA1 em linhagens celulares de câncer de mama.....	57
Figura 9. Frequência de mutação e análise de CNA de PHLDA1 em câncer de mama..	58
Figura 10. Expressão do microRNA hsa-miR-181a-5p em câncer de mama.....	60
Figura 11. Curva de sobrevida para pacientes com todos os subtipos moleculares estratificados pelo nível de expressão do gene PHLDA1.....	61
Figura 12. Valor prognóstico do gene PHLDA1 em câncer de mama. Análise de sobrevida em estratificada em 5 anos (long-rank).....	63
Figura 13. Interações gênicas de PHLDA1	64
Figura 14. Expressão de mRNA e perfil de metilação do gene PHLDA2 em amostra normal comparada com tumor de mama.	65
Figura 15. Expressão de mRNA de PHLDA2 em subtipos moleculares de câncer de mama.	66
Figura 16. Expressão de mRNA e perfil de metilação de PHLDA2 em tumores de mama estratificados pelo receptor de estrógeno.....	67
Figura 17. Perfil de expressão de mRNA de PHLDA2 em linhagens celulares de câncer de mama.....	68
Figura 18. Frequência de mutação e análise de CNA do gene PHLDA2 em câncer de mama..	69

Figura 19. Expressão do microRNA hsa-miR-193b-3p em câncer de mama.	70
Figura 20. Curva de sobrevida para os diferentes subtipos moleculares estratificados pelo nível de expressão do gene PHLDA2.	71
Figura 21. Valor prognóstico do gene PHLDA2 em câncer de mama. Análise de sobrevida estratificada em 5 anos.	73
Figura 22. Expressão de mRNA e perfil de metilação do gene PHLDA3 em amostra normal comparada com tumor de mama	74
Figura 23. Expressão de mRNA de PHLDA3 em subtipos moleculares de câncer de mama.	75
Figura 24. Expressão de mRNA e perfil de metilação de PHLDA3 em tumores de mama estratificados pelo receptor de estrógeno.	76
Figura 25. Perfil de expressão de mRNA de PHLDA3 em linhagens celulares de câncer de mama.	77
Figura 26. Frequência de mutação e análise de CNA de PHLDA3 em câncer de mama.	78
Figura 27. Curva de sobrevida para pacientes com todos os subtipos estratificados pelo nível de expressão do gene PHLDA3.	79
Figura 28. Valor prognóstico do gene PHLDA3 em câncer de mama. Análise de sobrevida estratificada em 5 anos.	80

LISTA DE TABELAS

Tabela 1. Bancos de dados para mineração de dados em câncer.....	29
Tabela 2. Conjuntos de dados usados nas análises	49

LISTA DE ABREVIATURAS E SIGLAS

mRNA RNA mensageiro

RNA ácido ribonucleico

WHO Organização Mundial de Saúde (World Health Organization)

vs versus

INCA Instituto Nacional do Câncer

DNA ácido desoxirribonucléico

SUMÁRIO

SUMÁRIO	16
1. INTRODUÇÃO	18
1.1 Câncer de mama	18
1.2 Morfogênese da glândula mamária; processo de iniciação e progressão do câncer de mama. .	23
1.3 Classificação molecular do câncer de mama.....	25
1.4 Mineração de dados	27
1.5 Família PHLDA	33
1.6 PHLDA1	34
1.7 PHLDA2	37
1.8 PHLDA3	40
2. OBJETIVO GERAL	42
2.1 Objetivos Específicos	42
3. MATERIAL E MÉTODOS	43
3.1 ANÁLISE DE EXPRESSÃO GENICA	43
3.2 ANÁLISE DE METILAÇÃO	45
3.3 CURVAS DE SOBREVIVÊNCIA.....	46
3.4 MicroRNA.....	47
ANÁLISE ESTATÍSTICA.....	49
4. RESULTADOS	50
4.1 Análise comparativa de expressão, metilação e proteína da família PHLDA em câncer de mama ..	50
4.2 Análise do perfil de expressão e de metilação do gene PHLDA1 em câncer de mama.....	52
4.2.1 Análise do perfil de expressão do gene PHLDA1 por subtipo molecular de câncer de mama. .	54
4.2.2 Análise do perfil de expressão do gene PHLDA1 por status do receptor de estrogênio em câncer de mama	55
4.2.3 Análise do perfil de expressão do gene PHLDA1 em linhagens celulares de câncer de mama	57
4.2.4 Análise de mutação e CNA no gene PHLDA1 em câncer de mama	58
4.2.5 Análise de microRNAs alvos de PHLDA1 em câncer de mama	59
4.2.6 Análise de sobrevivência do gene PHLDA1 em câncer de mama	60
4.2.7 Vias de interação do gene PHLDA1 em câncer de mama	64
4.3 Análise do perfil de expressão e de metilação do gene PHLDA2 em câncer de mama.....	65
4.3.1 Análise do perfil de expressão do gene PHLDA2 por subtipo molecular de câncer de mama .	66
4.3.2 Análise do perfil de expressão do gene PHLDA2 por status do receptor de estrogênio em câncer de mama.	67
4.3.3 Análise do perfil de expressão do gene PHLDA2 em linhagens celulares de câncer de mama	68
4.3.4 Análise de mutação e CNA no gene PHLDA2 em câncer de mama	68
4.3.5 Análise de microRNAs alvos de PHLDA2 em câncer de mama	69
4.3.6 Análise de sobrevivência do gene PHLDA2 em câncer de mama	70
4.4 Análise do perfil de expressão e de metilação do gene PHLDA3 em câncer de mama.....	74

4.4.1	Análise do perfil de expressão do gene PHLDA3 por subtipo molecular de câncer de mama	75
4.4.2	Análise do perfil de expressão do gene PHLDA3 por status do receptor de estrógeno em câncer de mama	76
4.4.3	Análise do perfil de expressão do gene PHLDA3 em linhagens celulares de câncer de mama	77
4.4.4	Análise de mutação e CNA no gene PHLDA3 em câncer de mama	77
4.4.5	Análise de sobrevivência do gene PHLDA3 em câncer de mama	78
4.4.6	Vias de interação do gene PHLDA3 em câncer de mama	81
5.	DISCUSSÃO	83
6.	CONCLUSÃO.....	97
7.	REFERENCIAS	98

1. INTRODUÇÃO

1.1 Câncer de mama

O câncer é uma doença de múltiplos passos envolvendo alterações genéticas e epigenéticas (DOUGLAS HANAHA* AND ROBERT A. WEINBERG, 2000) e que juntamente com os fatores ambientais contribuem para a heterogeneidade e complexidade da doença. Dos tipos de câncer, o de mama possui a maior incidência e a maior mortalidade na população feminina em todo o mundo com um estimado de 522 mil mortes por ano (FERLAY et al., 2015). Estes índices elevados de mortalidade são atribuídos ao diagnóstico tardio da doença que varia regionalmente e em cada país e difere entre hospitais públicos e privados, o que pode ser devido a fatores socioeconômicos como a carência de recursos, sistemas de saúde ineficientes, e às limitações no acesso ao rastreamento do câncer mamário (RENCK et al., 2014). Estima-se que nos países de alta renda, a taxa média de sobrevivência atinge 80%, em países de média renda cerca de 60% e em países de baixa renda de 40% (World Health Organization, 2013; INCA, 2014).

Para o Brasil, em 2016, foram esperados 57.960 casos novos de câncer de mama, com um risco estimado de 56,20 casos a cada 100 mil mulheres sendo as regiões sul e sudeste as mais incidentes (INCA, 2016). Estima-se que no Brasil apenas cerca de 20% dos casos de câncer de mama sejam diagnosticados nos estágios iniciais (GOSS et al., 2013) o que sugere a importância do rastreamento ativo desta doença, como forma de diagnóstico precoce, prevenção e redução dos índices de mortalidade (RENCK et al., 2014).

O desenvolvimento do câncer de mama tem origem multifatorial que pode envolver tanto fatores endógenos e não modificáveis, de predisposição hereditária ou dependente da constituição hormonal, como também de fatores exógenos e modificáveis relacionados com o estilo de vida, a vida reprodutiva e outros fatores externos capazes de favorecer o desenvolvimento do carcinoma (AMENDOLA; VIEIRA, 2005).

O risco de aparecimento do câncer de mama em mulheres incrementa com a idade durante a etapa reprodutiva. Dados epidemiológicos têm mostrado que a incidência de câncer de mama é relativamente baixa até antes dos 30 anos de idade enquanto que 85% dos casos aparecem após os 45 anos de idade. Esta relação é justificada pelo acúmulo de lesões no tecido mamário produto do processo contínuo de proliferação e diferenciação das células da mama normal, durante o ciclo de vida da mulher confirmando o papel hormonal na etiologia da doença (KLUTTIG; SCHMIDT-POKRZYWNIAK, 2009; RENCK et al., 2014).

A herança genética é o principal fator de risco não modificável para o câncer de mama representando em torno de 5 a 10% dos tumores de mama de tipo familiar. Os genes melhor caracterizados e mais frequentemente associados com câncer de mama de tipo familiar são os genes *BRCA1* e *BRCA2*. Estes genes desempenham funções relacionadas com diferentes processos celulares, incluindo ativação e regulação transcricional, o reparo de lesões no DNA, controle do ciclo celular, proliferação e diferenciação celular (BAI et al., 2014; BIRNEY et al., 2007; WANG; DI, 2014); Mutações nesses genes são de alta penetrância e apresentam um risco estimado de desenvolvimento da doença de 60 a 80%. (AMENDOLA; VIEIRA, 2005; DOUGLAS HANAHAN* AND ROBERT A. WEINBERG, 2000).

Depois da predisposição genética a exposição aos estrógenos constitui um dos mais importantes fatores de risco para o câncer de mama. A associação entre os hormônios e o risco de câncer de mama tem sido sustentada em numerosos estudos epidemiológicos, em animais e *in vitro*, incluindo experimentos de redução do risco após ooforectomia bilateral em camundongos (SAMAVAT; KURZER, 2015).

Os estrógenos são hormônios esteroídicos produzidos pelos folículos ovarianos maduros e cuja principal função é o desenvolvimento e diferenciação de órgãos reprodutivos como os tecidos mamários e os ovários. O efeito dos estrógenos é mediado principalmente pelos receptores de estrógenos (ER) que consistem de proteínas intracelulares que atuam como fatores de transcrição ligando-se no DNA para promover a proliferação celular (LOUIE; SEVIGNY, 2017).

Tumores com diferentes níveis de expressão dos receptores de estrógenos apresentam diferentes respostas a terapias, tempos de sobrevida e curso clínico da doença; por esta razão têm-se considerado como um dos mais importantes marcadores para a classificação molecular do câncer de mama. Existem duas teorias propostas para explicar o mecanismo molecular pelo qual os estrógenos influenciam o desenvolvimento de câncer de mama. A hipótese mais aceita explica que devido a que os estrógenos a traves dos seus receptores levam à proliferação celular, a exposição prolongada da glândula mamária a estes hormônios pode aumentar as chances de mutações, podendo, portanto, tornar as células susceptíveis à transformação maligna. (SANTEN et al., 2009).

A segunda hipótese propõe que os estrógenos podem ser metabolizados a compostos genotóxicos com geração de espécies reativas de oxigênio as quais podem danificar o DNA causando desta forma mutações (SANTEN et al., 2009). O resultado de ambos os processos pode levar então à alteração do ciclo celular, interrupção da

apoptose e reparação do DNA, três das 10 características essenciais propostas por Hanahan e Weinberg para o desenvolvimento neoplásico (HANAHAN; WEINBERG, 2011).

Diferentes níveis de exposição dos estrógenos na glândula mamária ao longo da etapa reprodutiva como a menarca precoce e a gravidez são chaves na determinação do risco de desenvolvimento do câncer mama. O início da produção estrogênica pelos ovários em mulheres jovens é determinado pela primeira menarca e é um importante indicador de risco para o câncer de mama (KLUTTIG; SCHMIDT-POKRZYWNIAK, 2009). Em mulheres com menarca precoce, a exposição aos hormônios no tecido mamário acaba sendo mais prolongada durante o tempo, o que aumenta as chances de aparecimento de mutações nas células. A menopausa por outro lado indica a cessação da produção estrogênica ovariana, portanto, a menopausa tardia também está associada com aumento de risco. Alguns estudos estimam que a cada ano de atraso na menopausa o risco aumenta em 3% (KLUTTIG; SCHMIDT-POKRZYWNIAK, 2009) e a cada diminuição de um ano na idade na menarca o risco de câncer de mama pode aumentar em até 5% (COLDITZ; BOHLKE; BERKEY, 2014; LIU et al., 2013).

A gravidez tardia é considerada um fator de risco para câncer de mama. O principal motivo se deve a que a gravidez precoce, induz diminuição no número de células luminais sensíveis aos hormônios, redução dos níveis de prolactina e de estrógenos e aumento dos níveis de globulina transportadora de hormônios sexuais, o que se traduz em um fator de proteção contra o câncer de mama (COLDITZ; BOHLKE; BERKEY, 2014); portanto, quanto maior o intervalo entre a menarca e a primeira gravidez, maior o risco de desenvolver câncer de mama.

Fatores ambientais ou externos como consumo de álcool e tabaco, alimentação, radiação e baixa atividade física são considerados fatores de risco que quando modificados podem reduzir substancialmente o risco de câncer de mama. Diversos estudos indicam que a ingestão de alimentos ricos em soja, frutas e vegetais podem diminuir a incidência de câncer de mama por possuírem propriedades antioxidantes que podem influir no reparo do DNA, na desintoxicação metabólica e no potencial antiestrogênico (FALAVIGNA et al., 2014), entretanto, os alimentos ricos em gorduras estão associados com aumento do risco (KLUTTIG; SCHMIDT-POKRZYWNIK, 2009). Radiações ionizantes têm sido reportadas de alto risco mesmo em baixas doses de exposição, especialmente na puberdade (VERONESI et al., 2005), pois nesse período o tecido mamário ainda não está totalmente diferenciado (KLUTTIG; SCHMIDT-POKRZYWNIK, 2009).

A relação entre obesidade e o risco de câncer de mama depende fortemente do estado da menopausa. Presume-se que em mulheres na pré-menopausa, a obesidade pode proteger contra o câncer de mama, fazendo os ciclos menstruais anovulatórios mais frequentes, o que por sua vez resulta em diminuição dos níveis de estradiol e progesterona. (KLUTTIG; SCHMIDT-POKRZYWNIK, 2009) por outro lado, na pós-menopausa, existe uma produção aumentada de estrógenos derivada principalmente da aromatização de precursores androgênicos no tecido adiposo e diminuição do hormônio sexual ligado à globulina (SHGB) (DE PERGOLA; SILVESTRIS, 2013; KEY et al., 2003).

O consumo de álcool é considerado pela Agência Internacional para Pesquisa sobre Câncer (IARC) como um fator predisponente para o câncer de mama (IARC, 2010). Segundo vários estudos por cada 10 gramas de álcool consumidos diariamente por mulheres adultas pode haver um aumento do risco de 7 até 10% (KEY;

VERKASALO; BANKS, 2001; KLUTTIG; SCHMIDT-POKRZYWNIAK, 2009; LIU et al., 2013; WENDY Y. CHEN, MD, MPH; BERNARD ROSNER et al., 2011). O fumo do cigarro por outro lado, contém aproximadamente 7000 compostos químicos dos quais 69 são carcinógenos e 20 são considerados carcinógenos mamários, (REYNOLDS, 2013).

1.2 Morfogênese da glândula mamária; processo de iniciação e progressão do câncer de mama.

As glândulas mamárias são órgãos secretores complexos compostos principalmente pelos ductos mamários, os lóbulos, e as células do estroma (MACIAS; HINCK, 2012). O ducto mamário maduro se compõe de um epitélio basal que contém uma camada de células mioepiteliais que formam uma barreira física que separa os compartimentos epiteliais e do estroma (POLYAK; KALLURI, 2010); e de um epitélio luminal formado por uma capa interior de células epiteliais luminais que se diferenciam em ductos e alvéolos produtores de leite (GJOREVSKI; NELSON, 2011; MACIAS; HINCK, 2012).

O estroma é uma estrutura dinâmica de suporte composta principalmente por tecido adiposo infiltrado por células endoteliais vasculares, fibroblastos e células do sistema imune (MACIAS; HINCK, 2012; WATSON; KHALED, 2008).

O desenvolvimento da glândula mamária compreende vários processos que obedecem a mudanças na etapa reprodutiva da mulher. Na puberdade, a glândula passa por um processo de morfogênese caracterizado por elevados níveis de hormônios que estimulam o alongamento, ramificação e bifurcação lateral dos ductos, até atingir o limite da camada de gordura (GJOREVSKI; NELSON, 2011; WATSON;

KHALED, 2008). Durante a gravidez após o aumento dos níveis de progesterona o epitélio alveolar se expande em forma de ramificação e se diferencia em alvéolos os quais sintetizam e secretam leite durante a lactação. Posteriormente após remodelamento do tecido, no processo de involução o processo de apoptose remove mais de 80% do epitélio; este processo de remodelação continua é repetido no ciclo gravidez-aleitamento-involução graças à estrutura dinâmica da glandula mamária (GJOREVSKI; NELSON, 2011).

Figura 1. Modelo hipotético de iniciação e progressão do câncer de mama. Os dutos da mama normal estão compostos da membrana basal e uma camada de células epiteliais luminas e mioepiteliais. As células do estroma estão compostas por leucócitos, fibroblastos, miofibroblastos e células endoteliais. Em carcinomas *in situ* as células mioepiteliais são epigeneticamente e fenotipicamente alteradas, cujo número diminui potencialmente devido à degradação da membrana basal. Ao mesmo tempo, o número de fibroblastos do estroma, miofibroblastos, linfócitos e células endoteliais aumentam. Perda de células mioepiteliais e membrana basal resultam em carcinomas invasivos, no qual as células tumorais podem invadir os tecidos circundantes e podem migrar para órgãos distantes, eventualmente levando a metástases (Adaptado de Poliak, 2007).

De acordo com o modelo da evolução clonal o processo de iniciação de câncer começa pelo acúmulo de alterações genéticas e epigenéticas numa única célula que lhe conferem vantagem seletiva com respeito às células normais (CAMPBELL; POLYAK, 2007; POLYAK, 2007). No estágio denominado carcinoma ductal *in situ*, as células epiteliais malignas estão confinadas ao sistema ductal da mama, sem evidências de invasão através da membrana basal no estroma circundante (SIZIOPIKOU, 2013). À medida que o tumor progride as células adquirem mutações adicionais que lhes conferem características próprias do câncer como capacidade proliferativa, resistência à apoptose, sobrevivência, invasão e metástase (HANAHAHAN; WEINBERG, 2011). Posteriormente as células mioepiteliais transformadas migram ao estroma onde em conjunto degradam a membrana basal resultando na progressão de carcinoma *in situ* a invasivo. Neste estágio as células podem invadir tecidos adjacentes e migrar a órgãos distantes eventualmente desenvolvendo metástase (POLYAK, 2007).

1.3 Classificação molecular do câncer de mama.

O advento das técnicas moleculares para análise de expressão gênica na última década transformou o entendimento sobre a diversidade biológica do câncer de mama redirecionando as decisões no tratamento clínico (ZEPEDA-CASTILLA, 2008). Ao longo do tempo a classificação tradicional que inicialmente era baseada na patologia foi refinada com a análise de expressão gênica o que permitiu a melhor compreensão e estratificação da doença.

Perou e colaboradores foram os pioneiros em analisar centenas de genes utilizando a técnica de cDNA *microarray* a partir de amostras de tecido mamário normal e tumoral de 42 indivíduos diferentes (PEROU et al., 2000). Esses pesquisadores correlacionaram as características de variação fenotípica com dados de expressão gênica surgindo a partir desse estudo cinco subtipos moleculares. Luminal A, luminal B, basal-*like*, her2+ e normal *like*.

O subtipo luminal A é o mais comum dos cinco subtipos representando entre 50-60% dos casos (Eroles; 2012). É caracterizado pela expressão dos receptores de estrógeno (ER) e progesterona (PR), citoqueratinas epiteliais luminais 8 e 18 e baixa expressão do marcador de proliferação celular Ki67 (menor a 14%). Estes tumores tendem a ser de baixo grau histológico, a ter baixas taxas proliferativas e baixas taxas de recidiva (PUSZTAI, 2006). O tratamento para este grupo de pacientes está baseado em terapias hormonais e possui o melhor prognóstico dentre os cinco subtipos (YERSAL; BARUTCA, 2014).

O subtipo Luminal B é classificado com base na expressão do ER e/ou PR, a expressão ou não de HER2 e citoqueratinas 8/18. Os subtipos luminais B possuem um fenótipo mais agressivo com alta taxa proliferativa e de alto grau principalmente pela expressão de genes altamente proliferativos como MKI67 (maior a 14%), ciclina B1, EGFR, FGFR1, PI3K e SRC. Compreendem entre 15-20% dos tumores de mama e apresentam um prognóstico menos favorável em comparação com os tumores luminais A (EROLES et al., 2012).

Os tumores de tipo basal ou triplos negativos representam de 10-20% dos tumores mamários (EROLES et al., 2012). A característica mais importante desses tumores é que eles contêm pouca ou nenhuma expressão de ER, PR, e HER2, aumento de expressão de genes característicos das células da camada basal epitelial

como citoqueratina 5, 6 e 14, p-cadherina EGFR e uma alta taxa de mutações no gene p53. Os tumores de tipo basal são frequentes em mulheres jovens, afroamericanas e pre-menopausadas e se apresentam com alta agressividade e alto índice metastático. A maioria dos tumores basais apresenta alta sensibilidade à quimioterapia, no entanto, eles têm o pior prognóstico dentre os cinco subtipos (ALIZART et al., 2012; PEROU et al., 2000).

O subtipo HER2 positivo mostra expressão aumentada de *HER2*, ausência ou mínima expressão de ER e PR e está associado a outros marcadores de mau prognóstico (SORLIE et al., 2001). Este subtipo é responsivo a terapias como trastuzumab, e está associado com baixa sobrevida.

O subtipo-normal é caracterizado pela expressão de genes associados com o tecido adiposo e outros tipos de células do estroma. São classificados como triplo negativos porque não expressam nenhum dos três receptores, mas são negativos para os outros marcadores basais. Este subtipo é controverso porque para alguns autores representa a contaminação com células normais e não um subtipo molecular (VUONG et al., 2014). Eles são pouco caracterizados e o prognóstico é intermédio entre o subtipo luminal e basal.

1.4 Mineração de dados

O perfil de expressão diferencial de genes tem sido uma das ferramentas mais utilizadas em investigação. Esta ferramenta tem revolucionado o mundo da ciência de tal forma que tem resultado no descobrimento de numerosas redes gênicas e vias regulatórias, em vários processos tumorais (ABBA et al., 2010).

Atualmente existem inumeráveis projetos tentando analisar o perfil de expressão de genes para entender o comportamento de diferentes tipos de câncer. No entanto, uma das grandes limitações da maioria desses estudos radica principalmente no tamanho da amostra necessário para a obtenção de resultados significativos (MODELSKA; QUATTRONE; RE, 2015).

Existem na atualidade milhões de dados experimentais de estudos gerados por tecnologias de alto desempenho publicados com informação de centos de genes em diferentes tipos de tumores. A integração e disponibilização desses dados tornou-se o foco de muitos pesquisadores, não só pelo aumento do poder estatístico que confere um número maior de amostras senão também porque permite a captura das possíveis variabilidades técnicas o que aumenta o nível de confiabilidade dos resultados (CLARKE et al., 2013).

Apesar das diferentes técnicas, os dados genômicos e epigenômicos obtidos a partir de diferentes populações, plataformas e laboratórios, contêm informações consistentes e permitem investigar um único gene ou um conjunto de genes em um dado tumor levando à geração de novos conhecimentos e formulação de hipóteses (JIANG; LIU, 2015).

Hoje em dia existem numerosos bancos de dados com ferramentas de bioinformática online disponíveis de forma gratuita para o uso da comunidade científica. Esses bancos permitem a mineração de dados individuais e integrados em diferentes tipos de tumores, para responder perguntas clínicas que exigiriam pesquisas de alto custo usando métodos tradicionais (JURCA et al., 2016).

TCGA é um dos principais e mais conhecidos na comunidade científica (DENG et al., 2016). Esta plataforma disponibiliza uma vasta quantidade de dados de múltiplos estudos em diferentes tipos de câncer para extração, análise e visualização

de dados (CHENG; DUMMER; LEVESQUE, 2015). TCGA inclui dados de Alteração de número de cópia (CNA), mRNA, miRNA, mutação, metilação, proteína e dados clínicos relacionados a sobrevida. Ferramentas web para análise de dados de TCGA incluem cBioportal, PROGeneV2, UCSC Cancer browser.

TCGA contém dados desde o nível 1 até o nível 3 dependendo do grau de pré-processamento ou modificação computacional aplicada aos dados.

Tabela 1. Bancos de dados para mineração de dados em câncer

Banco de dados	Link	Tipo de dado	Tipo de informação
TCGA	http://cancergenome.nih.gov/dataportal	Nível I-III	Aberto e controlado
GEO	http://ncbi.nlm.nih.gov/geo	Nível I-III	Aberto
ONCOMINE	http://www.oncomine.org	Nível IV	Aberto e controlado
COSMIC	http://www.sanger.ac.uk/genetics/CGP/cosmic	Nível III-IV	Aberto
MirTarBase	http://mirtarbase.mbc.nctu.edu.tw	Nível I-III	Aberto
METHHC	http://methhc.mbc.nctu.edu.t	Nível I-III	Aberto
Kmplotter	www.kmplot.com	Nível I-III	Aberto
GOBO	http://co.bmc.lu.se/gobo/	Nível IV	Aberto

Os dados crus ou de nível I são dados sem nenhum tipo de processamento. Os dados processados ou de nível II, são aqueles cujas medidas de intensidade de sondas de um mesmo locus foram combinadas para permitir a sua comparação. O processamento também se refere aos processos de remoção de qualquer variação técnica que possa interferir com o valor de intensidade. Os dados de nível III são dados com resultado biológico como por exemplo perfis de expressão genica ou perfis de número de cópia do genoma inteiro. Os dados de nível IV ou resumidos, se caracterizam por serem dados interpretados a partir de um conjunto de amostras onde métodos estatísticos foram aplicados. Dados do nível quatro são geralmente encontrados em bancos como COSMIC, UCSC e cBioportal (CHIN et al., 2011).

Gene expression omnibus (GEO), outra ferramenta para análise e visualização de perfis de expressão gênica e epigenética, nasceu como uma iniciativa da comunidade científica para arquivar o grande volume de dados de microarray produzidos a partir de publicações de estudos (TANYA, 2013). GEO arquiva aproximadamente 40,000 estudos sendo que 90% correspondem a estudos de expressão gênica em diversos processos biológicos incluindo o câncer e o restante corresponde a estudos genômicos funcionais e epigenômicos como metilação do genoma, estrutura da cromatina, CNVs e interações genoma proteína (TANYA, 2013).

A interfase do GEO oferece uma completa variedade de ferramentas que inclui pacotes estatísticos do projeto bioconductor para comparação de conjuntos de dados, geração de lista de genes diferencialmente expressos, download de arquivos, análise de perfis de expressão gênica etc.

GOBO (*Gene expression-based outcome for breast cancer online*); é uma plataforma simples e acessível que inclui uma coleção de 1881 casos provenientes de onze estudos para validação de genes ou conjuntos de genes em câncer de mama (RINGNER et al., 2011). A interfase contém uma ferramenta de análise de expressão genica em diferentes categorias como status do receptor de estrógeno, status nodal, grau histológico, idade e tamanho do tumor. Adicionalmente conta com um conjunto de dados para análise de expressão em linhagem celular.

ONCOMINE, uma base de dados de nível IV é uma excelente opção para determinar a expressão de genes ou um conjunto de genes em conjuntos de dados de microarray em câncer. ONCOMINE contém 65 conjuntos de dados de expressão que correspondem a 48 milhões de dados de expressão genica em 4700 experimentos (RHODES et al., 2004).

Recentemente, um número considerável de ferramentas web tem sido desenvolvidas para integrar e correlacionar dados de metilação de DNA, microRNA e CNA com expressão gênica e sobrevida em pacientes com câncer de mama. BreastMark, Kmplotter, Survexpress, prognoscan, são alguns exemplos. O potencial dos microRNAs como marcadores em câncer também tem promovido a criação de bancos de dados para predição de microRNAs alvos de genes. TargetScan, Mirtarbase, PicTar e Miranda são os mais conhecidos (MONTERISI et al., 2014).

Diversos estudos de larga escala têm aportado em grande maneira à compreensão do câncer de mama. Muitos desses estudos encontram-se depositados em bancos de dados públicos para o livre uso da comunidade científica bem seja para reanálise desses dados ou para o teste de novas hipóteses.

Estudos como o de Curtis 2000, por exemplo, tem sido analisado e reanalisado por diferentes grupos o que tem facilitado a identificação de novos subtipos intrínsecos de câncer de mama, a predição de sobrevida e a resposta a drogas. Curtis integrou a análise do genoma e transcriptoma de 2000 tumores de mama identificando 10 subgrupos que representam subtipos moleculares de mama (CURTIS et al., 2012). Usando a mesma coorte de dados Joe e Hojung Nam (2016) correlacionaram 4 módulos de expressão com prognóstico em câncer de mama e o grupo de Fertig (2016) caracterizaram a expressão de 5 moléculas associadas com metástases nos 5 subtipos moleculares de câncer de mama (FERTIG et al., 2015; JOE; NAM, 2016).

Usando dados de outro estudo depositado no TCGA com 1000 casos de câncer de mama, o RNA não codificante de cadeia longa LINC00657 (lncRNAs), foi identificado como um biomarcador em câncer de mama (LIU et al., 2016).

FERNANDEZ et al., 2011 analisaram o perfil de metilação de 1628 amostras normais, tumorais e não tumorais de diferentes tipos de tecidos dentre eles o da mama. Nesse trabalho encontraram alterações no padrão de metilação das células cancerosas e em outras doenças humanas analisadas. Os dados desse trabalho estão depositados na plataforma GEO para o livre uso para estudos adicionais.

O valor prognóstico da expressão do gene DCYTB em câncer de mama, por exemplo, foi avaliado usando três bancos de dados; GOBO e GEO para a análise de expressão de mRNA, e KMplotter para a análise de sobrevida (LEMLER et al., 2017). Nesse trabalho o gene DCYTB foi identificado como um potencial supressor de tumor.

Conjuntos de dados de diversas plataformas têm sido utilizados adicionalmente para a identificação de novos marcadores de resistência a drogas em câncer de mama (ZHUO et al., 2014), marcadores de risco metastático em câncer de mama (ZHANG et al., 2013) identificar vias de regulação importantes em câncer de mama entre outros.

Os dados de escala genômica tem tornado a biologia na ciência dos dados. Hoje a área de bioinformática tem crescido de maneira surpreendente em resposta à demanda da grande quantidade de dados que há alguns anos parecia restringida somente a dados complexos de finanças, meteorologia e outros campos (NEVINS; POTTI, 2007).

É evidente que o entendimento da heterogeneidade do câncer de mama tem melhorado substancialmente, no entanto, ainda há uma grande necessidade de identificar marcadores que permitam melhor caracterizar esta doença e aumentar o prognóstico e sobrevida dos pacientes (WEIGEL; DOWSETT, 2010).

Numerosos genes têm-se mostrado como bons candidatos sendo alguns amplamente caracterizados, no entanto, existe uma boa quantidade que tem sido pouco explorado. Este é o caso dos genes da família PHLDA. Estes genes têm sido identificados repetidamente em vários estudos de perfil transcricional em câncer de mama por diferentes grupos incluindo o nosso (NAGAI et al., 2003) e reconhecidos por seu papel em vários processos, principalmente na biologia do câncer devido ao seu potencial papel na apoptose.

1.5 Família PHLDA

A família PHLDA compreende um grupo de três genes; o *PHLDA1*, localizado no cromossomo 12q21.2, o gene *PHLDA2* localizado no cromossomo 11p15.5 e o gene *PHLDA3* localizado no cromossomo 1q31. Esses genes apresentam uma estrutura comum composta principalmente pelo domínio de homologia à plectrina que compreende uma região de aproximadamente 120 aminoácidos, encontrado em muitas proteínas com funções na sinalização celular (LEMMON; FERGUSON, 2000; LEMMON, 2007). Aproximadamente 15% destes domínios se ligam com alta afinidade e especificidade ao fosfatidil inositol localizado no interior da membrana plasmática onde medeiam o recrutamento de proteínas às membranas celulares, o que é de suma importância para a transdução de sinal (LEMMON; FERGUSON, 2000).

Os genes *PHLDA2* e *PHLDA3* apresentam uma estrutura composta majoritariamente na porção central pelo domínio plectrina. À diferencia destes, *PHLDA1* possui um domínio central plectrina interrompido por um trato de poliglutamina (QQ), e uma região carboxi terminal composta por extensões ricas em

prolina-glutamina (PQ) e prolina-histidina (PH) não encontradas nos outros dois membros da família (FRANK et al., 1999).

Figura 2. Estrutura da proteína dos três membros da família PHLDA. A estrutura da proteína de PHLDA2 e PHLDA3 está composta principalmente pelo domínio plecstrina, à diferença de PHLDA1 que contém o domínio interrompido por repetições de glutamina-prolina e histidina. PHL: PH-like domain; Q: domínio rico em glutamina; PQH: extensão C-terminal rica em prolina, glutamina e histidina. (Adaptado de Frank, 1999).

1.6 PHLDA1

O gene *PHLDA1* codifica uma proteína de 401 aminoácidos localizada principalmente no citoplasma e no núcleo (HOSSAIN et al., 2003; NEEF et al., 2002; SAKTHIANANDESWAREN et al., 2011)

PHLDA1 foi descrito inicialmente como um regulador importante da expressão de Fas na apoptose induzida por ativação em um hibridoma de células T *in vitro*, entretanto, estudos posteriores demonstraram que *PHLDA1* não é essencial para a regulação de Fas e apoptose *in vivo* (RHO et al., 2001).

Em cultura de células endoteliais vasculares a expressão de *PHLDA1* é induzida por choque térmico e agentes como ditiotreitol e tunicamicina causando mudanças na morfologia celular, diminuição da adesão e morte celular programada por perda de ancoragem (HOSSAIN et al., 2003; PARK et al., 1996).

Adicionalmente, a expressão de *PHLDA1* responde a estímulos estrogênicos, fator de crescimento insulínico (IGF-1), fator de crescimento fibroblástico e à ativação de RAS.

Alterações na expressão de *PHLDA1* têm sido descritas em diferentes tipos de tumores e linhagens celulares com funções que variam dependendo do tipo de células e do contexto. Em melanoma metastático, câncer de mama, linhagens celulares neuronais, células de carcinoma cervical, endotelial e células T; a expressão de *PHLDA1* tem sido associada com indução de apoptose (NEEF et al., 2002; PARK et al., 1996). Em células de carcinoma oral (MURATA et al., 2014) células de fibroblastos de rato (PARK et al., 2013) e em tumor de intestino (SAKTHIANANDESWAREN et al., 2011) pelo contrário, é associado com um efeito anti-apoptótico.

Em um estudo do nosso grupo, a diminuição de expressão de *PHLDA1* se associou com um fenótipo agressivo e mal prognóstico em pacientes com tumores primários de mama (NAGAI et al., 2007). Em outro trabalho, nosso grupo avaliou o papel dos estrógenos na regulação de *PHLDA1* em linhagem celular MCF7; os resultados mostraram que após exposição a 17 B-estradiol os níveis mRNA de *PHLDA1* aumentaram significativamente, indicando que *PHLDA1* pode ser regulado pela via do receptor de estrógeno (MARCHIORI; CASOLARI; NAGAI, 2008).

Em células normais de mama MCF-10A a supressão de expressão do gene *PHLDA1* induz um comportamento agressivo nessas células com alterações na morfologia e aumento na proliferação e migração celular (BONATTO et al., 2017).

Em células de câncer de mama MDA-MB-231, *PHLDA1* atua como sustrato da Aurora A, uma proteína quinase que se encontra superexpressa em vários tipos de câncer de mama. Nessas células *PHLDA1* é fosforilado e posteriormente degradado,

inibindo assim suas funções pro-apoptóticas (JOHNSON et al., 2011). *PHLDA1* regula processos de apoptose e autofagia em células T47D tratadas com rapamicina (MOAD et al., 2013). A proteína *PHLDA1* se encontra com baixa expressão em carcinoma de células basais, e diferencia esses tumores do tricoepitelioma desmoplástico que expressam *PHLDA1* (SELLHEYER; KRAHL, 2011).

Várias vias de sinalização foram propostas como reguladoras de *PHLDA1*. No trabalho de Oberst et al., (2008) a proteína *PHLDA1* foi identificada como alvo da via de sinalização de RAS onde pode atuar como supressor da sinalização de ERK em células epiteliais de mama. Por outro lado, Lyu et al., (2016) demonstrou que *PHLDA1* está regulado pela via JAK2/STAT3 sendo ERK1/2 um importante mediador entre JAK2 e STAT3. A hipótese é que a ativação de TLR2 estimula JAK2/STAT3 via ERK 1/2 e eventualmente aumenta a expressão de *PHLDA1*. No trabalho de Li et al., (2014), *PHLDA1* foi identificado como um gene alvo da via *ERBB2* em linhagens celulares de câncer de mama *HER2* positivo; e foi apontado como um possível alvo de novas terapias contra este tipo de câncer (LI et al., 2014a; LYU et al., 2016; OBERST et al., 2008).

Deficiência de *PHLDA1* em células fibroblásticas embrionárias de rato resultou em produção de elementos de resposta a estresse oxidativo e ativação da caspase 3, induzindo a morte celular por apoptose (PARK et al., 2013). *PHLDA1* é crucial na formação e crescimento de mamosferas de células de câncer de mama receptor de estrógeno positivo. Nessas células *PHLDA1* é regulado positivamente por ER e NFkB que por sua vez contribuem com fenótipo metastático e mau prognóstico (KASTRATI; CANESTRARI; FRASOR, 2014). O papel do gene *PHLDA1* ainda não está bem definido, mas alguns estudos mostram que alterações na expressão deste gene podem contribuir com o processo tumorigênico fazendo dele um potencial

marcador prognóstico. Entretanto, pouco se sabe sobre os mecanismos que estão associados à redução de expressão deste gene.

1.7 PHLDA2

O gene *PHLDA2* (*Pleckstrin homology-like domain family A member 2*) também denominado *IPL* e *TSSC3* (LEE; FEINBERG, 1998), está localizado na região cromossômica 11p15.5, que é considerada uma importante região de localização de genes supressores de tumores (QIAN et al., 1997).

Este gene à diferença dos seus homólogos realiza expressão monoalélica, a partir do alelo materno por meio de um processo denominado *imprinting* genômico (LEE; FEINBERG, 1998). *Imprinting* é um mecanismo epigenético por meio do qual apenas um dos dois alelos herdados, o paterno ou o materno, expressa um gene específico. Este fenômeno ocorre em menos de 1% dos genes no organismo e é essencial para o crescimento e metabolismo (JOVANOVIC et al., 2010). A marcação do alelo é feita através da metilação do DNA que consiste na adição de um grupo metil no carbono 5' do nucleotídeo citosina, pela ação de uma enzima denominada DNA metil transferase (PIEDRAHITA, 2011).

Alteração no padrão de metilação do DNA resulta em silenciamento ou superexpressão de genes. A hipermetilação está associada à compactação da cromatina e posterior silenciamento de genes que geralmente desempenham funções supressoras de tumores como controle do ciclo celular, reparação do DNA, apoptose, adesão e migração celular (BARBARESCO et al., 2012), em contrapartida hipometilação está relacionada com superexpressão de proto-oncogenes que atuam na invasão celular e metástase.

O gene *PHLDA2* codifica uma proteína de 262 aminoácidos cujas funções estão associadas com controle da morte celular por apoptose e desenvolvimento fetal e placentário (APOSTOLIDOU et al., 2007). Este gene foi descrito como o primeiro gene pro-apoptótico regulado por “*imprinting*” genômico (LEE; FEINBERG, 1998). O gene *PHLDA2* foi associado a “*imprinting*” em placenta e fígado denominando-se *ipl* do inglês (*imprinted in placenta and liver*) (QIAN et al., 1997) não obstante, LEE; FEINBERG, (1998); encontraram *imprinting* em quase todos os tecidos fetais testados exceto no coração e nos testículos.

Vários trabalhos têm mostrado que alterações na expressão deste gene podem contribuir com o desenvolvimento de diversos tipos de tumores. A perda de expressão de *PHLDA2* é observada durante a gravidez em uma doença denominada mola hidatiforme (SCHWIENBACHER et al., 2000) e em neuroblastoma de Wilm (SAXENA et al., 2003). O aumento de expressão é frequentemente encontrado em osteosarcoma e em restrição de crescimento intrauterino em humanos (IUGR) (APOSTOLIDOU et al., 2007; BASSERI et al., 2013; LI et al., 2014b).

DAI e colaboradores (2012), associaram a expressão de *PHLDA2* com funções pró-apoptóticas em tumores primários de osteosarcoma e em linhagem celular de osteosarcoma SaOS2 através da inibição de crescimento e indução da apoptose mitocondrial *in vivo* e *in vitro*. Em outro estudo com células SaOS2 tratadas com um agente inibidor da DNA metil transferase conhecido como 5-Aza-CdR, a expressão do gene *PHLDA2* foi reativada, indicando hipermetilação da região promotora (LI et al., 2014b).

Schwiebacher et al., 2000 mostrou em um estudo que *PHLDA2* encontrava-se com baixa expressão em tumor de Wilm e que o mecanismo de regulação do gene nesse tumor era independente de metilação. Em outro estudo Li et al., 2008;

realizaram uma análise do padrão de metilação de células osteoblásticas transformadas com dois agentes carcinogênicos. Os resultados mostraram que a região promotora dos genes *H19* e *PHLDA2* se encontravam hipometiladas, e como consequência, esses genes apresentavam aumento de expressão, sugerindo que alterações no padrão de metilação estão envolvidos na progressão tumoral. Nocaute de Dicer, uma enzima descrita na regulação da cromatina resultou na superexpressão de *PHLDA2* em células HEK293 (TANG et al., 2007).

Em linhagens celulares de câncer de mama HER2 positivo tratadas com lapatinib, a expressão de mRNA e proteína de *PHLDA2* é baixa sugerindo a regulação do gene pela via ERBB2/EGFR (WANG et al., 2015). Em células de pulmão, superexpressão de *PHLDA2* inibe a ativação de AKT. Em câncer gástrico o silenciamento de *PHLDA2* leva à diminuição de expressão de P53 e aumento dos níveis de AKT (KOH et al., 2017). Em linhagem celular de trofoblasto JEG-3 a expressão de *PHLDA2* tem um efeito inibidor de crescimento, invasão e migração das células (JIN et al., 2016).

Perda de heterozigose (LOH) na região cromossômica 11p15.5 é um evento frequente em tumores de mama. Estudo de Negrini e colaboradores (1994) demonstraram que a transferência do cromossomo 11 normal para células de câncer de mama MCF7, reduz o potencial tumorigênico dessas células. Esses dados sugerem a presença de genes supressores de tumor no cromossomo 11 associados ao processo de tumorigênese da mama. Entretanto, poucas são as informações sobre a expressão ou o padrão de metilação do gene *PHLDA2* em câncer de mama.

1.8 PHLDA3

O gene *PHLDA3* também denominado *Tih1* possui uma estrutura similar aos seus homólogos. Codifica uma proteína de 127 aminoácidos composta na maior parte pelo domínio PH. A proteína se localiza na membrana plasmática onde desempenha funções pro-apoptóticas (KAWASE et al., 2009b).

O transcrito de *PHLDA3* é ubíquo em tecidos fetais, com exceção do fígado fetal. Em tecido adulto, o mRNA de *PHLDA3* é expresso fortemente em músculo esquelético e pulmão e em níveis moderados em outros tecidos do corpo (FRANK et al., 1999; KERLEY-HAMILTON et al., 2005); identificaram *PHLDA3* como um gene indutor de *P53*, no entanto, KAWASE et al., 2009; demonstraram que *PHLDA3* é um alvo direto do gene *P53*.

PHLDA3 é induzido em resposta à ativação de *P53* (KAWASE, PAPPAS), tratamento com cisplatina (KERLEY-HAMILTON et al., 2005), exposição à radiação gama e outros agentes de dano ao DNA.

PHLDA3 tem sido descrito como um regulador negativo da via AKT em vários tumores. Em tumor pulmonar e em tumores neuroendócrinos, *PHLDA3* compete com o domínio PH da proteína AKT para a união de lipídeos de membrana, levando à ativação e indução da apoptose dependente de *P53*. Em linhagens celulares de osteosarcoma, de mama, de pulmão, de próstata entre outras, *PHLDA3* pode induzir morte celular independentemente do status de *P53* (KAWASE et al., 2009b).

Perda de heterozigose e hipermetilação resulta na supressão da transcrição do gene em tumores neuroendócrinos pancreáticos e se correlaciona com baixa sobrevida (OHKI et al., 2013). Alterações no número de cópia também são encontradas em

tumores endócrinos de pulmão e levam à diminuição de expressão do gene (TAKIKAWA; OHKI, 2017).

Baixa expressão de PHLDA3 está associada com mal prognóstico em tumores de células esofágicas (MUROI et al., 2015). Superexpressão de p53 induz aumento de expressão de PHLDA3 em linhagens celulares de mama MDA-MB-468 (KAWASE et al., 2009b).

São escassos os estudos do perfil de expressão e o valor prognóstico de PHLDA3 em câncer de mama. Contudo, os mecanismos associados a este gene no desenvolvimento tumoral ainda são desconhecidos e merecem maiores investigações.

2. OBJETIVO GERAL

Identificar e utilizar bases de dados disponíveis para coleta de informações sobre a expressão e o potencial valor prognóstico dos genes da família PHLDA (1,2,3) em câncer de mama.

2.1 Objetivos Específicos

- Compilar dados para análise do perfil de expressão dos genes *PHLDA1*, *PHLDA2*, e *PHLDA3* em tumores de mama e linhagens de células derivadas de câncer de mama e avaliar o potencial valor prognóstico da expressão desses genes em câncer de mama.
- Compilar dados para análise do perfil de metilação dos genes *PHLDA1*, *PHLDA2*, e *PHLDA3* a partir de bancos de dados selecionados.
- Identificar e utilizar bancos de dados públicos para análise de interação molecular de *PHLDA1*, *PHLDA2* e *PHLDA3*.

3. MATERIAL E MÉTODOS

3.1 ANÁLISE DE EXPRESSÃO GENICA

Os dados de expressão gênica da família PHLDA em relação a subtipos moleculares, status de receptor de estrógeno, grau do tumor e subtipo histológico foram extraídos usando três diferentes bancos de dados: TCGA (The Cancer Genome Atlas), GEO (Gene Expression Omnibus) e GOBO (Gene Expression-Based Outcome for Breast Cancer).

O banco TCGA foi acessado pela plataforma cBioPortal (www.cbioportal.org/), uma plataforma web que facilita a exploração e análise de dados gerados de estudos genômicos de larga escala. Esta plataforma contém diversos estudos em câncer de mama com dados processados, normalizados e interpretados (nível 3).

Para obtenção dos dados duas ferramentas foram usadas: o pacote estatístico CGDS-R da plataforma R que contém uma série de funções para a busca, extração e análise de diferentes tipos de dados a partir do servidor de dados genômicos em câncer CGDS; e a ferramenta de busca disponível no site do cBioPortal acessando ao link (<https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>).

Os dados de expressão de mRNA da família PHLDA por subtipos moleculares, status do receptor de estrógeno, grau do tumor e subtipo histológico foram extraídos do conjunto de dados do estudo METABRIC que contém dados de expressão e mutação de aproximadamente 2000 pacientes com câncer de mama.

O banco de dados GEO foi acessado através do link (<https://www.ncbi.nlm.nih.gov>). Para a busca de dados de expressão de mRNA da família

PHLDA em câncer de mama, foram usadas palavras chaves como: “*Breast cancer molecular subtypes*”, “*Breast cancer estrogen receptor*”, na ferramenta de busca “*GEO dataset*” de GEO.

Os critérios de busca definidos para a seleção dos dados foram: n maior do que 50, dados com distribuição normal, e dados com expressão dos três membros da família PHLDA. Somente dados que cumprissem com os critérios definidos foram escolhidos. Para a extração dos valores de expressão gênica foi usada a plataforma de análise de dados do GEO inserindo o número de acesso do conjunto de dados na opção “*profile graph*” e posteriormente o ID dos genes da família PHLDA. Os dados foram posteriormente analisados e visualizados usando os pacotes da plataforma estatística R.

O perfil de expressão gênica em linhagem de células de câncer de mama foi realizado a partir de um conjunto de dados selecionado a partir da plataforma GEO. Os critérios de busca foram: dados com pelo menos 30 amostras e que representassem os subtipos Luminais, Her2 e Basal. As palavras usadas para a busca foram “*Breast cancer cell line*” e o filtro de tipo de estudo “*expression profiling by array*” foi aplicado.

A base de dados GOBO (*Gene expression-based outcome for breast cancer online*; <http://co.bmc.lu.se/gobo/>) também foi escolhida para a análise de expressão gênica. GOBO foi desenvolvida por Ringnér e colaboradores (2011) com o objetivo de fornecer uma ferramenta online para validação do valor prognóstico de genes individuais, conjuntos de genes ou preditores simples em um conjunto de dados de câncer da mama. Esta plataforma disponibiliza dados sumarizados (nível 4) a partir de uma coleção de 1881 casos provenientes de onze estudos utilizando cDNA microarray U133A Affymetrix para avaliação do perfil de expressão de genes em câncer de mama.

Este banco de dados utiliza dois classificadores para a subtipagem dos tumores de mama em Luminal A, luminal B, Basal-like, HER2/⁺ER⁻ e Normal-like. A lista de genes baseada nos dados de Hu et al., (2006) e no PAM50, os quais foram incluídos nas análises para comparação. Além disso, dados de status do receptor de estrógeno e do grau do tumor também foram incluídos nesta análise. A extração dos dados analisados foi realizada inserindo na ferramenta de busca do site do GOBO o nome dos genes e posteriormente os arquivos foram descarregados e copiados.

3.2 ANÁLISE DE METILAÇÃO

A análise de metilação foi realizada usando as plataformas TCGA, GEO e METHHC. Na plataforma TCGA, a análise de metilação foi realizada usando o conjunto de dados do estudo TCGA provisional que contém 963 amostras de câncer de mama. Usando dados deste estudo foram investigadas as associações entre o padrão de metilação da família PHLDA e o status de receptor de estrógeno, o estágio do tumor e o subtipo molecular em câncer de mama. Estes dados foram extraídos acessando à plataforma de busca do site cBioPortal (<https://tcga-data.nci.nih.gov/tcga/dataAccessMatrix.htm>).

Para a análise de metilação em tecido tumoral comparado com amostra de tecido normal foi utilizada a plataforma METHHC. Este banco coleta dados do portal TCGA com informações do perfil de metilação da região promotora, das ilhas CpG, das regiões intergênicas entre outras, de genes em diferentes tipos de câncer para correlacioná-los com dados de expressão gênica. Para extração dos valores de metilação (Beta value) em amostra normal e em tumor, o ID dos genes foi inserido na

plataforma de busca e posteriormente os dados foram analisados usando pacotes estatísticos da plataforma R. Apenas os dados pareados foram incluídos.

O perfil de metilação no banco de dados GEO foi obtido inserindo na plataforma de busca palavras chaves como “*Breast cancer methylation*”, “*Breast cancer*” e posteriormente filtradas no tipo de estudo por “*Methylation profiling by array*”. O conjunto de dados contendo como mínimo 100 amostras, com valores para os três membros da família e com distribuição normal, foi escolhido. Para a extração dos dados de metilação foi usada a plataforma de análise de dados do GEO inserindo o número de acesso do conjunto de dados na opção “*profile graph*” e posteriormente o ID dos genes da família PHLDA. Os dados foram extraídos e posteriormente visualizados e analisados usando os pacotes da plataforma estatística R.

3.3 CURVAS DE SOBREVIDA

O valor prognóstico da expressão de mRNA da família PHLDA em câncer de mama foi avaliado usando o Kaplan-Meier plotter (www.kmplot.com) que é uma base de dados online que contém dados de expressão gênica e de sobrevida total e livre de recorrência em câncer extraídas a partir de GEO (Affymetrix microarrays only), EGA e TCGA. Os dados de busca foram inseridos no link <http://kmplot.com/analysis/index.php?p=service&cancer=breast> para a obtenção dos gráficos e dos valores de log rank p e de Hazard ratio com intervalo de 95% de confiança. A análise foi feita com o melhor *cucutoff* incluindo indivíduos censurados. A análise foi restrita aos subtipos intrínsecos Basal-like, luminal A, luminal B e HER2 positivo.

3.4 MicroRNA

A busca de microRNAs alvos da família PHLDA foi realizada usando o banco de dados mirTarBase (mirtarbase.mbc.nctu.edu.tw). Esta base de dados permite a identificação de microRNAs alvos de genes validados experimentalmente. Além disso, este banco permite a visualização de redes de interação de microRNAs e genes alvo, e a visualização de perfis de expressão de microRNAs em diferentes conjuntos de dados, entre outras funções. A identificação de genes alvos da família PHLDA1 foi realizada inserindo o nome dos genes na plataforma de busca do banco. Os microRNAs validados pelo maior número de métodos ou validados por um método de validação com “evidencia forte” foi escolhido para cada gene da família PHLDA.

O perfil de expressão de microRNAs em tumor comparado com amostra normal e por subtipo molecular, foi realizado usando um conjunto de dados do banco GEO. As palavras de busca inseridas foram: “*Breast cancer microRNA*”, “*Breast cancer microRNA molecular subtypes*”. Seguidamente o resultado da busca foi filtrado selecionando na opção tipo de estudo: “*non-coding RNA profiling by array*”. Conjuntos de dados com menos de 100 amostras foram descartados.

Para a extração dos valores de expressão gênica foi usada a plataforma de análise de dados do GEO, inserindo o número de acesso do conjunto de dados na opção “*profile graph*” e posteriormente o ID dos microRNAs em análise. Os dados foram posteriormente visualizados e analisados usando os pacotes da plataforma estatística R.

O valor prognóstico de microRNAs alvos da família PHLDA foi determinado usando a plataforma miRpower (www.kmplot.com/mirpower/). MiRpower é uma ferramenta web para validação de microRNAs associados a sobrevida utilizando conjuntos dados de expressão gênica de METABRIC, GEO e TCGA. Os dados de

busca foram inseridos na plataforma KM Plotter, para a obtenção dos gráficos e dos valores de log rank p e de Hazard ratio com intervalo de 95% de confiança. O conjunto de dados escolhido para a análise foi o de METABRIC e os parâmetros definidos para a busca incluíram indivíduos censurados, tempo de seguimento de 20 meses, e análise para todos os subtipos moleculares em conjunto.

REDES GÊNICAS

As possíveis vias de interação dos genes da família PHLDA em câncer de mama foram analisadas usando Netdecoder (da Rocha et al.,2016), uma plataforma computacional para construção e análise de redes biológicas. Netdecoder é baseado em um algoritmo de fluxo de rede e formulado como um problema de otimização de fluxo de custo mínimo para identificar caminhos de interação molecular e reguladores chave dentro de sub-redes específicas de doenças. Genes fonte “source”, alvos “targets” e fatores de transcrição podem ser identificados dentre as vias representando um fluxo de redes biológicas. Genes com impacto em doença e genes chaves podem também ser identificados dentre as redes. Para analisar redes gênicas foram escolhidos três conjuntos de dados na plataforma GEO com dados de expressão por subtipos moleculares Luminal, HER2 e Basal. Os dados crus foram extraídos e posteriormente normalizados. Para identificação da lista de genes diferencialmente expressos em doença comparados com o controle foi utilizada a ferramenta web GEO2R. Considerando que nenhum dos genes da família PHLDA apareceu diferencialmente expresso nas comparações, eles foram inseridos manualmente. As redes foram customizadas e visualizadas utilizando a ferramenta Cytoscape.

ANÁLISE ESTATÍSTICA

A significância estatística nas comparações entre dois grupos foi determinada usando o test t com um valor de $p < 0.05$ aceito como estatisticamente significativo. A significância estatística entre os grupos foi realizada usando o test ANOVA e o posteste Tukey. As análises foram realizadas usando os pacotes da plataforma estatística R.

Tabela 2. Conjuntos de dados usados nas análises

Número de acesso	Referência	Formato dos dados	n	Tipo de plataforma (probe number)
GSE20713	Dedeurwaerder <i>et al</i> , 2011	Processados	250	Illumina HumanMethylation27 BeadChip
GSE37751	Terunuma <i>et al</i> , 2014	Processados	108	Affymetrix Human Gene 1.0 ST Array
GSE22133	Jönsson <i>et al</i> , 2012; Holm <i>et al</i> , 2012	Processados	359	SWEGENE H_v2.1.1 55K
GSE81002	aure <i>et al</i> , 2017	Processados	375	Agilent-Human miRNA Microarray v14 Rev.2
GSE19783	Enerly <i>et al</i> , 2011	Processados	101	Agilent- Human miRNA Microarray 2.0 G4470B
GSE73002	Shimomura <i>et al</i> , 2016	Processados	4,113	3D-Gene Human miRNA V20_1.0.0
METABRIC	Curtis <i>et al</i> , 2012; Pereira <i>et al</i> , 2016	Processados	2,509	Affymetrix SNP 6.0 and Illumina HT-12 v3
TCGA provisional	-	Processados	1,105	Illumina HumanMethylation450 (HM450)

4. RESULTADOS

4.1 Análise comparativa de expressão, metilação e proteína da família PHLDA em câncer de mama

Utilizando bancos de dados públicos foram levantados dados de expressão de mRNA e de metilação da família PHLDA para análise de coexpressão dos genes *PHLDA1*, *PHLDA2* e *PHLDA3* em câncer de mama. A análise de correlação da família PHLDA usando dados de expressão de mRNA do estudo METABRIC e do TCGA provisional mostrou que há uma relação positiva fraca tanto na expressão quanto na metilação dos genes da família PHLDA em câncer de mama; e uma relação inversa de expressão entre o gene *PHLDA1* e *PHLDA3* ($r = -0.0444$) (Figura 3).

A partir de dados do estudo METABRIC e TCGA provisional foi investigado o nível de expressão de mRNA, proteína e de metilação dos genes *PHLDA1*, *PHLDA2* e *PHLDA3* em câncer de mama. Os resultados mostram que o perfil de expressão de mRNA e proteína entre *PHLDA1* e *PHLDA2* é bastante similar ($p = 0.5446$) (Figura 4A e B), enquanto que *PHLDA3* tem uma tendência a ter maiores níveis de expressão quando comparado com seus homólogos ($p < 0,0001$) (Figura 4A e C). A avaliação do perfil de metilação pelo contrário, mostrou que há diferenças significativas no nível de metilação entre os três genes da família PHLDA; o gene *PHLDA1* é o mais metilado dos três enquanto que o *PHLDA2* é o menos metilado ($p < 0,0001$) (Figura 4C).

Figura 3. Análise de correlação da expressão de mRNA e status de metilação entre os membros da família PHLDA em câncer de mama. (A) Expressão de *PHLDA1* vs *PHLDA2* (B) Expressão de *PHLDA2* vs *PHLDA3* (C) Expressão de *PHLDA1* vs *PHLDA3* (D) Metilação de *PHLDA1* vs *PHLDA2* (E) Metilação de *PHLDA2* vs *PHLDA3* (F) Metilação de *PHLDA3* vs *PHLDA1*. Os valores de correlação r estão indicados.

Figura 4. Análise comparativa de expressão de mRNA, proteína e status de metilação da família PHLDA no banco de dados TCGA (A) Expressão de mRNA (B) Expressão de proteína (C) status de metilação. Os valores P-ANOVA estão indicados.

4.2 Análise do perfil de expressão e de metilação do gene *PHLDA1* em câncer de mama

A análise de expressão de mRNA e metilação de *PHLDA1* em amostra normal e de tumor de mama foi realizada usando dois bancos de dados diferentes; o TCGA e o GEO. Nos dois bancos, encontramos que os transcritos de *PHLDA1* encontram-se com baixa expressão nas amostras de tumor de mama quando comparado com amostras de tecido mamário normal (TCGA, $p=4.594 \times 10^{-13}$); (GEO, $p=0,0283$) (Figura 5A e 5B). A análise do perfil de metilação da figura 5C foi obtida a partir de dados do TCGA depositados na plataforma METHHC e as análises da figura 5B e D foram realizadas usando duas coortes de pacientes com câncer de mama e pacientes com amostra normal a partir dos trabalhos depositados na plataforma GEO com o número de acesso: GSE37751 e GSE20713 respectivamente. Como observado na figura 5D e 5C, os

resultados obtidos a partir das duas bases de dados são diferentes. Os dados gerados a partir da plataforma TCGA não mostram diferença significativa no padrão de metilação do gene PHLDA1 entre as amostras de tecido tumoral e normal de mama (Figura 5C), entretanto, os dados obtidos na base de dados GEO, mostram uma frequência maior de metilação do gene PHLDA1 em amostras de tumor de mama quando comparado com amostra de tecido normal (Figura 5D, $p=0.00024$).

Figura 5. Expressão de mRNA e perfil de metilação do gene PHLDA1 em amostra normal comparada com tumor de mama. Perfil de expressão de PHLDA1 em amostra normal comparada com amostra de tumor usando dados do banco TCGA (A) Expressão de mRNA de PHDA1 usando dados do banco GEO (B) perfil de metilação do gene PHLDA1 usando dados do banco TCGA (C) perfil de metilação do gene PHLDA1 usando o banco de dados GEO (D).

4.2.1 Análise do perfil de expressão do gene PHLDA1 por subtipo molecular de câncer de mama.

O perfil de expressão da família PHLDA por subtipos moleculares de câncer de mama foi investigado usando três bancos de dados: GOBO, TCGA e GEO. Os dados de expressão apresentados na Figura 6A foram gerados com base na classificação dos subtipos intrínsecos de câncer de mama que utiliza uma lista de 306 genes (Intrinsic/UNC), definida no estudo de Hu et al.,(2006), (GOBO database); enquanto que os dados das figuras 6B, C e D foram gerados com base na classificação dos subtipos intrínsecos de câncer de mama com base na plataforma PAM50 (Prosigna® (PAM50) Gene Signature Assay), que mede os níveis de expressão de 50 genes em amostras de câncer de mama e permite a classificação dos tumores em um dos quatro subtipos intrínsecos, como luminal A, luminal B, HER2-enriquecido, e Basal-like (Staaf et al. 2010). A análise mostrou que o perfil de expressão do gene PHLDA1 por subtipo molecular é similar entre todos os bancos avaliados observando-se em todos uma tendência dos subtipos Basais apresentarem maior expressão dos transcritos de PHLDA1 do que os demais subtipos de câncer de mama e os tumores do subtipo her2+ apresentarem menor expressão. Subtipos Her2+ vs basal e normal like, ($p < 0,00001$), her2 vs lumA, ($p < 0,01$) her2 vs luminal B ($p = 0,006$) (Figura 6D).

Figura 6. Expressão de mRNA de PHLDA1 em subtipos moleculares de câncer de mama. Box plots representando o perfil de expressão do gene *PHLDA1* em tumores de mama estratificados de acordo com subtipos moleculares em diferentes bancos de dados. (A) dados de expressão em amostras estratificadas de acordo com a lista de genes do Hu et al.,(2006) usando o banco de dados GOBO (B) dados de expressão em amostras estratificadas pela plataforma PAM50 usando o banco de dados GOBO (C) dados de expressão em amostras estratificadas pela plataforma PAM50 usando o banco de dados da GEO (D) dados de expressão em amostras estratificadas pela plataforma PAM50 usando o banco de dados TCGA. Os valores P-ANOVA estão indicados.

4.2.2 Análise do perfil de expressão do gene PHLDA1 por status do receptor de estrógeno em câncer de mama

O perfil de expressão do gene PHLDA1 de acordo com o status do receptor de estrógeno foi investigado nos bancos TCGA, GEO e GOBO (Figura 7A, B e C). Consistentemente, os três bancos mostraram que tumores com presença de receptor de estrógeno expressam níveis significativamente menores de PHLDA1 comparados com os

tumores que não expressam o receptor de estrogênio ($p = 1.786e-14$), ($p = 0.0321$), ($p < 0.00001$) (Figura 7A, B e C). Concordando com este resultado, os tumores de mama ER positivos apresentam altos níveis de metilação do gene *PHLDA1* quando comparados com os ER negativos ($p = 3.762e-15$) (Figura 7D).

Figura 7. Expressão de mRNA e perfil de metilação de *PHLDA1* em tumores de mama estratificados pelo receptor de estrogênio. (A) Expressão de mRNA de *PHLDA1* estratificados de acordo ao status de receptor de estrogênio usando o banco TCGA (B) expressão de mRNA de *PHLDA1* estratificados de acordo ao status de receptor de estrogênio usando o banco GEO (C) Expressão de *PHLDA2* estratificados de acordo ao status de receptor de estrogênio usando o banco GOBO (D) Nível de metilação de *PHLDA2* em tumores estratificados de acordo ao status de receptor de estrogênio usando o banco de dados TCGA.

4.2.3 Análise do perfil de expressão do gene *PHLDA1* em linhagens celulares de câncer de mama

O perfil de expressão de mRNA do gene *PHLDA1* foi investigado em um conjunto de dados com expressão em linhagens celulares de mama. Comparando com a linhagem celular de mama normal MCF-10A, as linhagens celulares hormônio responsivas triplo negativas e HER2 positivas têm uma tendência a ter baixos níveis de expressão de *PHLDA1* com exceção de algumas linhagens triplo negativas como MDA231, HTERT e SUM102.

Figura 8. Perfil de expressão de mRNA de *PHLDA1* em linhagens celulares de câncer de mama. Histogramas representando o perfil de expressão do gene *PHLDA1* em 46 linhagens celulares individuais de câncer de mama agrupadas por subtipos moleculares; triplo negativas (TN, azul); receptor de hormônio positivo (HR, azul) e HER2-positivo (HER2, cinza), e baseado nos dados de anotação de Kao et al., (2009).

4.2.4 Análise de mutação e CNA no gene *PHLDA1* em câncer de mama

A frequência de mutação do gene *PHLDA1* em câncer de mama foi avaliada usando o banco de dados COSMIC em 2446 amostras de câncer de mama. Segundo as análises deste banco apenas uma porcentagem de 0.2% de mutações do tipo *missense* são encontradas no gene *PHLDA1* nas amostras analisadas (Figura 9A). Alterações no número de cópia (CNA) também foram investigadas para o gene *PHLDA1* usando um conjunto de dados do estudo METABRIC depositado na plataforma cBioPortal e mostrou que há aumento na expressão de mRNA de *PHLDA1* apenas quando há deleções heterozigotas (*Shallow deletions*) ($p=0.0115$) e nenhuma diferença quando há ganhos, amplificações e deleções homozigotas (*deep deletions*) no gene em câncer de mama (Figura 9B).

Figura 9. Frequência de mutação e análise de CNA de *PHLDA1* em câncer de mama. (A) Porcentagem de mutações no gene *PHLDA1* em câncer de mama (B) CNA vs expressão de mRNA de *PHLDA1*. Os valores P-ANOVA estão indicados. * comparação estatisticamente significativa $p < 0,05$.

4.2.5 Análise de microRNAs alvos de *PHLDA1* em câncer de mama

Para determinar o possível mecanismo envolvido na regulação do gene *PHLDA1* em câncer de mama, a participação de microRNAs neste processo foi investigada. Para este propósito foi usado o banco MirTarBase. Este banco coleta dados a partir de artigos e trabalhos com microRNAs validados a partir de diferentes experimentos de western blot, reporter assay, microarray e sequenciamento de nova geração, para a coleta de informação de microRNAs alvos de genes. Usando este banco de dados foi possível identificar 2 microRNAs que têm como alvo o *PHLDA1*; o miR-181a-5p validado por qPCR e microarray e o miR-375 validado pela técnica de microarray. Seguidamente, o perfil de expressão destes microRNAs em câncer de mama foi investigado, usando um conjunto de dados depositado no GEO com número de acesso GSE73002. Nesse trabalho foram comparadas amostras de soro de pacientes com câncer de mama e doença benigna de mama com amostras não tumorais. Para a análise do perfil de expressão do miR-181a-5p por subtipo molecular foi selecionado o conjunto de dados com número de acesso GSE81002 e outro conjunto de dados contendo dados pareados de expressão de mRNA e microRNA também foi utilizado (GSE19783). Na figura 10A, observa-se que amostras de soro de pacientes sem câncer de mama expressam níveis significativamente baixos de miR181a-5p quando comparadas com amostras de pacientes com lesões benignas e com câncer de mama ($p < 0,0001$). Quando comparada a expressão do microRNA com respeito à expressão do gene *PHLDA1* em amostras de tumor e em linhagem celular, observam-se diferenças significativas ($p = 7.183e-07$) ($p = 2.7e-12$) (Figura 10C e D) respectivamente. Em subtipos moleculares de câncer de mama se observaram diferenças significativas de expressão ($p = 0.041$) (Figura 10B). Interessantemente a curva de sobrevida, obtida utilizando-se o KM Plotter, mostrou que

altos níveis do miR-181a-5p estão associados a uma pior sobrevida nesses pacientes (Figura 10F).

Figura 10. Expressão do microRNA hsa-miR-181a-5p em câncer de mama. (A) Expressão do miR-181a-5p em soro de pacientes com câncer de mama e doença benigna de mama comparados com amostra normal (B) Expressão do miR-181a-5p em subtipos moleculares de câncer de mama (C) Expressão do miR-181a-5p vs expressão de *PHLDA1* (D) correlação da expressão de miR-181a-5p vs expressão de *PHLDA1* em linhagens celulares de mama (E) Rede de interação gênica do miR-181a-5p. (D) curvas de sobrevida para pacientes com altos e baixos níveis de expressão de miR-181a-5p em câncer de mama. Os valores P-ANOVA para as figuras 10A, B e C estão indicados, curvas construídas com o método de Kaplan meier.

4.2.6 Análise de sobrevida do gene *PHLDA1* em câncer de mama

O potencial valor prognóstico da expressão dos transcritos dos membros da família *PHLDA* foi investigado utilizando-se a base de dados KM Plotter. Foram geradas curvas de sobrevida total e livre de recorrência para todos os pacientes e para todos os subtipos de câncer de mama: Luminal A, Luminal B, Her2+ e Basal-like.

Pacientes com tumores de mama dos subtipos basais e Her2+ apresentaram aumento na sobrevida total e livre de recorrência quando os níveis de expressão dos transcritos de *PHLDA1* estavam por baixo da média (logrank $p=0.034$, logrank $p=0.0097$) (Figura 12C e D); (logrank $p=0.068$; logrank $p=2e-04$) (Figura 12E e F). Por outro lado, baixos níveis de expressão do gene *PHLDA1* foram correlacionados com aumento de sobrevida livre de recorrência nos subtipos Luminais A ($p=0.019$) (Figura 12H), Luminais B ($p=0.0011$) (Figura 12J) e nos subtipos em conjunto ($p=0.027$) (Figura 11B).

Figura 11. Curva de sobrevida para pacientes com todos os subtipos moleculares estratificados pelo nível de expressão do gene *PHLDA1*. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de *PHLDA1*. A) sobrevida total para pacientes com todos os subtipos B) Sobrevida livre de recorrência para pacientes com todos os subtipos.

Figura 12. Valor prognóstico do gene PHLDA1 em câncer de mama. Análise de sobrevida em estratificada em 5 anos (long-rank). Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA1. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

Figura 12. Valor prognóstico do gene PHLDA1 em câncer de mama. Análise de sobrevida em estratificada em 5 anos (long-rank). Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA1. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

4.2.7 Vias de interação do gene *PHLDA1* em câncer de mama

As possíveis vias de interação gênica de *PHLDA1* foram investigadas usando o software NetDecoder que permite a construção de redes específicas de um contexto biológico. Para a análise de interação foi usado o conjunto de dados com o número de acesso GSE29431 onde foram comparados tumores com amplificação do gene *ERBB2* com amostras controle. Em amostras controle, *EGFR* interage positivamente com o gene *PHLDA1* (Figura 13A), e em tumores com amplificação de *ERBB2*, *EGFR* interage com *PHLDA1* de forma negativa (Figura 13B).

Figura 13. Interações gênicas de *PHLDA1*. Rede gênica construída com o software NetDecoder (www.fte.com/products/NetDecoder.aspx). Genes ou produtos de genes são representados como nós e a relação biológica entre os nós é representada como uma malha. A cor dos nós representa a diferença de fluxo entre os fenótipos (controle e doença) e a forma o papel dos genes dentre as redes sendo o círculo (genes intermediários), os diamantes (genes fonte) e os quadrados (genes alvos). A cor vermelha das linhas indica correlação positiva e a cor azul correlação negativa entre proteínas, a intensidade da cor o grau de correlação (A) subrede de interação no controle (B) subrede de interação na doença.

4.3 Análise do perfil de expressão e de metilação do gene *PHLDA2* em câncer de mama

A análise de expressão de *PHLDA2* nos bancos de dados mostrou que contrário ao perfil de expressão do *PHLDA1*, *PHLDA2* encontra-se mais expresso nos tumores que em amostras de tecido normal ($p = 9.864e-08$) (Figura 14A); o perfil de metilação revelou que amostras tumorais apresentam um menor nível de metilação em comparação com amostra normal ($p = <2.2e-16$) (Figura 14D).

Figura 14. Expressão de mRNA e perfil de metilação do gene *PHLDA2* em amostra normal comparada com tumor de mama. Expressão de mRNA de *PHLDA2* em amostra normal comparada com amostra de tumor no banco TCGA (A) Expressão de mRNA de *PHLDA2* usando dados do banco GEO (B) perfil de metilação do gene *PHLDA2* usando o banco de dados TCGA (C) perfil de metilação do gene *PHLDA2* usando o banco de dados GEO (D).

4.3.1 Análise do perfil de expressão do gene *PHLDA2* por subtipo molecular de câncer de mama

A figura 15 representa o perfil de expressão por subtipo molecular do gene *PHLDA2* em três bancos de dados (GOBO, TCGA, GEO) e três plataformas diferentes, (HU; et al, PAM50) anteriormente descritos, todas elas mostrando o mesmo padrão de expressão. Observa-se em todas as análises que os subtipos Her2+ tendem a expressar maiores níveis de transcritos de *PHLDA2* quando comparados com os subtipos normal-like ($p=0.0189$) e o subtipo Luminal A ($p=0.0000017$) (Figura 15D) e que tumores do subtipo luminal A tendem a ter a menor expressão desse gene (Figura 15A, B, C e D).

Figura 15. Expressão de mRNA de *PHLDA2* em subtipos moleculares de câncer de mama. Box plots representando o perfil de expressão do gene *PHLDA2* em tumores de mama estratificados de acordo com subtipos moleculares de mama em diferentes bancos de dados. (A) Amostras estratificadas de acordo com a lista de genes do Hu et al., (2006) usando o banco de dados GOBO (B) Amostras estratificadas pela plataforma PAM50 usando o banco de dados GOBO (C) Amostras estratificadas pela plataforma PAM50 usando o banco de dados da GEO (D) Amostras estratificadas pela plataforma PAM50 usando o banco de dados TCGA. Os valores P-ANOVA estão indicados.

4.3.2 Análise do perfil de expressão do gene *PHLDA2* por status do receptor de estrógeno em câncer de mama.

A avaliação do perfil de expressão e metilação do gene *PHLDA2* em tumores estratificados pelo receptor de estrógeno mostrou que em tumores ER negativos a expressão dos transcritos de *PHLDA2* é mais elevada do que nos ER positivos ($p=3.379\text{e-}16$), ($p=4\text{e-}05$), ($p=0.01168$) (Figura 16A, B e C), entretanto, não foram encontradas diferenças significativas entre metilação e status de receptor ($p=4\text{e-}05$) (Figura 16D).

Figura 16. Expressão de mRNA e perfil de metilação de *PHLDA2* em tumores de mama estratificados pelo receptor de estrógeno. (A) expressão de mRNA de *PHLDA2* estratificados de acordo ao status de receptor de estrógeno usando o banco TCGA (B) expressão de mRNA de *PHLDA2* estratificados de acordo ao status de receptor de estrógeno usando o banco GEO (C) Expressão de *PHLDA2* estratificados de acordo ao status de receptor de estrógeno usando o banco GOBO (D) Nível de metilação de *PHLDA2* em tumores estratificados de acordo ao status de receptor de estrógeno usando o banco de dados TCGA.

4.3.3 Análise do perfil de expressão do gene *PHLDA2* em linhagens celulares de câncer de mama

A análise de expressão dos transcritos de *PHLDA2* em linhagens de células derivadas de câncer de mama mostrou que linhagens celulares triplo negativas e HER2+ têm uma tendência a apresentarem diminuição de expressão de *PHLDA2* em relação à linhagem de células epiteliais mamárias MCF10A. Linhagens celulares luminais e derivadas de fibroblasto pelo contrário tenderam a ter uma maior expressão.

Figura 17. Perfil de expressão de mRNA de *PHLDA2* em linhagens celulares de câncer de mama. Histogramas representando o perfil de expressão do gene *PHLDA1* em 46 linhagens celulares individuais de câncer de mama agrupadas por subtipos moleculares; triplo negativos (TN, azul); receptor de hormônio positivo (HR, azul) e HER2-positivo (HER2, cinza), e baseado nos dados de anotação de Kao et al., (2009).

4.3.4 Análise de mutação e CNA no gene *PHLDA2* em câncer de mama

A análise de mutação no banco de dados COSMIC mostrou que apenas 0.04% das amostras de câncer de mama analisadas apresentavam mutações do tipo *missense* no gene *PHLDA2* (Figura 18A). A análise de CNA revelou que ganhos no gene *PHLDA2* estão

associados com maior expressão dos transcritos de *PHLDA2* em câncer de mama ($p=0.0349$) (Figura 18B).

Figura 18. Frequência de mutação e análise de CNA do gene *PHLDA2* em câncer de mama. (A) Porcentagem de mutações no gene *PHLDA2* em câncer de mama (B) CNA vs expressão de mRNA. * comparação estatisticamente significativa $p<0,05$.

4.3.5 Análise de microRNAs alvos de *PHLDA2* em câncer de mama

A figura 19 mostra o perfil de expressão, rede de interação e análise de sobrevivência para o miR-193b-3p em câncer de mama. O miR-193b-3p foi identificado como um inibidor do gene *PHLDA2* no MirTarBase e foi validado em um trabalho pela técnica de microarray. Soro de pacientes com doenças benignas de mama e com câncer de mama apresentam maior expressão do miR-193b-3p ($p<2e-16$) (Figura 19A), sendo menos predominante nos subtipos Basais (basal vs her2+, luminal A e normal-like ($p<0.000001$)) (Figura 19B). A expressão do miR-193b-3p em amostras de tumor de mama é menor comparada com a expressão de *PHLDA2* (Figura 19C); e a rede mostrou a interação direta do miR-193b-3p com o gene *PHLDA2*. A análise de sobrevivência mostrou que pacientes com altos níveis de expressão do miR-193b-3p apresentam uma tendência de um pior curso clínico da doença (Figura 19F).

Figura 19. Expressão do microRNA hsa-miR-193b-3p em câncer de mama. (A) Expressão do miR-193b-3p em soro de pacientes com câncer de mama e doença benigna de mama comparados com amostra normal (B) Expressão do miR-193b-3p em subtipos moleculares de câncer de mama (C) Expressão do miR-193b-3p vs expressão de *PHLDA2* (D) correlação da expressão de miR-193b-3p vs expressão de *PHLDA2* em linhagens celulares de mama (E) Rede de interação gênica do miR-193b-3p (F) curvas de sobrevida para pacientes com altos e baixos níveis de expressão de em câncer de mama. Os valores P-ANOVA para as figuras 19A, B e C estão indicados.

4.3.6 Análise de sobrevida do gene *PHLDA2* em câncer de mama

A partir da base de dados KM Plotter, os dados de sobrevida total e sobrevida livre de recorrência de pacientes com câncer de mama foram investigados em relação ao nível de expressão de *PHLDA2*. Baixos níveis de expressão de *PHLDA2* foram significativamente associados com aumento de sobrevida total e sobrevida livre de recorrência para pacientes com todos os subtipos moleculares (logrank $p = 1.3 \times 10^{-6}$; logrank $p = 1.7 \times 10^{-13}$) (Figura 20A e B), com subtipo Basal (logrank $p = 0.02$; logrank $p = 0.0045$) (Figura 21C e D), Luminal A (logrank $p = 0.0052$; logrank $p = 0.00066$) (Figura 21G e H) e com subtipo luminal B (logrank $p = 0.01$; logrank $p = 1.9 \times 10^{-6}$) (Figura 21I e J). Em pacientes

com subtipos Her2 não houve diferenças significativas (logrank $p=0.47$; logrank $p=0.13$ (Figura 21E e F).

Figura 20. Curva de sobrevida para os diferentes subtipos moleculares estratificados pelo nível de expressão do gene PHLA2. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLA2. A) sobrevida total para pacientes com todos os subtipos moleculares B) Sobrevida livre de recorrência para pacientes com todos os subtipos moleculares.

Figura 21. Valor prognóstico do gene PHLDA2 em câncer de mama. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA2. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

Figura 21. Valor prognóstico do gene PHLDA2 em câncer de mama. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA2. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

4.4 Análise do perfil de expressão e de metilação do gene PHLDA3 em câncer de mama

A análise de dados de expressão nas plataformas TCGA, GOBO e GEO revelaram que tumores de mama apresentam maior expressão de PHLDA3 em relação ao tecido normal ($p = 5.322\text{e-}12$) (Figura 22A), ($p = 5.248\text{e-}07$) (Figura 22 B); da mesma forma a análise de metilação mostrou que no grupo de amostras de tumores de mama a frequência de metilação é maior do que no grupo de amostras de tecido normal ($p = <2.2\text{e-}16$) (Figura 22C), ($p = 1.212\text{e-}05$) (Figura 22D).

Figura 22. Expressão de mRNA e perfil de metilação do gene PHLDA3 em amostra normal comparada com tumor de mama. Expressão de mRNA de PHLDA3 em amostra normal comparada com amostra de tumor no banco TCGA (A) Expressão de mRNA de PHDA3 usando dados do banco GEO (B) perfil de metilação do gene PHLDA3 usando o banco de dados TCGA (C) perfil de metilação do gene PHLDA3 usando o banco de dados GEO (D).

4.4.1 Análise do perfil de expressão do gene PHLDA3 por subtipo molecular de câncer de mama

O perfil de expressão do gene PHLDA3 por subtipos moleculares mostrou que PHLDA3 apresentou o nível de expressão mais baixo nos tumores de subtipo basal quando comparado com os subtipos Her2+, luminais A, luminais B e normais no banco GEO ($p < 0,0001$) (Figura 23C) e no banco TCGA ($p < 0,0001$) (Figura 23D).

Figura 23. Expressão de mRNA de PHLDA3 em subtipos moleculares de câncer de mama. Box plots representando o perfil de expressão do gene PHLDA3 em tumores de mama estratificados de acordo com subtipos moleculares de mama em diferentes bancos de dados. (A) Amostras estratificadas de acordo com a lista de genes do Hu et al (2006) usando o banco de dados GOBO (B) Amostras estratificadas pela plataforma PAM50 usando o banco de dados GOBO (C) Amostras estratificadas pela plataforma PAM50 usando o banco de dados da GEO (D) Amostras estratificadas pela plataforma PAM50 usando o banco de dados TCGA. Os valores P-ANOVA estão indicados.

4.4.2 Análise do perfil de expressão do gene PHLDA3 por status do receptor de estrógeno em câncer de mama

O perfil de expressão do gene PHLDA3 em tumores de mama pelo status do receptor de estrógeno mostrou que em tumores ER positivos a expressão de PHLDA3 é predominante em relação aos ER negativos nos três bancos avaliados ($p = 0.0017$) (Figura 24A), ($p = 4.106 \times 10^{-5}$) (Figura 24B), ($p < 0.00001$) (Figura 24C). O perfil de metilação por outro lado, revelou que tumores ER negativos apresentaram um nível de metilação maior no gene PHLDA3 do que em tumores ER positivos ($p = 3.414 \times 10^{-10}$), (Figura 24D).

Figura 24. Expressão de mRNA e perfil de metilação de PHLDA3 em tumores de mama estratificados pelo receptor de estrógeno. (A) expressão de mRNA de PHLDA3 estratificados de acordo ao status de receptor de estrógeno usando o banco TCGA (B) expressão de mRNA de PHLDA3 estratificados de acordo ao status de receptor de estrógeno usando o banco GEO (C) expressão de mRNA de PHLDA3 estratificados de acordo ao status de receptor de estrógeno

usando o banco GOBO (D) Nível de metilação de PHLDA3 em tumores estratificados de acordo ao status de receptor de estrogênio usando o banco de dados TCGA.

4.4.3 Análise do perfil de expressão do gene PHLDA3 em linhagens celulares de câncer de mama

A análise de expressão de PHLDA3 em linhagens celulares derivadas de câncer de mama mostrou que em linhagens celulares de câncer de mama triplo negativas e her2 positivas a expressão de PHLDA3 tende a ser menor quando comparado com a linhagem celular de mama normal MCF-10A. Em contraposição, linhagens celulares hormônio dependentes tendem a ter um maior nível de expressão desse gene.

Figura 25. Perfil de expressão de mRNA de PHLDA3 em linhagens celulares de câncer de mama. Histogramas representando o perfil de expressão do gene PHLDA3 em 46 linhagens celulares individuais de câncer de mama agrupadas por subtipos moleculares; triplo negativos (TN, azul); receptor de hormônio positivo (HR, azul) e HER2-positivo (HER2, cinza). Análise baseado nos dados de anotação de Kao et al., (2009).

4.4.4 Análise de mutação e CNA no gene PHLDA3 em câncer de mama

A frequência de mutação de PHLDA3 em câncer de mama foi investigada usando o banco de dados COSMIC numa coorte de 2446 amostras. Esta análise revelou que o

gene apresenta uma frequência de menos de 1% de mutações do tipo missense. A análise de CNA (copy number alteration) foi realizada usando o conjunto de dados do estudo METABRIC depositado na plataforma cBioPortal e mostrou que há aumento na expressão de mRNA de PHLDA3 quando há ganhos ($p < 0,0001$) e amplificações ($p < 0,0001$) e nenhuma quando há deleções heterozigotas (*shallow deletions*) no gene em câncer de mama ($p = 0.3448$) (Figura 26B).

Figura 26. Frequência de mutação e análise de CNA de PHLDA3 em câncer de mama. (A) Porcentagem de mutações no gene PHLDA3 em câncer de mama (B) CNA vs expressão de Mrna de PHLDA3. * comparação estatisticamente significativa $p < 0,05$.

4.4.5 Análise de sobrevida do gene PHLDA3 em câncer de mama

Nenhuma associação foi encontrada entre expressão de *PHLDA3* (alta ou baixa) e sobrevida total na população total de pacientes com câncer de mama incluindo todos os subtipos moleculares, nas pacientes com subtipo basal-like e pacientes com subtipos Luminais A e B. (logrank $p = 0,32$; logrank $p = 0,14$; logrank $p = 0,098$; logrank $p = 0,91$) (Figura 27A, 28C, G e I), entretanto, redução da expressão de *PHLDA3* foi associada

com redução de sobrevida livre de recorrência em pacientes com câncer de mama em todos os subtipos moleculares (logrank $p = 1.1 \times 10^{-16}$; logrank $p = 0.031$; logrank $p = 0.0019$; logrank $p = 6.3 \times 10^{-8}$; logrank $p = 3.5 \times 10^{-7}$, (Figura 27B, 28D, F, H e J) respectivamente.

Figura 27. Curva de sobrevida para pacientes com todos os subtipos estratificados pelo nível de expressão do gene PHLDA3. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA2. A) sobrevida total para pacientes com todos os subtipos B) Sobrevida livre de recorrência para pacientes com todos os subtipos.

Figura 28. Valor prognóstico do gene PHLDA3 em câncer de mama. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA3. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

Figura 28. Valor prognóstico do gene PHLDA3 em câncer de mama. Análise de sobrevida estratificada em 5 anos. Em vermelho: pacientes com alta expressão e em preto, pacientes com baixa expressão dos transcritos de PHLDA3. Curvas de sobrevida total e livre de recorrência para pacientes com o subtipo molecular basal-like (C, D) subtipo HER2+ (E, F) subtipo luminal A (G, H) e subtipo Luminal B (I, J). Dados analisados usando o banco de dados Kaplan meier plotter.

4.4.6 Vias de interação do gene PHLDA3 em câncer de mama

As possíveis vias de interação gênica de *PHLDA3* foram investigadas usando o software Netdecoder. A subrede de interação da figura 29 corresponde a análise do conjunto de dados com o número de acesso GSE29431 onde foram comparados tumores com amplificação do gene *ERBB2* com amostras controle. Na rede aparece *PHLDA3* como um gene fonte conectando genes da via do p53 e genes que interagem com o gene UBC (Ubiquitin C).

Figura 29. Interações gênicas de PHLDA3. Rede gênica construída com o software netdecoder. Genes ou produtos de genes são representados como nós e a relação biológica entre os nós é representada como uma malha. A cor dos nós representa a diferença de fluxo entre os fenótipos (controle e doença) e a forma o papel dos genes dentre as redes sendo o círculo (genes intermediários), os diamantes (genes fonte) e os quadrados (genes alvos). A cor vermelha das linhas indica correlação positiva e a cor azul correlação negativa entre proteínas, a intensidade da cor o grau de correlação (A) subrede de interação no controle (B) subrede de interação na doença.

5. DISCUSSÃO

Informações sobre o papel funcional, o perfil de expressão e o potencial valor prognóstico dos genes membros da família PHLDA (Pleckstrin homology-like domain, family A) em câncer de mama são bastante limitados na literatura. Assim sendo, o principal objetivo deste trabalho foi o de utilizar análises *in silico*, para avaliar o papel funcional e o potencial valor prognóstico de *PHLDA1*, *PHLDA2* e *PHLDA3* em câncer de mama.

A pesar da alta homologia de sequência que os membros da família PHLDA compartilham estes genes apresentam níveis de expressão e metilação não correlacionados em câncer de mama. O gene *PHLDA3* por exemplo tem uma tendência a ter maiores níveis de expressão de mRNA e proteína do que os outros membros, e os genes *PHLDA1* e *PHLDA3* são significativamente mais metilados do que o gene *PHLDA2* em câncer de mama. Estes dados sugerem que a expressão de cada gene da família PHLDA é diferente e varia em câncer de mama.

Para avaliar o perfil de expressão de *PHLDA1* em câncer de mama, nós procuramos dados depositados em duas plataformas diferentes contendo dados de expressão de mRNA em câncer de mama e encontramos em ambos os bancos expressão reduzida de mRNA de *PHLDA1* em tumores comparado com amostra normal de mama. Estudos prévios na literatura confirmam nossos resultados. Redução de expressão de proteína de PHLDA1 foi evidenciada em tumores de mama primário e associado com progressão e mau prognóstico pelo nosso grupo em um estudo prévio (NAGAI et al., 2007). JHONSON et al 2011, analisou expressão de proteína de PHLDA1 em 114 tumores usando TMA e observou que 66% das amostras analisadas apresentaram marcação negativa para PHLDA1; de maneira similar (LI et al., 2014a), encontraram

expressão de proteína de *PHLDA1* significativamente reduzida em tecido tumoral de mama comparada com tecido epitelial normal adjacente.

Contrário ao reportado na literatura, neste estudo, nós observamos que *PHLDA1* é menos expresso nos tumores ER+ do que nos ER-, e que em subtipos moleculares luminais, alta expressão de *PHLDA1* está associada com aumento de sobrevida livre de recorrência. De forma similar em linhagens celulares ER positivas observamos que a expressão de *PHLDA1* foi menor do que na linhagem celular normal MCF-10A. (MARCHIORI; CASOLARI; NAGAI, 2008), identificaram um sitio putativo de ligação do receptor de estrógeno na região promotora do gene *PHLDA1*. Consequentemente o grupo mostrou que *PHLDA1* é regulado pelos estrógenos após observar um aumento de expressão do transcrito em células tratadas com estrógenos e inibição de expressão após tratamento com um composto antiestrogênico. JHONSON et al, 2011, realizou análise multivariada a partir de expressão de proteína e dados clínicos de pacientes com tumores primários de mama e demonstrou que expressão de *PHLDA1* está correlacionada com expressão de estrógeno nesses tumores. Usando bancos de dados públicos, KASTRATI et al., 2014 associaram alta expressão de mRNA de *PHLDA1* com metástase e agressividade de tumores ER positivos.

A análise por subtipo molecular nos três bancos de dados avaliados revelou que *PHLDA1* foi significativamente mais expresso nos subtipos basais e que alta expressão do gene nesse subtipo está associada com mau prognóstico em câncer de mama. Tumores de tipo basal são os mais agressivos entre os subtipos moleculares de mama devido à falta de terapias alvo específicas para este tipo de tumor. Geralmente são de progressão rápida e apresentam baixa sobrevida. Concordando com o dado de expressão, tumores basais foram significativamente menor metilados do que os outros

subtipos moleculares o que indica que neste subtipo, *PHLDA1* pode estar sendo regulado por este mecanismo.

Os subtipos HER2 positivos apresentaram uma tendência a ter menor expressão de *PHLDA1* entre todos os subtipos moleculares. A relação entre *PHLDA1* e HER2 já foi avaliada por LI et al., 2014, quem correlacionou positivamente os níveis de expressão de HER2 com os de *PHLDA1*. LI et al., 2014 mostrou que após inibição dos níveis de proteína de EGFR ou HER2, os níveis de proteína de *PHLDA1* são reduzidos. O mesmo grupo também mostrou que em tumores de mama HER2 positivo a expressão de *PHLDA1* é maior quando comparado com os HER2 negativos.

Nós analisamos as possíveis vias de interação gênica de *PHLDA1* usando um conjunto de dados depositado no banco GEO contendo dados de tumores com e sem amplificação de HER2. Contrário aos achados de LI et al, 2014, e concordando com nossos achados, nós encontramos que *PHLDA1* interage com EGFR de forma negativa em tumores com amplificação de HER2 comparado com o controle onde interage de forma positiva. Estes resultados demonstram que evidentemente há uma regulação do gene *PHLDA1* pela via HER2/EGFR não entanto os mecanismos precisam ser explorados a profundidade.

Até o momento nenhum estudo na literatura reportou mutações no gene *PHLDA1* em câncer de mama. Neste estudo nos encontramos que mutações no gene *PHLDA1* são pouco frequentes e se apresentam em 1% dos dados analisados em COSMIC. Usando dados de alteração no número de cópia depositados na plataforma TCGA, encontramos que deleções no DNA podem alterar a expressão do gene *PHLDA1* em câncer de mama.

Visando identificar o mecanismo de regulação do gene *PHLDA1* em câncer de mama a participação de microRNAs foi investigada. Os microRNAs são pequenas moléculas de RNA não codificantes que modulam a ação dos genes degradando ou inibindo a tradução dos seus mRNAs alvos (ZHANG et al., 2015). Expressão aberrante de microRNAs em diferentes tipos de câncer tem sido associada com processos de proliferação celular, migração, invasão, apoptose e sobrevida, motivo pelo qual têm sido propostos como possíveis biomarcadores em câncer.

PHLDA1 foi identificado como alvo do miR-181a-5p através de banco MIRTARBASE que usa uma ferramenta para predição de microRNAs alvos de genes. KASTRATI et al., 2015 demonstrou que o gene *PHLDA1* é um alvo direto do miR-181 experimentalmente, usando inibidores de microRNAs chamados target protector. O grupo de KASTRATI et al., 2015 desenharam essas moléculas exclusivamente para bloquear a ligação do miR-181-a à região 3'UTR de *PHLDA1*. Depois de transfectar células MCF-7 com moléculas sintéticas do miR-181a, a expressão de *PHLDA1* foi restabelecida confirmando a regulação de *PHLDA1* por parte do miR-181a.

O miR181a-5-p pertence à família miR-181s que incluem três membros miR-181a, b, c, e d. Esta família de genes tem sido implicada em desenvolvimento, progressão, metástase, prognóstico e resposta a terapias em câncer de endométrio, câncer de ovário, glioma, câncer de fígado, câncer coloretal, câncer gástrico, câncer de pulmão, câncer de mama, carcinoma cervical, câncer pancreático, osteosarcoma, carcinoma de células escamosas orais, linfoma de células B, câncer de tireoide (DESIGN, 2015). O miR181a-5-p regula um amplo número de genes alvos como BCL2/K-RAS/GATA6/ CDX2, envolvidos em processos de proliferação e apoptose nas células (DESIGN, 2015)

Nós escolhemos um conjunto de dados depositado no banco GEO com aproximadamente 4000 amostras de soro de pacientes com e sem tumor de câncer de mama para analisar a expressão do miR181a. Usando esses dados nós observamos que a expressão de miR-181a-5p é alta em tumores de mama e está associada com baixa sobrevida. Também observamos que quando a expressão do miR-181a é alta a do *PHLDA1* é baixa sugerindo que a expressão de *PHLDA1* pode estar sendo regulada por este microRNA. De maneira similar (BISSO et al., 2013) avaliou o perfil de expressão do miR-181a em câncer de mama e reportou que a superexpressão de miR181a-5-p está correlacionada com tumores agressivos, alto grau histológico e alto índice proliferativo. Em tumores de mama triplo negativos, alta expressão do miR181a-5-p foi proposto como um potencial marcador de quimioresistencia. Expressão de miR-181a-5p foi associado com indução de fenótipo metastático em linhagens celulares de mama MDA-MB-231 tratadas com TGF- β (TAYLOR et al., 2013).

Nossos dados junto com os achados na literatura nos levam a formular a hipótese que o miR-181a-5p, junto com alteração do padrão de metilação regulam a expressão de *PHLDA1* em câncer de mama.

O gene *PHLDA2* está localizado na região 11p15.5 uma região cromossômica alvo de *imprinting*, processo crucial no desenvolvimento fetal. De acordo com uma das hipóteses o processo de *imprinting* surgiu como um mecanismo de competição paterna-materna pelo sucesso reprodutivo. Consistentemente, genes com expressão a partir do alelo paterno têm funções promotoras do crescimento enquanto que os maternos apresentam funções relacionadas à inibição de crescimento; porque genes silenciados que se expressam a partir do alelo materno particularmente nesta região tendem a inibir a proliferação celular (QIAN et al., 1997).

Em vista de que o gene *PHLDA2* é altamente expresso na placenta de mamíferos, diversos estudos têm investigado o papel deste gene em doenças relacionadas com desenvolvimento fetal e placentário. Alta expressão de *PHLDA2* tem sido encontrada em bebês com baixo peso ao nascimento em vários estudos independentes. Expressão de *PHLDA2* foi associada com baixa velocidade de crescimento de fêmur fetal (LEWIS et al., 2012; OHKI et al., 2014a); Deleção do gene resulta em alargamento placentário em camundongos (SAXENA et al., 2002), e perda de imprinting afeta negativamente o crescimento do feto e da placenta (SALAS et al., 2004; TUNSTER; TYCKO; JOHN, 2010).

PHLDA2 foi descrito como o primeiro gene apoptótico alvo de imprinting (LEE; FEINBERG, 1998). *PHLDA2* foi relacionado com indução de apoptose em câncer gástrico pela regulação da via PI3K/AKT (SUNG et al., 2017). Em células de osteosarcoma SaOS2, superexpressão de *PHLDA2* induz supressão de crescimento e apoptose associada com a expressão de caspase 3 (DAI et al., 2012). Perda de expressão de *PHLDA2* em osteosarcoma humano foi correlacionada com a presença de metástases (YAN; LV; GUO, 2016).

A perda de expressão de *PHLDA2* tem sido reportada em tumor de Wilms (SAXENA et al., 2003), mola hidatiforme completa (SCHWIENBACHER et al., 2000) e osteosarcoma onde tem sido definido como um potencial supressor de tumor.

Contrariando a hipótese do papel supressor do gene *PHLDA2*, os resultados neste estudo mostram que altos níveis de expressão de *PHLDA2* estão associados a subtipos com fenótipos mais agressivos como HER2 positivos, alto grau histológico e mau prognóstico sugerindo que *PHLDA2* é um potencial preditor de mal prognóstico em câncer de mama.

Os subtipos HER2 positivos são considerados um dos mais agressivos entre os subtipos moleculares de mama. Estes subtipos se caracterizam pela superexpressão de HER2, um oncogene da família de receptores tirosina quinase associado com baixa sobrevida e mau prognóstico em pacientes com câncer de mama. Ativação do oncogene depende da ativação de vias como AKT/PI3K e as vias ERK/MAPK que são efetores imediatos de HER2 e levam a estimulação do crescimento celular, progressão, proliferação, bloqueio de apoptose, angiogênese e invasão (WANG et al., 2015).

Wang et al., 2015, identificaram *PHLDA2* repetidamente em vários estudos de perfil transcricional como um gene induzido pela via de sinalização de EGFR e HER2. Em linhagens celulares SKBR3 tratadas com lapatinib, um inibidor de HER2, a expressão de mRNA e proteína de *PHLDA2* foi reprimida em comparação com as células controle tratadas com DMSO. Em linhagens celulares de mama SKBR3 e linhagem celulares de pulmão H1975 tratadas com inibidores de HER2 (LI et al., 2014a), *PHLDA2* apareceu como um dos genes mais diferencialmente expressos. Em nossos dados de expressão em linhagem celular, encontramos que as linhagens celulares HCC1954 e SKBR3 que superexpressam HER2 tendem a ter maior expressão de *PHLDA2* quando comparadas com a linhagem de mama normal MCF-10A, o que confirma os dados do estudo de WANG et al., 2015.

Interessantemente em tumores de tipo luminal A, que carecem de expressão de HER2, com baixos índices proliferativos e de melhor prognóstico, a expressão de *PHLDA2* é menor do que em todos os subtipos moleculares.

O gene *PHLDA2* é um gene regulado por um mecanismo epigenético que envolve metilação, o que nos levou a especular que o possível mecanismo responsável pela alteração do perfil de expressão desse gene em câncer de mama seja alteração do padrão de metilação do gene.

Nós analisamos o perfil de metilação de *PHLDA2* em tumor comparado com amostra normal, nos cinco subtipos de câncer de mama, e por status do receptor de estrógeno e não encontramos nenhuma diferença na metilação. Adicionalmente a análise de correlação entre metilação e expressão de mRNA foi negativa fraca o que indica que outros mecanismos diferentes de perda de metilação poderiam estar associados com a alta expressão desse gene em câncer de mama.

Schwienbacher et al., 2000 mostraram em um estudo que *PHLDA2* encontrava-se com baixa expressão em tumor de Wilms e que o mecanismo de regulação do gene nesse tumor era independente de metilação. Por outro lado, hipometilação foi descrito como o principal mecanismo de inativação do gene *PHLDA2* em células de osteosarcoma humano. Em outro estudo (LI; MENG; GUO, 2008); realizaram uma análise do padrão de metilação de células osteoblásticas transformadas com dois agentes carcinogênicos. Os resultados mostraram que a região promotora dos genes *H19* e *PHLDA2* se encontravam hipometiladas, e como consequência, esses genes apresentavam aumento de expressão. Estes dados nos levam a pensar que a regulação da expressão do gene depende do tipo de tumor.

Nós investigamos também a frequência de mutações no gene *PHLDA2* e alterações no número de cópia (CNV) do gene em câncer de mama. Segundo os dados de COSMIC mutações nesse gene se apresentam em baixa frequência; por outro lado, ganhos no DNA de *PHLDA2* foram associados com aumento de expressão do gene.

A busca pelo possível mecanismo regulatório envolvido na expressão de *PHLDA2* em câncer de mama nos levou à procura de microRNAs alvos de *PHLDA2*.

PHLDA2 foi identificado no banco MIRTARBASE como alvo do microRNA mir-193b-3p. O mir-193b-3p tem sido descrito como inibidor de proliferação, migração e invasão em células de câncer gástrico (WANG et al., 2016), em células de carcinoma hepatocelular (XU et al., 2010), em células de melanoma (CHEN et al., 2010), em células de carcinoma renal (TREVISANI et al., 2016), em células de carcinoma epidermal escamoso (INSTITUTET et al., 2013) em câncer de próstata (RAUHALA et al., 2010) e em células de câncer de mama (LI; YAN; SHAO, 2009).

O mir-193b-3p foi proposto como um potencial marcador de diagnóstico em câncer colorectal (YONG; LAW; WANG, 2013), marcador preditivo de doença crônica de rim (TREVISANI et al., 2016) e biomarcador de resposta a imatinib em pacientes com leucemia crônica (RAMACHANDRAN et al., 2017).

O mir-193b-3p tem sido identificado como regulador de vários oncogenes em câncer de ovário, em leucemias agudas (GAO et al., 2011).

Leivonen et al 2008, identificou o miR-193b como modulador do ER alfa e inibidor do crescimento celular em células MCF-7. Mais tarde o mesmo autor em outro trabalho identificou 4 genes alvos do miR-193b envolvidos em progressão tumoral e demonstrou mediante ensaios funcionais o efeito inibitório do crescimento nas células MCF-7 mediado pela regulação do miR-193b nesses genes (KOHONEN et al., 2009).

UpA, um gene definido como promotor de crescimento, invasão e metástase, foi identificado como alvo do miR-193b. Em linhagens celulares de mama e em modelo xenográfico a inibição dos níveis de miR-193b levou a aumento de expressão de UpA e induziu potencial metastático. Diminuição do efeito invasivo, migratório e proliferativo foi observado após superexpressão do miR-193b mostrando o potencial papel funcional do microRNA na regulação desse gene em câncer de mama (LI; YAN; SHAO, 2009) 2009).

Nossos resultados mostraram que miR-193b-3p é mais expresso em soro de pacientes com tumor de mama do que em soro de pacientes sem doença.

Nós analisamos um conjunto de amostras de 100 tumores de mama com dados de expressão de microRNA e mRNA para correlacionar o perfil de expressão do miR-193b-3p com a expressão de *PHLDA2*. A análise mostrou que há uma relação inversa entre a expressão do mir-193b-3p e a expressão do gene *PHLDA2* no conjunto de amostras analisado. Adicionalmente a rede de interação mostrou que mir-193b-3p regula genes relacionados com tumorigenese como *CCND1* que codifica a proteína Ciclina D1, o gene *ESR1* que codifica o receptor de estrógenos.

Todos estes achados indicam que o mir-193b-3p é um microRNA com funções associadas a supressão de tumores, não obstante nas amostras analisadas nos encontramos resultados inconclusivos.

Expressão elevada do mir-193b-3p em tumor pode ser devida a um mecanismo do miRNA para regular a expressão de *PHLDA2* em câncer de mama no conjunto de amostras analisado. A discrepância observada nos nossos resultados poderia ser também em parte pelo tipo de amostra analisada neste estudo, que neste caso consiste de soro. Os microRNAs circulantes estáveis estão presentes no soro, plasma e sangue de pacientes pelo qual numerosos estudos tem procurado a detecção de microRNAs circulantes com potencial no diagnóstico e prognóstico do câncer e outras doenças. No entanto, alguns estudos têm mostrado que não há concordância entre os microRNAs detectados em tumor com os detectados no sangue. Estas diferenças têm sido atribuídas principalmente à heterogeneidade de cada paciente e aos procedimentos de extração dos microRNAs circulantes o que pode contribuir à variabilidade de expressão (ZHANG et al., 2015). Apesar disso, os trabalhos descritos na literatura indicam que o miR-193b-3p é um microRNA importante na regulação de genes envolvidos em progressão tumoral e

potencial biomarcador de diagnóstico, prognóstico e tratamento; portanto, a relação deste microRNA com o gene *PHLDA2* em câncer de mama merece maiores análises.

Os resultados neste estudo demonstram que contrário ao descrito na literatura em outros modelos de tumor, o gene *PHLDA2* pode exercer funções relacionadas à promoção de tumor e mau prognóstico em pacientes com câncer de mama.

O gene *PHLDA3* vem chamando bastante a atenção de diferentes investigadores devido a seu potencial papel como regulador do gene p53 e do AKT em processos tumorigênicos. O papel do gene *PHLDA3* tem sido avaliado em tumores de próstata, em tumores neuroendócrinos, em tumores pulmonares e em osteosarcoma, no entanto, em câncer de mama somente três trabalhos tem explorado apenas o nível de expressão do gene *PHLDA3*. Nesses trabalhos, a expressão de *PHLDA3* é reduzida em tumores de mama em relação a tecido normal (CHRISTGEN et al., 2012; OHKI et al., 2014a; SOUNG et al., 2011a). Concordando com os dados desses grupos, nos também encontramos que *PHLDA3* é significativamente mais expresso em tecido normal, e adicionalmente observamos alta expressão em tumores ER positivos, em subtipos menos agressivos como os luminais, e em tumores de grau 1.

Dentre os subtipos moleculares, os basais destacaram-se por apresentarem o menor nível de expressão de *PHLDA3*. Os subtipos basais caracterizam-se porque apresentam a maior taxa de mutações no gene p53 sendo detectadas em até 88% dos casos (BERTHEAU et al., 2013).

Já é bem estabelecida na literatura a ligação entre baixa expressão do gene *PHLDA3* com mutação no gene p53. Kawase et al., 2009 mostrou aumento de expressão do gene em células MDA-MB-468 com superexpressão ectópica de p53 selvagem. Em um conjunto de amostras com acúmulo de p53, indicativa de mutação deletérea no gene TP53, a expressão de *PHLDA3* foi baixa ou ausente (CHRISTGEN et al., 2012).

PAPPAS et al., 2017 examinou 4899 casos de 18 tipos de câncer com mutações no gene p53 e encontrou que mutação no gene PT53 está significativamente associada com baixos níveis de expressão de *PHLDA3* em todos os tumores incluindo câncer de mama. No mesmo estudo, usando ChIP-seq em células epiteliais não transformadas MCF10A; p53 selvagem liga-se a *PHLDA3*.

Dentre todos os genes envolvidos em processos de carcinogênese, p53 é o de maior importância. Este gene é considerado um gene supressor de tumor por sua participação no reparo do DNA, apoptose e controle do ciclo celular pelo qual é amplamente utilizado como alvo de terapias e marcador de prognóstico clínico (CHEUNG, 2005).

p53 pode exercer seus efeitos através da regulação de genes em diferentes tecidos. Muitos desses genes alvos de p53 têm sido identificados como reguladores de vários processos envolvidos na prevenção da tumorigênese (KAWASE et al., 2009a). Por exemplo P53 interage com p21 cujo produto proteico é um inibidor de quinase dependente de ciclina, que promove o arraste do ciclo celular, favorecendo o reparo do DNA danificado; o p53 e o p14ARF, um outro potente supressor, estão funcionalmente ligados e relacionados à controle do ciclo celular e apoptose. Alteração desses genes está associada com patogênese de diferentes tipos de tumores.

PHLDA3 pode inibir a ativação de Akt por se ligar, via o domínio PH-like, a fosfolípidos na face interna da membrana plasmática (KAWASE et al., 2009b). Em conjunto esses dados sugerem que em câncer de mama p53 pode exercer seu papel de gene supressor de tumor através da regulação do gene *PHLDA3*, um gene regulador de AKT.

A análise de metilação revelou que *PHLDA3* é um gene metilado em câncer de mama. O silenciamento génico por hipermetilação é um importante mecanismo de carcinogênese com grande potencial na prevenção e terapia do câncer de mama (JOVANOVIC et al., 2010). Muitos supressores de tumor e outros genes relacionados a câncer tem sido encontrados com hipermetilação em células de câncer e tumores primários. A maioria desses genes aberrantemente metilados tem papel na regulação do ciclo celular, reparo de DNA e apoptose.

Nos encontramos diferenças na metilação de *PHLDA3* entre os subtipos basais e os não basais. Um resultado similar foi obtido pelo grupo de OHKI et al., 2014 que encontraram diminuição de metilação e aumento de expressão de *PHLDA3* em células neuroendócrinas de pulmão após tratamento com 5AZA. Estes achados reforçam a importância da metilação como mecanismo regulador do gene *PHLDA3* em câncer.

Similar aos seus homólogos, *PHLDA3* apresenta uma porcentagem baixa de mutações; por outro lado, nós encontramos que ganhos e amplificações no DNA de *PHLDA3* influem na expressão de mRNA do gene em câncer de mama. OHKI et al., 2014b, associou perda de heterozigose e hipermetilação com supressão da transcrição de *PHLDA3* em tumores neuroendócrinos pancreáticos. TAKIKAWA; OHKI, 2017, encontraram alterações no número de cópia do gene *PHLDA3* em tumores endócrinos de pulmão e a associou com diminuição de expressão do gene. SOUNG et al., 2011b; avaliaram mutações somáticas em vários tipos de tumores dentre eles carcinoma de mama e somente uma mutação silenciosa foi encontrada em adenocarcinoma de pulmão indicando que mutações somáticas de *PHLDA3* é um evento raro em vários tipos de câncer humanos.

Nossos achados sugerem que a expressão do gene *PHLDA3* em câncer de mama pode ser regulada principalmente pela expressão de p53, por mecanismos epigenéticos como metilação e por alterações no número de cópia como ganhos e amplificações.

Somente dois trabalhos tem avaliado o potencial prognóstico de *PHLDA3* em processos tumorigênicos. TA et al., 2015, analisaram a sobrevida específica de câncer e sobrevida livre de doença de 84 pacientes com carcinoma de células escamosas esofágicas com os níveis de expressão do gene *PHLDA3*. As análises mostraram que a sobrevida aumenta quando os níveis de expressão de *PHLDA3* são altos sugerindo que *PHLDA3* pode ser um importante marcador em pacientes com alto risco da doença. Em pacientes com tumores neuroendócrinos pancreáticos perda de heterozigose no locus de *PHLDA3* é significativamente associada com fenótipo maligno (OHKI et al., 2014b).

Até o momento nosso estudo é o primeiro em analisar a expressão da família de genes *PHLDA1*, *PHLDA2* e *PHLDA3* em subtipos moleculares de câncer de mama e por status do receptor de estrógeno associando com prognóstico clínico de pacientes. Embora estudos clínicos e funcionais avaliando o papel destes genes em câncer de mama ainda são necessários para confirmar esses achados, os resultados apresentados neste estudo indicam o envolvimento dos membros da família *PHLDA* no processo de tumorigênese da mama e o potencial valor prognóstico de *PHLDA1*, *PHLDA2* e *PHLDA3* em câncer de mama. Nossos dados são importantes e servem como ferramenta para novos estudos clínicos e experimentais, alguns dos quais já estão em desenvolvimento em nosso laboratório, visando melhor entender o potencial valor prognóstico e preditivo de resposta desses genes como biomarcadores em câncer de mama e em outros tipos de tumores.

6. CONCLUSÃO

Os dados compilados até o momento, por mineração de dados disponíveis em bancos de dados públicos, para avaliar o perfil de expressão e potencial valor prognóstico dos membros da família PHLDA em câncer de mama, demonstram que alterações na expressão dos genes da família PHLDA tem um impacto na progressão do câncer de mama. Além disso, podemos concluir que:

- ✓ Existe uma correlação inversa na metilação e expressão de mRNA e proteína dos genes *PHLDA1* e *PHLDA3* em tumores de mama.
- ✓ O gene *PHLDA1* apresenta expressão diferencial dependendo do subtipo molecular sendo mais expresso nos tumores mais agressivos como os basais. Nesses tumores alta expressão confere baixa sobrevida. A expressão deste gene pode ser regulada por diversos mecanismos como metilação do DNA e participação do miR 181a-5p.
- ✓ Nossos resultados sugerem que contrário ao encontrado em outros tumores, em câncer de mama *PHLDA2* tem características de oncogene e está associado com mau prognóstico. O possível mecanismo de regulação deste gene pode estar determinado por mutações no gene e a participação do micro RNA 193b-3p.
- ✓ A expressão do gene *PHLDA3* em subtipos basais é baixa e está associada com diminuição de sobrevida livre de recorrência. Nossos dados corroboram os achados na literatura que demonstram que o gene *PHLDA3* está intimamente ligado com o gene TP53, frequentemente mutado neste tipo de tumores.

7. REFERENCIAS

ABBAS, M. et al. Breast Cancer Biomarker Discovery in the Functional Genomic Age: A Systematic Review of 42 Gene Expression Signatures. **Biomarker Insights**, v.5, p.103-118, 2010.

ALIZART, M. et al. Molecular classification of breast carcinoma. **Diagnostic Histopathology**, v. 18, n. 3, p. 97–103, 2012.

AMENDOLA, L. C. B.; VIEIRA, R. A contribuição dos genes BRCA na predisposição hereditária ao câncer de mama. **Revista Brasileira de Cancerologia**, v. 51, n. 4, p. 325–330, 2005.

APOSTOLIDOU, S. et al. Elevated placental expression of the imprinted PHLDA2 gene is associated with low birth weight. **Journal of Molecular Medicine**, v. 85, n. 4, p. 379–387, 2007.

BAI, X. et al. PIK3CA and TP53 gene mutations in human breast cancer tumors frequently detected by ion torrent DNA sequencing. **PLoS ONE**, v. 9, n. 6, 2014.

BARBARESCO, A. A. et al. Metilação aberrante de DNA, câncer cervical e HPV. **Femina**, v. 40, n. 5, 2012.

BASSERI, S. et al. Loss of TDAG51 results in mature-onset obesity, hepatic steatosis, and insulin resistance by regulating lipogenesis. **Diabetes**, v. 62, n. 1, p. 158–69, jan. 2013.

BERTHEAU, P. et al. P53 in Breast Cancer Subtypes and New Insights Into Response To Chemotherapy. **Breast (Edinburgh, Scotland)**, v. 22 Suppl 2, p. S27-9, 2013.

BIRNEY, E. et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. **Nature**, v. 447, n. 7146, p. 799–816, 2007.

BISSO, A. et al. Oncogenic miR-181a / b affect the DNA damage response in aggressive breast cancer. v. 4101, n. December 2017, 2013.

BONATTO, N. et al. PHLDA1 (pleckstrin homology-like domain, family A, member 1) knockdown promotes migration and invasion of MCF10A breast epithelial cells. **Cell Adhesion and Migration**, v. 6918, n. June, p. 1–10, 2017.

CAMPBELL, L. L.; POLYAK, K. Breast tumor heterogeneity: Cancer stem cells or clonal evolution? **Cell Cycle**, v. 6, n. 19, p. 2332–2338, 2007.

CHEN, J. et al. MicroRNA-193b Represses Cell Proliferation and Regulates Cyclin D1 in Melanoma. v. 176, n. 5, p. 2520–2529, 2010.

CHENG, P. F.; DUMMER, R.; LEVESQUE, M. P. Data mining The Cancer Genome Atlas in the era of precision cancer medicine. **Swiss Medical Weekly**, v. 145, n.

September, p. 1–5, 2015.

CHEUNG, K.-Y. J. AND V. G. Identification of novel p53 target genes in ionizing radiation response. **Cancer research**, v. 65, n. 17, p. 7666–73, 1 set. 2005.

CHIN, L. et al. Making sense of cancer genomic data. p. 534–555, 2011.

CHRISTGEN, M. et al. IPH-926 lobular breast cancer cells harbor a p53 mutant with temperature-sensitive functional activity and allow for profiling of p53-responsive genes. **Laboratory Investigation**, v. 92, n. 11, p. 1635–1647, 2012.

CLARKE, C. et al. Correlating transcriptional networks to breast cancer survival: A large-scale coexpression analysis. **Carcinogenesis**, v. 34, n. 10, p. 2300–2308, 2013.

COLDITZ, G. A.; BOHLKE, K.; BERKEY, C. S. Breast cancer risk accumulation starts early: Prevention must also. **Breast Cancer Research and Treatment**, v. 145, n. 3, p. 567–579, 2014.

CURTIS, C. et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. **Nature**, v. 486, n. 7403, p. 346–352, 2012.

DAI, H. et al. TSSC3 overexpression associates with growth inhibition, apoptosis induction and enhances chemotherapeutic effects in human osteosarcoma. **Carcinogenesis**, v. 33, n. 1, p. 30–40, 1 jan. 2012.

DE PERGOLA, G.; SILVESTRI, F. Obesity as a Major Risk Factor for Cancer. **J Obes**, v. 2013, p. 291546, 2013.

DENG, M. et al. Web-TCGA: An online platform for integrated analysis of molecular cancer data sets. **BMC Bioinformatics**, v. 17, n. 1, p. 1–7, 2016.

DESIGN, D. Hsa-microRNA-181a is a regulator of a number of cancer genes and a biomarker for endometrial carcinoma in patients : a bioinformatic and clinical study and the therapeutic implication. 2015.

DOUGLAS HANAHAN* AND ROBERT A. WEINBERG. The Hallmarks of Cancer. **Cell**, v. 100, p. 57–70, 2000.

EROLE, P. et al. Molecular biology in breast cancer: Intrinsic subtypes and signaling pathways. **Cancer Treatment Reviews**, v. 38, n. 6, p. 698–707, 2012.

FALAVIGNA, M. et al. Effects of lifestyle modification after breast cancer treatment: a systematic review protocol. **Syst Rev**, v. 3, n. 1, p. 72, 2014.

FERLAY, J. et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. **International Journal of Cancer**, v. 136, n. 5, p. E359–E386, 2015.

FERNANDEZ, A. F. et al. A DNA Methylation Fingerprint of 1 , 628 Human Samples. **Genome Research**, p. 407–419, 2011.

FERTIG, E. J. et al. Analysis of gene expression of secreted factors associated with

breast cancer metastases in breast cancer subtypes. **Scientific Reports**, v. 5, n. January, 2015.

FRANK, D. et al. A novel pleckstrin homology-related gene family defined by Ipl / Tssc3 ,TDAG51, and Tih1: tissue-specific expression, chromosomal location, and parental imprinting. v. 1159, p. 1150–1159, 1999.

GAO, X. et al. MicroRNA-193b regulates c-Kit proto-oncogene and represses cell proliferation in acute myeloid leukemia. v. 35, p. 1226–1232, 2011.

GJOREVSKI, N.; NELSON, C. M. Integrated morphodynamic signalling of the mammary gland. **Nature reviews. Molecular cell biology**, v. 12, n. 9, p. 581–593, 2011.

GOSS, P. E. et al. Planning cancer control in Latin America and the Caribbean. **The Lancet Oncology**, v. 14, n. 5, p. 391–436, 2013.

HANAHAN, D.; WEINBERG, R. A. Hallmarks of cancer: The next generation. **Cell**, v. 144, n. 5, p. 646–674, 2011.

HOSSAIN, G. S. et al. TDAG51 is induced by homocysteine, promotes detachment-mediated programmed cell death, and contributes to the development of atherosclerosis in hyperhomocysteinemia. **The Journal of biological chemistry**, v. 278, n. 32, p. 30317–27, 8 ago. 2003.

INSTITUTET, S. K. et al. MiR-193b / 365a cluster controls progression of epidermal squamous cell carcinoma. n. October 2015, 2013.

JIANG, P.; LIU, X. S. Big data mining yields novel insights on cancer. **Nature Genetics**, v. 47, n. 2, p. 103–104, 2015.

JIN, F. et al. Lentivirus-mediated PHLDA2 overexpression inhibits trophoblast proliferation, migration and invasion, and induces apoptosis. **International Journal of Molecular Medicine**, p. 949–957, 2016.

JOE, S.; NAM, H. Prognostic factor analysis for breast cancer using gene expression profiles. **BMC Medical Informatics and Decision Making**, v. 16, n. Suppl 1, 2016.

JOHNSON, E. O. et al. PHLDA1 is a crucial negative regulator and effector of Aurora A kinase in breast cancer. **Journal of cell science**, v. 124, n. Pt 16, p. 2711–22, 15 ago. 2011.

JOVANOVIC, J. et al. The epigenetics of breast cancer. **Molecular Oncology**, v. 4, n. 3, p. 242–254, 2010.

JURCA, G. et al. Integrating text mining , data mining , and network analysis for identifying genetic breast cancer trends. **BMC Research Notes**, p. 1–35, 2016.

KASTRATI, I.; CANESTRARI, E.; FRASOR, J. PHLDA1 expression is controlled by an estrogen receptor-NFκB-miR-181 regulatory loop and is essential for formation of

ER+ mammospheres. **Oncogene**, n. April, p. 1–8, 23 jun. 2014.

KAWASE, T. et al. PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt. **Cell**, v. 136, n. 3, p. 535–550, 2009a.

KAWASE, T. et al. PH Domain-Only Protein PHLDA3 Is a p53-Regulated Repressor of Akt. **Cell**, v. 136, n. 3, p. 535–550, fev. 2009b.

KERLEY-HAMILTON, J. S. et al. A p53-dominant transcriptional response to cisplatin in testicular germ cell tumor-derived human embryonal carcinoma. **Oncogene**, v. 24, n. 40, p. 6090–6100, 2005.

KEY, T. J. et al. Body Mass Index, Serum Sex Hormones, and Breast Cancer Risk in Postmenopausal Women. **JNCI Journal of the National Cancer Institute**, v. 95, n. 16, p. 1218–1226, 2003.

KEY, T. J.; VERKASALO, P. K.; BANKS, E. Full-Text. v. 2, n. 0, p. 133–140, 2001.

KLUTTIG, A.; SCHMIDT-POKRZYWNIAK, A. Established and suspected risk factors in breast cancer aetiology. **Breast Care**, v. 4, n. 2, p. 82–87, 2009.

KOHONEN, P. et al. Protein lysate microarray analysis to identify microRNAs regulating estrogen receptor signaling in breast cancer cell lines. n. April, p. 3926–3936, 2009.

LEE, M. P.; FEINBERG, A. P. Genomic Imprinting of a Human Apoptosis Gene Homologue , Genomic Imprinting of a Human Apoptosis Gene Homologue , TSSC31. p. 1052–1056, 1998.

LEMLER, D. J. et al. DCYTB is a predictor of outcome in breast cancer that functions via iron-independent mechanisms. **Breast Cancer Research**, v. 19, n. 1, p. 1–18, 2017.

LEMMON, M. A. Domains and Phosphoinositides. **Society**, v. 93, n. 74, p. 81–93, 2007.

LEMMON, M. A.; FERGUSON, K. M. Signal-dependent membrane targeting by pleckstrin homology (PH) domains. **Biochemical Journal**, v. 350, n. 1, p. 1–18, 2000.

LEWIS, R. M. et al. Relationship between placental expression of the imprinted PHLDA2 gene , intrauterine skeletal growth and childhood bone mass ☆. **Bone**, v. 50, n. 1, p. 337–342, 2012.

LI, G. et al. Modulation of ErbB2 blockade in ErbB2-positive cancers: the role of ErbB2 Mutations and PHLDA1. **PloS one**, v. 9, n. 9, p. e106349, jan. 2014a.

LI, X.; YAN, P.; SHAO, Z. Downregulation of miR-193b contributes to enhance urokinase-type plasminogen activator (uPA) expression and tumor progression and invasion in human breast cancer. p. 3937–3948, 2009.

LI, Y. et al. Epigenetic regulation of the pro-apoptosis gene TSSC3 in human

osteosarcoma cells. **Biomedicine and Pharmacotherapy**, v. 68, n. 1, p. 45–50, 2014b.

LI, Y.; MENG, G.; GUO, Q. N. Changes in genomic imprinting and gene expression associated with transformation in a model of human osteosarcoma. **Exp Mol Pathol**, v. 84, n. 3, p. 234–239, 2008.

LIU, H. et al. Long non-coding RNAs as prognostic markers in human breast cancer. **Doi.Org**, v. 7, n. 15, p. 20584–20596, 2016.

LIU, Y. et al. Alcohol intake between menarche and first pregnancy: A prospective study of breast cancer risk. **Journal of the National Cancer Institute**, v. 105, n. 20, p. 1571–1578, 2013.

LOUIE, M. C.; SEVIGNY, M. B. Steroid hormone receptors as prognostic markers in breast cancer. **Am J Cancer Res**, v. 7, n. 8, p. 1617–1636, 2017.

LYU, J. H. et al. Regulation of PHLDA1 Expression by JAK2-ERK1/2-STAT3 Signaling Pathway. **Journal of cellular biochemistry**, v. 117, n. 2, p. 483–90, fev. 2016.

MACIAS, H.; HINCK, L. Mammary gland development. **Wiley Interdisciplinary Reviews: Developmental Biology**, v. 1, n. 4, p. 533–557, 2012.

MARCHIORI, A. C.; CASOLARI, D. A.; NAGAI, M. A. Transcriptional up-regulation of PHLDA1 by 17beta-estradiol in MCF-7 breast cancer cells. **Brazilian journal of medical and biological research = Revista brasileira de pesquisas médicas e biológicas / Sociedade Brasileira de Biofísica ... [et al.]**, v. 41, n. 7, p. 579–82, jul. 2008.

MOAD, A. I. H. et al. Rapamycin induces apoptosis when autophagy is inhibited in T-47D mammary cells and both processes are regulated by Phlda1. **Cell biochemistry and biophysics**, v. 66, n. 3, p. 567–87, jul. 2013.

MODELSKA, A.; QUATTRONE, A.; RE, A. Molecular portraits: The evolution of the concept of transcriptome-based cancer signatures. **Briefings in Bioinformatics**, v. 16, n. 6, p. 1000–1007, 2015.

MONTERISI, S. et al. ScienceDirect Mining cancer gene expression databases for latent information on intronic microRNAs. **Molecular Oncology**, v. 9, n. 2, p. 473–487, 2014.

MURATA, T. et al. Differential susceptibility to hydrogen sul fi de-induced apoptosis between PHLDA1-overexpressing oral cancer cell lines and oral keratinocytes : Role of PHLDA1 as an apoptosis suppressor. v. 20, p. 247–257, 2014.

MUROI, H. et al. Low PHLDA3 expression in oesophageal squamous cell carcinomas is associated with poor prognosis. **Anticancer Research**, v. 35, n. 2, p. 949–954, 2015.

NAGAI, M. A. et al. Down-regulation of PHLDA1 gene expression is associated with breast cancer progression. **Breast Cancer Research and Treatment**, v. 106, n. 1, p.

49–56, 28 set. 2007.

NAGAI M A. et al. Differentially expressed genes and estrogen receptor status in breast cancer. *Int J Oncol*, v 23 p. 1425–1430. 2003.

NEEF, R. et al. Identification of the human PHLDA1/TDAG51 gene: Down-regulation in metastatic melanoma contributes to apoptosis resistance and growth deregulation. **Cancer Research**, v. 62, n. 20, p. 5920–5929, 2002.

NEVINS, J. R.; POTTI, A. Mining gene expression profiles: Expression signatures as cancer phenotypes. **Nature Reviews Genetics**, v. 8, n. 8, p. 601–609, 2007.

OBERST, M. D. et al. TDAG51 is an ERK signaling target that opposes ERK-mediated HME16C mammary epithelial cell transformation. **BMC cancer**, v. 8, p. 189, jan. 2008.

OHKI, R. et al. PHLDA3 is a novel tumor suppressor of pancreatic neuroendocrine tumors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 23, p. E2404-13, 2014a.

OHKI, R. et al. PHLDA3 is a novel tumor suppressor of pancreatic neuroendocrine tumors. **Proceedings of the National Academy of Sciences of the United States of America**, v. 111, n. 23, p. E2404-13, 2014b.

PAPPAS, K. et al. p53 Maintains Baseline Expression of Multiple Tumor Suppressor Genes. **Molecular Cancer Research**, v. 15, n. 8, p. 1051–1062, 2017.

PARK, C. G. et al. A novel gene product that couples TCR signaling to Fas(CD95) expression in activation-induced cell death. **Immunity**, v. 4, n. 6, p. 583–91, jun. 1996.

PARK, E.-S. et al. TDAG51 deficiency promotes oxidative stress-induced apoptosis through the generation of reactive oxygen species in mouse embryonic fibroblasts. **Experimental & molecular medicine**, v. 45, n. 8, p. e35, jan. 2013.

PEROU, C. M. et al. Molecular portraits of human breast tumours. **Nature**, v. 406, n. 6797, p. 747–752, 2000.

PIEDRAHITA, J. A. The Role of Imprinted Genes in Fetal Growth Abnormalities. v. 692, n. June, p. 682–692, 2011.

POLYAK, K. Science in medicine Breast cancer : origins and evolution. **Cell**, v. 117, n. 11, p. 3155–63, 2007.

POLYAK, K.; KALLURI, R. The role of the microenvironment in mammary gland development and cancer. **Cold Spring Harb Perspect Biol**, v. 2, n. 11, p. a003244, 2010.

PUSZTAI, L. Molecular Classification of Breast Cancer: Limitations and Potential. **The Oncologist**, v. 11, n. 8, p. 868–877, 2006.

QIAN, N. et al. The IPL gene on chromosome 11p15.5 is imprinted in humans and mice

and is similar to TDAG51, implicated in Fas expression and apoptosis. **Human Molecular Genetics**, v. 6, n. 12, p. 2021–2029, 1997.

RAMACHANDRAN, S. S. et al. miR-505-5p and miR-193b-3p : potential biomarkers of imatinib response in patients with chronic myeloid leukemia. v. 8194, n. December, 2017.

RAUHALA, H. E. et al. suppressor in prostate cancer. v. 1372, p. 1363–1372, 2010.

RENCK, D. V. . et al. Equity in access to breast cancer screening in a mobile mammography program in southern Rio Grande do Sul state, Brazil [Equidade no acesso ao rastreamento mamográfico do câncer de mama com intervenção de mamógrafo móvel no sul do rio grande do Sul, Brasil. **Cadernos de Saude Publica**, v. 30, n. 1, p. 88–96, 2014.

REYNOLDS, P. Smoking and breast cancer. **Journal of Mammary Gland Biology and Neoplasia**, v. 18, n. 1, p. 15–23, 2013.

RHO, J. et al. TDAG51 is not essential for Fas/CD95 regulation and apoptosis in vivo. **Molecular and cellular biology**, v. 21, n. 24, p. 8365–70, 2001.

RHODES, D. R. et al. ONCOMINE: A Cancer Microarray Database and Integrated Data-Mining Platform. **Neoplasia**, v. 6, n. 1, p. 1–6, 2004.

RINGNER, M. et al. GOBO: Gene expression-based outcome for breast cancer online. **PLoS ONE**, v. 6, n. 3, p. 1–11, 2011.

SAKTHIANANDESWAREN, A. et al. PHLDA1 expression marks the putative epithelial stem cells and contributes to intestinal tumorigenesis. **Cancer Research**, v. 71, n. 10, p. 3709–3719, 2011.

SALAS, M. et al. Placental growth retardation due to loss of imprinting of Phlda2. **Mechanisms of Development**, v. 121, n. 10, p. 1199–1210, 2004.

SAMAVAT, H.; KURZER, M. S. Estrogen metabolism and breast cancer. **Cancer Letters**, v. 356, n. 2, p. 231–243, 2015.

SANTEN, R. et al. Estrogen mediation of breast tumor formation involves estrogen receptor-dependent, as well as independent, genotoxic effects. **Annals of the New York Academy of Sciences**, v. 1155, p. 132–140, 2009.

SAXENA, A. et al. Phosphoinositide binding by the pleckstrin homology domains of Ip1 and Tih1. **Journal of Biological Chemistry**, v. 277, n. 51, p. 49935–49944, 2002.

SAXENA, A. et al. The product of the imprinted gene IPL marks human villous cytotrophoblast and is lost in complete hydatidiform mole. **Placenta**, v. 24, n. 8–9, p. 835–842, 2003.

SCHWIENBACHER, C. et al. Abnormal RNA expression of 11p15 imprinted genes and kidney developmental genes in Wilms' tumor. **Cancer Research**, v. 60, n. 6, p.

1521–1525, 2000.

SELLHEYER, K.; KRAHL, D. PHLDA1 (TDAG51) is a follicular stem cell marker and differentiates between morphoeic basal cell carcinoma and desmoplastic trichoepithelioma. **The British journal of dermatology**, v. 164, n. 1, p. 141–7, jan. 2011.

SIZIOPIKOU, K. P. Ductal carcinoma in situ of the breast: current concepts and future directions. **Archives of pathology & laboratory medicine**, v. 137, n. 4, p. 462–6, 2013.

SORLIE, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proceedings of the National Academy of Sciences of the United States of America**, v. 98, n. 19, p. 10869–74, 2001.

SOUNG, Y. H. et al. Expressional and mutational analysis of PHLDA3 gene in common human cancers. **Human Pathology**, v. 43, n. 6, p. 895–900, 2011a.

SOUNG, Y. H. et al. Expressional and mutational analysis of PHLDA3 gene in common human cancers. **Human Pathology**, v. 43, n. 6, p. 895–900, 2011b.

SUNG et al. HGF mediated upregulation of PHLDA2 is associated with apoptosis in gastric cancer [abstract]. In: Proceedings of the American Association for Cancer Research Annual Meeting 2017; 2017 Apr 1-5; Washington, DC. Philadelphia (PA): AACR; Cancer Res 2017;77(13 Suppl):Abstract nr 523. doi:10.1158/1538-7445.AM2017-523.

TA, M. A. et al. Low PHLDA3 Expression in Oesophageal Squamous Cell Carcinomas Is Associated with Poor Prognosis. v. 954, p. 949–954, 2015.

TAKIKAWA, M.; OHKI, R. A vicious partnership between AKT and PHLDA3 to facilitate neuroendocrine tumors. **Cancer Science**, v. 108, n. 6, p. 1101–1108, 2017.

TANG, K. F. et al. Upregulation of PHLDA2 in Dicer knockdown HEK293 cells. **Biochimica et Biophysica Acta - General Subjects**, v. 1770, n. 5, p. 820–825, 2007.

TANYA, B. Gene Expression Omnibus. **Gse69034**, n. Md, p. 1–10, 2013.

TAYLOR, M. A. et al. TGF- β upregulates miR-181a expression to promote breast cancer metastasis. v. 123, n. 1, 2013.

TREVISANI, F. et al. MicroRNA 193b-3p as a predictive biomarker of chronic kidney disease in patients undergoing radical nephrectomy for renal cell carcinoma. n. November, p. 1343–1350, 2016.

TUNSTER, S. J.; TYCKO, B.; JOHN, R. M. The Imprinted Phlda2 Gene Regulates Extraembryonic Energy Stores □ †. v. 30, n. 1, p. 295–306, 2010.

VERONESI, U. et al. Breast cancer. **Lancet (London, England)**, v. 365, n. 9472, p. 1727–41, 2005.

VUONG, D. et al. Molecular classification of breast cancer. **Virchows Archiv**, v. 465, n. 1, p. 1–14, 2014.

WANG, L. et al. Acta Histochemica MicroRNA-193b inhibits the proliferation , migration and invasion of gastric cancer cells via targeting cyclin D1. **Acta Histochemica**, v. 118, n. 4, p. 323–330, 2016.

WANG, L.; DI, L. J. BRCA1 and estrogen/estrogen receptor in breast cancer: where they interact? **International journal of biological sciences**, v. 10, n. 5, p. 566–575, 2014.

WANG, X. et al. PHLDA2 is a key oncogene-induced negative feedback inhibitor of EGFR / ErbB2 signaling via interference with AKT signaling. p. 1–13, 2015.

WATSON, C. J.; KHALED, W. T. Mammary development in the embryo and adult: a journey of morphogenesis and commitment. **Development (Cambridge, England)**, v. 135, p. 995–1003, 2008.

WEIGEL, M. T.; DOWSETT, M. Current and emerging biomarkers in breast cancer: Prognosis and prediction. **Endocrine-Related Cancer**, v. 17, n. 4, 2010.

WENDY Y. CHEN, MD, MPHBERNARD ROSNER, P. et al. Moderate AlcoholConsumption DuringAdult Life, Drinking Patterns, and Breast Cancer Risk. **JAMA**, p. 1884–1890, 26 dez. 2011.

XU, C. et al. MicroRNA-193b regulates proliferation , migration and invasion in human hepatocellular carcinoma cells. **European Journal of Cancer**, v. 46, n. 15, p. 2828–2836, 2010.

YAN, G.-N.; LV, Y.-F.; GUO, Q.-N. Advances in osteosarcoma stem cell research and opportunities for novel therapeutic targets. **Cancer Letters**, v. 370, n. 2, p. 268–274, 2016.

YERSAL, O.; BARUTCA, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. **World journal of clinical oncology**, v. 5, n. 3, p. 412–24, 2014.

YONG, F. L.; LAW, C. W.; WANG, C. W. Potentiality of a triple microRNA classifier : early detection of colorectal cancer. 2013.

ZEPEDA-CASTILLA. molecular classification of breast cancer. **Cirugía y Cirujanos**, v. 76, p. 87–93, 2008.

ZHANG, L. et al. A circulating miRNA signature as a diagnostic biomarker for non-invasive early detection of breast cancer. **Breast Cancer Research and Treatment**, v. 154, n. 2, p. 423–434, 2015.

ZHANG, Y. et al. MiR-126 and miR-126 * repress recruitment of mesenchymal stem cells and inflammatory monocytes to inhibit breast cancer metastasis. **Nature Cell Biology**, v. 15, n. 3, p. 284–294, 2013.

ZHUO, W. L. et al. Identifying differentially expressed genes and screening small molecule drugs for Lapatinib-resistance of breast cancer by a bioinformatics strategy. **Asian Pacific Journal of Cancer Prevention**, v. 15, n. 24, p. 10847–10853, 2014.

ANEXO A

Análise de expressão de mRNA e de metilação de PHLDA1 por estágio por subtipo 3-gene e por subtipo molecular em tumor de mama. (A) Perfil de expressão de PHLDA1 por estágio de tumor usando dados do banco GOBO (B) Expressão de mRNA de PHDA1 em amostras estratificadas por 3-gene classifier usando dados do banco TCGA (C) correlação entre o perfil de expressão e metilação do gene PHLDA1 usando dados do banco TCGA (D) perfil de metilação por subtipos moleculares do gene PHLDA1 usando o banco de dados TCGA.

ANEXO B

Análise de expressão de mRNA e de metilação de PHLDA2 por estágio por subtipo 3-gene e por subtipo molecular em tumor de mama. (A) Perfil de expressão de PHLDA2 por estágio de tumor usando dados do banco GOBO (B) Expressão de mRNA de PHDA2 em amostras estratificadas por 3-gene classifier usando dados do banco TCGA (C) correlação entre o perfil de expressão e metilação do gene PHLDA2 usando dados do banco TCGA (D) perfil de metilação por subtipos moleculares do gene PHLDA2 usando o banco de dados TCGA.

ANEXO C

Análise de expressão de mRNA e de metilação de PHLDA3 por estágio por subtipo 3-gene e por subtipo molecular em tumor de mama. (A) Perfil de expressão de PHLDA3 por estágio de tumor usando dados do banco GOBO (B) Expressão de mRNA de PHDA3 em amostras estratificadas por 3-gene classifier usando dados do banco TCGA (C) correlação entre o perfil de expressão e metilação do gene PHLDA3 usando dados do banco TCGA (D) perfil de metilação por subtipos moleculares do gene PHLDA3 usando o banco de dados TCGA.