



**Ministério da Saúde  
Instituto Nacional de Câncer  
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER  
Pós-Graduação em Oncologia**

*MARIANA GOMES DA SILVA ARAUJO*

**O PAPEL DE ASXL2 NO CÂNCER DE MAMA**

**Orientadora:** Prof<sup>ª</sup>. Dra. Stephany Cristiane Corrêa  
**Co-Orientadora:** Prof<sup>ª</sup>. Dra. Eliana Saul Furquim Werneck Abdelhay

**RIO DE JANEIRO  
2019**



**Ministério da Saúde  
Instituto Nacional de Câncer  
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER  
Pós-Graduação em Oncologia**

*MARIANA GOMES DA SILVA ARAUJO*

**O PAPEL DE ASXL2 NO CÂNCER DE MAMA**

Dissertação apresentada ao Instituto Nacional de Câncer como parte dos requisitos para obtenção do título de Mestre em Oncologia

**Orientadora:** Prof<sup>ª</sup>. Dra. Stephany Cristiane Corrêa

**Co-Orientadora:** Prof<sup>ª</sup>. Dra. Eliana Saul Furquim Werneck Abdelhay

**RIO DE JANEIRO  
2019**

A658p Araujo, Mariana Gomes da Silva.  
O papel de ASXL2 no câncer de mama / Mariana Gomes da Silva Araujo. –  
Rio de Janeiro, 2019.  
224 f. : il. color.

Dissertação (Mestrado em Oncologia) – Instituto Nacional de Câncer José  
Alencar Gomes da Silva, 2019.

Orientadores: Stephany Cristiane Corrêa; Eliana Saul Furquim Werneck  
Abdelhay.

1. Neoplasias da Mama. 2. Epigenômica. 3. Montagem e Desmontagem da  
Cromatina. 4. ASXL2. I. Corrêa, Sthephany Cristiane (Orient.). II. Abdelhay,  
Eliana Saul Furquim Werneck (Orient.). III. Instituto Nacional de Câncer José  
Alencar Gomes da Silva. IV. Título.

CDD 616.99449042



**Ministério da Saúde  
Instituto Nacional de Câncer  
Coordenação de Pós-graduação**

**INSTITUTO NACIONAL DE CÂNCER  
Pós-Graduação em Oncologia**

*MARIANA GOMES DA SILVA ARAUJO*

**O PAPEL DE ASXL2 NO CÂNCER DE MAMA**

**ORIENTADORA:** Prof<sup>ª</sup>. Dra. Stephany Cristiane Corrêa

**CO-ORIENTADORA:** Prof<sup>ª</sup>. Dra. Eliana Saul Furquim Werneck Abdelhay

**Aprovada em:**

**EXAMINADORES:**

**Prof<sup>ª</sup>. Dra. Claudia Esther Alicia Rocio Hassan** - Presidente

**Prof<sup>ª</sup>. Dra. Sheila Coelho Soares Lima**

**Prof<sup>ª</sup>. Dra. Rosane Silva**

**Prof. Dr. Júlio César Madureira de Freitas Júnior** – Suplente Interno

**Prof<sup>ª</sup>. Dra. Nathalia de Oliveira Meireles da Costa** – Suplente Externa

**RIO DE JANEIRO  
2019**

Ao meu pai, Luiz Augusto Araujo.

## AGRADECIMENTOS

Agradeço a Deus por todas as bênçãos que tive e tenho em minha vida, pelas oportunidades que me foram apresentadas e por ser meu guia nos caminhos em que percorro.

À minha mãe Marina, meu maior exemplo de força e determinação, que sempre me impulsionou para que eu alcançasse meus sonhos, e cuja memória estará para sempre comigo, me motivando a seguir em frente e buscar o melhor para mim. Ao meu pai Luiz Augusto, que me abraçou quando chorei, me acalmou quando me desesperei e sempre acreditou em mim, e ao meu irmão Luiz Henrique.

À minha orientadora, Dra. Stephany Corrêa, por me incentivar, apoiar e motivar desde a Iniciação Científica. Grande parte do meu crescimento, acadêmico e pessoal, se devem a você.

À Dra. Eliana Abdelhay pela oportunidade que me foi dada e por todo o apoio.

À Dra. Renata Binato, por ser sempre solícita quando eu tinha dúvidas e por me auxiliar em diversos experimentos.

Às minhas amigas, Brenda, Juliana, Gabriela, Vanessa, Dayane, Dayara, Mariana Martins, Ana Luiza, Jéssica, Aline, Joana e Mariana Stutz por todos os anos de amizade, pelos conselhos e puxões de orelha, pelas risadas intermináveis e momentos de fuga da realidade. Amo vocês.

Um agradecimento especial à Vanessa, que também segue na vida acadêmica e passa junto comigo pelos perrengues, dificuldades e momentos de felicidade que a pós-graduação oferece. Juntas somos fortes e vamos conseguir viver nossos sonhos.

À minha afilhada Carolina, que me traz luz e felicidade, e que, com sua inocência de criança, sempre me faz lembrar que a vida deve ser levada de uma forma leve.

Aos amigos do LCT, Andressa, Taisnara, Everton e Pedro. Sem vocês, meus dias no laboratório não seriam tão divertidos. Obrigada por ouvirem meus desabafos, me auxiliarem nos experimentos, pelos doces e cafés. Ao Bruno, Gerson e Taísa, por todo o apoio e ajuda.

Agradeço também à Dra. Karina Lani, ao Dr. Matheus Rajão, às Dras. Morgana Castelo Branco e Grasiella Matioszek, e ao Dr. Rafael Lindoso, que me auxiliaram nos experimentos de citometria de fluxo, microscopia e contagem de microvesículas.

A todos da Pós-Graduação em Oncologia do INCA, secretaria, corpo docente e discente, funcionários do CEMO, por contribuírem imensamente para minha formação.

Por fim, agradeço às agências de fomento pelo financiamento que viabilizou a realização deste trabalho.



**Ministério da Saúde  
Instituto Nacional de Câncer  
Coordenação de Pós-graduação**

## **INSTITUTO NACIONAL DE CÂNCER**

### **O PAPEL DE ASXL2 NO CÂNCER DE MAMA**

#### **RESUMO**

#### **DISSERTAÇÃO DE MESTRADO**

**Mariana Gomes da Silva Araujo**

Os *Enhancers of Trithorax and Polycomb* (ETPs) são proteínas capazes de regular efetores de modificações de histona e remodelamento de cromatina, o mecanismo epigenético menos estudado em câncer. Este mecanismo pode envolver diversas proteínas e modificações específicas, causando efeitos variados na expressão gênica, como ativação e silenciamento. Em câncer de mama (CM), o maior causador de mortes por câncer em mulheres no mundo, o papel dos ETPs é ainda menos explorado. O CM pode ser classificado em subtipos moleculares de acordo com a expressão de receptores de membrana para estrógeno, progesterona e HER2, originando, respectivamente, os subtipos Luminal A, Luminal B, e HER2. Na ausência destes receptores, classifica-se como subtipo Triplo Negativo (TN), mais agressivo e de pior prognóstico. Recentemente, a proteína ASXL2, de função ETP descrita em *Drosophila*, foi descrita como potencial reguladora da transcrição de genes alvo do receptor de estrógeno em CM, através de sua interação com modificadores de histona. A fim de avaliar o papel desta proteína no CM, realizamos uma análise inicial da expressão e localização de ASXL2 em linhagens celulares representativas dos subtipos moleculares da doença. Nos modelos MCF-7 e MDA-MB-231, que representam os subtipos Luminal A e TN, há maior expressão de ASXL2, que está localizado também no núcleo celular. Além disso, membros de famílias de modificadores de histona e remodeladores de cromatina tais como Polycomb, Trithorax, SWI/SNF e NURF, foram encontrados diferencialmente expressos e localizados nestas linhagens, e alguns deles, como YY1 e MLL, colocalizados com ASXL2. Para avaliar vias celulares e processos biológicos associados a esta proteína, realizamos seu silenciamento transiente nas linhagens Luminal A e TN, e a identificação dos perfis de expressão gênica global nos modelos silenciados comparados a seus controles mostrou mais de mil genes diferencialmente expressos (DE) em ambos os modelos. Análises *in silico* revelaram fatores de transcrição envolvidos com pluripotência como potenciais reguladores dos genes DE, bem como o enriquecimento de vias de sinalização como Wnt e TGF-beta, associadas com proliferação e migração, dentre outros processos. Alguns dos genes DE foram validados por RT-qPCR, e ensaios funcionais apontaram que o silenciamento de ASXL2 tem impacto na migração celular e formação de microvesículas no modelo TN, ambos processos associados com agressividade no câncer. Potenciais parceiros de ASXL2 foram identificados em ambos os modelos por imunoprecipitação seguida de análise proteômica, e incluem proteínas envolvidas em complexos de silenciamento e ativação da expressão gênica. Após experimento de imunoprecipitação de cromatina, ASXL2 foi encontrado associado a uma marca característica de silenciamento da cromatina (H3K27me3) no modelo Luminal A, indicando uma possível influência na regulação de expressão gênica via remodelamento de cromatina. Assim, de modo geral, observamos que ASXL2 pode ter um papel na regulação da expressão gênica no CM, podendo atuar em diversos níveis, através de sua associação com modificadores de histona e remodeladores de cromatina, bem como na regulação direta ou indireta de genes e vias de sinalização associados com câncer.

**Palavras-chave:** Câncer de mama, Epigenética, Remodelamento de Cromatina, ASXL2.



**Ministério da Saúde  
Instituto Nacional de Câncer  
Coordenação de Pós-graduação**

## **INSTITUTO NACIONAL DE CÂNCER**

### **THE ROLE OF ASXL2 IN BREAST CANCER**

#### **ABSTRACT**

#### **DISSERTAÇÃO DE MESTRADO**

**Mariana Gomes da Silva Araujo**

The Enhancers of Trithorax and Polycomb (ETPs) are proteins capable of regulating histone modifiers and chromatin remodelers, the least studied epigenetic mechanism in cancer. This mechanism can involve different proteins and specific modifications, causing a diversity of effects on gene expression, such as activation and silencing. In breast cancer (BC), the leading cause of death by cancer in women worldwide, the role of the ETPs is even less explored. BC can be stratified into molecular subtypes according to the expression of membrane receptors for estrogen, progesterone and HER2, giving rise, respectively, to the Luminal A, Luminal B and HER2 subtypes. The absence of these receptors characterizes the Triple Negative (TN) subtype, more aggressive and with worse prognosis. Recently, the ASXL2 protein, which has ETP function in *Drosophila*, was described as a potential regulator of the transcription of the estrogen receptor's target genes in BC, through its interaction with histone modifiers. In order to evaluate the role of this protein in BC, we performed an initial analysis of the expression and localization of ASXL2 in cell lines representing the disease's molecular subtypes. In the MCF-7 and MDA-MB-231 models, representing the Luminal A and TN subtypes, there is higher expression of ASXL2, which is also present in the cell nucleus. Furthermore, members of histone modifier and chromatin remodeler families, including Polycomb, Trithorax, SWI/SNF and NURF, were found differentially expressed and localized in these cell lines, and some of them, such as YY1 and MLL1, were colocalized with ASXL2. To evaluate cell pathways and biological processes associated to this protein, we performed its transient silencing in the Luminal A and TN cell lines, and the identification of the global gene expression profiles of the silenced models compared to their controls showed more than a thousand differentially expressed (DE) genes in both models. *In silico* analyses revealed transcription factors involved in pluripotency as potential regulators of the DE genes, as well as an enrichment of signaling pathways such as Wnt and TGF-beta, associated with proliferation and migration, among other processes. Some of the DE genes were validated by RT-qPCR, and functional assays showed that the silencing of *ASXL2* has an impact in cell migration and microvesicle formation in the TN subtype, both processes related to aggressiveness in cancer. *ASXL2*'s potential partners were identified in both models by chromatin immunoprecipitation followed by proteomic analysis, and include proteins involved in gene silencing and gene activation complexes. Following a chromatin immunoprecipitation assay, *ASXL2* was found associated to a histone mark (H3K27me3), characteristic of gene silencing, in the Luminal A subtype, indicating its possible influence in regulating gene expression through chromatin remodeling. Altogether, we observed that *ASXL2* can have a role in regulating gene expression in BC, possibly acting in different levels, through its association with histone modifiers and chromatin remodelers, as well as in the direct or indirect regulation of genes and signaling pathways involved in cancer.

**Keywords:** Breast Cancer, Epigenetics, Chromatin Remodeling, *ASXL2*.



## LISTA DE ABREVIATURAS

ac – acetilação

ACN – acetonitrila

ADH – do inglês *yeast alcohol dehydrogenase*, álcool desidrogenase de levedura

APS - persulfato de amônia

ATP - adenosina trifosfato

BSA – do inglês *bovine serum albumine*, albumina do soro bovino

cDNA - DNA complementar

ChIP - imunoprecipitação de cromatina

CM - Câncer de Mama

CO<sub>2</sub> - gás carbônico

CpG - dinucleotídeo de citosina seguido por guanina

Ct - do inglês *threshold cycle*, ciclo de *threshold*

DAPI - 4',6-diamidino-2-fenilindol

DNA - ácido desoxirribonucleico

DNase - desoxirribonuclease

DNMT - DNA metil-transferase

dNTP - desoxirribonucleotídeo trifosfatado

dsRNA - RNA de fita dupla

DTT- ditioneitol

DZNep - 3-Deazaneplanocin A

EDTA - ácido etilenodiamino tetra-acético

ETP - do inglês *enhancers of trithorax and polycomb*, potencializadores de polycomb e trithorax

FC - do inglês *fold change*

FDR – do inglês *false discovery rate*, taxa de falsos positivos

FISH - do inglês *fluorescence in situ hybridization*, hibridização *in situ* por fluorescência

fw - do inglês *forward*, senso

g - unidade de aceleração ( $g = 9,81 \text{ m/s}^2$ )

h - hora(s)

H2A - histona H2A  
H2B - histona H2B  
H3 - histona H3  
H4 - histona H4  
HAT - histona acetiltransferase  
HCl - ácido clorídrico  
HDAC - histona desacetilase  
HDM - histona demetilase  
HER2 - do inglês *human epidermal growth factor receptor 2*, receptor do fator epidermal humano 2  
HMT - histona metiltransferase  
HRP - do inglês *horseradish peroxidase*  
HS - do inglês *high salt*, alta concentração de sal  
IAA - iodoacetamida  
IgG - Imunoglobulina G  
INCA - Instituto Nacional de Câncer  
K – lisina  
LC-MS - cromatografia líquida-espectrometria de massas  
lncRNA - do inglês *long non-coding RNA*, RNA longo não-codificante  
LMA - leucemia mielóide aguda  
KCl - cloreto de potássio  
me - metilação  
MgCl<sub>2</sub> - cloreto de magnésio  
min - minuto(s)  
miRNA - micro RNA  
mL - mililitro  
mM – milimolar  
NaCl - cloreto de sódio  
Na<sub>3</sub>VO<sub>4</sub> - ortovanadato de sódio

NCBI - do inglês *National Center for Biotechnology Information*, centro nacional de informação em biotecnologia

ncRNA - do inglês *non-coding RNA*, RNA não-codificante

ng – nanograma

NH<sub>4</sub>FA - formiato de amônio

nm - nanômetro

NuRD - complexo *nucleosome remodeling deacetylase*

NURF - do inglês *nucleosome remodeling factor*, fator remodelador do nucleossomo

OMS - Organização Mundial da Saúde

PBS - do inglês *phosphate-buffered saline*, tampão fosfato-salino

Pc - Polycomb

PHD - do inglês *plant homeodomain*

PI - iodeto de propídio

PIC - coquetel inibidor de protease

PCR - do inglês *polymerase chain reaction*, reação em cadeia da polimerase

PRC1 - do inglês *Polycomb repressive complex 1*, complexo repressivo Polycomb 1

PRC2 - do inglês *Polycomb repressive complex 2*, complexo repressivo Polycomb 2

R - arginina

RE - receptor de estrógeno

RIN – do inglês *RNA Integrity Number*, número de integridade de RNA

RNA - ácido ribonucleico

RNAm - RNA mensageiro

RNAr - RNA ribossomal

RNase - ribonuclease

RP - receptor de progesterona

RT-qPCR - do inglês *real time quantitative PCR*, PCR quantitativo em tempo real

rv - do inglês *reverse*, antissenso

S - serina

SBF - soro bovino fetal

SDS - do inglês *sodium dodecyl sulfate*, dodecil sulfato de sódio

SDS-PAGE – do inglês *sodium dodecyl sulfate–polyacrylamide gel electrophoresis*, eletroforese em gel de poliacrilamida–dodecil sulfato de sódio

seg - segundo(s)

siRNA - pequeno RNA de interferência

snoRNA - pequeno RNA nucleolar

SWI/SNF - complexo *SWItch/Sucrose Non-Fermentable*

SUS - Sistema Único de Saúde

TA - temperatura ambiente

TBS-T – do inglês *tris buffered saline with tween*, solução salina tris tamponada com tween

TEMED - tetrametiletilenodiamina

TN - Triplo Negativo

Trx - Trithorax

U - unidade

ub - ubiquitinação

WB - Western Blot

µg - micrograma

µL - microlitro

µM - micromolar

°C - graus Celsius

ΔG - delta G, energia livre de Gibbs

## LISTA DE GENES E PROTEÍNAS

- AKT - do inglês *AKT serine/threonine kinase*, AKT serina/treonina quinase
- ANKRD - do inglês *ankyrin repeat domain*, domínio de repetição de anquirina
- ASH2L - do inglês *ASH2 like histone lysine methyltransferase complex subunit*, *ASH2 like*, subunidade do complexo lisina histona metiltransferase
- ASX - *additional sex combs*
- ASXL - do inglês *additional sex combs like, transcriptional regulator*, adicional *sex combs like*, regulador transcricional
- ARID - do inglês *AT-rich interaction domain*, domínio de interação rico em AT
- ATM - do inglês *ATM serine/threonine kinase*, ATM serina/treonina quinase
- BAP1 - do inglês *BRCA1 associated protein 1*, proteína 1 associada a BRCA
- BAZ2B - do inglês *bromodomain adjacent to zinc finger domain 2B*, bromodomínio adjacente ao domínio dedo de zinco 2B
- BCL11A - *BCL11A, BAF complex component*
- BCOR - do inglês *BCL6 corepressor*, corepressor de BCL6
- BCSG1 - também chamado SNCG, do inglês *synuclein gamma*
- BMI1 - BMI1 proto-oncogene, *polycomb ring finger*
- BMP - *bone morphogenetic protein*
- BPTF - do inglês *bromodomain PHD finger transcription factor*, fator de transcrição bromodomínio PHD *finger*
- BRCA1 - do inglês *BRCA1, DNA repair associated*, BRCA1, associado ao reparo de DNA
- BRCA2 - do inglês *BRCA2, DNA repair associated*, BRCA2, associado ao reparo de DNA
- BRG1 - do inglês *brahma-related gene 1*, gene 1 relacionado a brahma
- BRM - brahma
- CBX - do inglês *chromobox*
- CCND1 - ciclina D1
- CDH1 - do inglês *cadherin 1*, caderina 1
- CDKN1C - do inglês *cyclin dependent kinase inhibitor 1C*, inibidor 1C de quinase dependente de ciclina

CHD3 - do inglês *chromodomain helicase DNA binding protein 3*, proteína de ligação ao DNA cromodomínio helicase 3

CHEK2 - do inglês *checkpoint kinase 2*, quinase de ponto de checagem 2

CTNNB1 - beta-catenina 1

DICER1 - DICER1, ribonuclease III

DKK1 - do inglês *dickkopf WNT signaling pathway inhibitor 1*, inibidor da via de sinalização de WNT *dickkopf* 1

DLL1 - do inglês *delta like canonical Notch ligand 1*, ligante delta like 1 da via canônica de Notch

DNMT - DNA metiltransferase

EED - do inglês *embryonic ectoderm development*, desenvolvimento do ectoderma embrionário

EPC - *enhancer* de polycomb

ESR1 - receptor de estrogênio 1

EZH - do inglês *enhancer of zeste polycomb repressive complex 2 subunit*, subunidade do complexo repressivo polycomb 2, *enhancer* de zeste

FAM208A - do inglês *family with sequence similarity 208, member A*, família com similaridade de sequência 208, membro A

FOXC1 - *forkhead box C1*

FZD - receptor da classe *frizzled*

GAPDH - gliceraldeído-3-fosfato desidrogenase

GATA2 - proteína de ligação a GATA 2

GTF3C4 - fator de transcrição geral IIIC subunidade 4

HBO1 - também chamado KAT7, lisina acetiltransferase 7

HCFC1 - *host cell factor C1*

HDAC2 - histona deacetilase 2

HER2 - do inglês *human epidermal growth factor receptor 2*, receptor do fator epidermal humano 2

HEY2 - *Hes related family BHLH transcription factor with YRPW motif 2*

HNF4A - *hepatocyte nuclear factor 4 alpha*

HOX - do inglês *homeobox*

HP1 - do inglês *heterocromatin protein 1*, proteína 1 da heterocromatina  
*hsp70 - heat shock protein 70*

IL - interleucina

ISWI - do inglês *imitation SWI*, imitação de SWI

ITGB - integrina subunidade beta

JADE - *JADE family PHD finger*

KAT6 - lisina acetiltransferase 6

Ki-67 - do inglês *marker of proliferation Ki-67*, marcador de proliferação Ki-67

KDM5A - lisina demetilase 5A

KMT2 - do inglês *lysine methyltransferase 2*, lisina metiltransferase 2

KREMEN1 - *kringle containing transmembrane protein 1*

LEF1 - *lymphoid enhancer binding factor 1*

LSD1 - também chamada KDM1A, lisina demetilase 1A

MAP1B - proteína associada a microtúbulo 1B

MLL - do inglês *mixed lineage leukemia*, leucemia linhagem mista

MMP - metaloproteinase

MYC - MYC proto-oncogene, fator de transcrição BHLH

NANOG - Nanog *homeobox*

Oct3/4 - *octamer binding transcription factor 3/4*

OTX2 - *orthodenticle homeobox 2*

p16 - proteína supressora de tumor codificada pelo gene CDKN2A, do inglês *cyclin dependent kinase inhibitor 2A*, inibidor de quinase ciclina-dependente 2A

PALB2 - do inglês *partner and localizer of BRCA2*, parceiro e localizador de BRCA2

PCGF - do inglês *polycomb group ring finger*

PHC - do inglês *polyhomeotic homolog*, homólogo polihomeótico

PORCN - porcupina o-aciltransferase

PRICKLE1 - do inglês *prickle planar cell polarity 1*, polaridade celular planar *prickle 1*

PRMT6 - do inglês *protein arginine methyltransferase 6*, proteína arginina metiltransferase 6

RAD51 – RAD51 recombinase

RASSF1 - do inglês *Ras association domain family member 1*, membro 1 da família domínio de associação a Ras

RBBP - do inglês *RB binding protein, chromatin remodeling factor*, proteína de ligação a RB, fator remodelador de cromatina

RBPJ - do inglês *recombination signal binding protein for immunoglobulin kappa J region*

RING1 - do inglês *ring finger protein 1*

RNF2 - do inglês *ring finger protein 2*

RSBN1 - *round spermatid basic protein 1*

RSBN1L - *round spermatid basic protein 1 like*

RUNX - do inglês *runt related transcription factor*, fator de transcrição relacionado a *runt*

SCML - *Scm polycomb group protein like 2*

SIRT1 - do inglês *sirtuin 1*, sirtuína 1

SLUG - também chamada de SNAI2, *snail family transcriptional repressor 2*

SMAD - *mothers against decapentaplegic homolog*

SMARC - do inglês *SWI/SNF related, matrix associated, actin dependent regulator of chromatin*, relacionado a SWI/SNF, associado à matriz, regulador da cromatina dependente de actina

SNAIL - também chamada de SNAI1, *snail family transcriptional repressor 1*

SOST - esclerostina

SOX - *SRY-box*

STK11 - do inglês *serine/threonine kinase 11*, serina/treonina quinase 11

SUZ12 - do inglês *SUZ12 polycomb repressive complex 2 subunit*, subunidade do complexo repressivo polycomb 2, SUZ12

SYT - sinaptotagmina

TCF - *transcription factor*

TDG - timina DNA glicosilase

TET - TET metilcitosina dioxigenase

TGFB - *transforming growth factor beta*

TGFBR - *transforming growth factor beta receptor*

TIMP - *TIMP metallopeptidase inhibitor*

TP53 - do inglês *tumor protein p53*, proteína tumoral p53



TWIST - *Twist family BHLH transcription factor*

UTX - também chamada KDM6A, lisina demetilase 6A

WDR5 - do inglês *WD repeat domain 5*, domínio 5 da repetição WD

WNT - *wingless type*

WWC1 - *WW and C2 domain containing 1*

YY1 - do inglês *YY1 transcription factor*, fator de transcrição YY1

ZMYND11 - *zinc finger MYND-type containing 11*

## LISTA DE FIGURAS

Figura 1.1 - Classificação dos subtipos moleculares de CM .....	4
Figura 1.2 - Modificações pós-traducionais covalentes das caudas de histonas .....	14
Figura 1.3 - Complexos PRC2 e PRC1, da família Polycomb .....	16
Figura 1.4 - Complexo MLL, da família Trithorax .....	17
Figura 1.5 - Níveis de regulação mediados por ETPs .....	22
Figura 5.1 - Expressão de ASXL2 em linhagens celulares de CM .....	50
Figura 5.2 - Localização de ASXL2 em MCF-7 e MDA-MB-231 .....	51
Figura 5.3 - Expressão de membros de PRC2 em linhagens celulares de CM .....	53
Figura 5.4 - Localização de membros de PRC2 em MCF-7 .....	55
Figura 5.5 - Localização de membros de PRC2 em MDA-MB-231 .....	56
Figura 5.6 - Expressão de membros de PRC1 em linhagens celulares de CM .....	57
Figura 5.7 - Localização de membros de PRC1 em MCF-7 e MDA-MB-231 .....	58
Figura 5.8 - Expressão de membros de MLL em linhagens celulares de CM .....	59
Figura 5.9 - Localização de MLL1 em MCF-7 e MDA-MB-231 .....	60
Figura 5.10 - Expressão de membros de NURF em linhagens celulares de CM .....	61
Figura 5.11 - Localização de SMARCA1 em MCF-7 e MDA-MB-231 .....	61
Figura 5.12 - Expressão de membros de SWI/SNF em linhagens celulares de CM .....	62
Figura 5.13 - Expressão de ASXL1 em linhagens celulares de CM .....	63
Figura 5.14 - Colocalização de ASXL2 com YY1 em linhagens celulares de CM .....	64
Figura 5.15 - Colocalização de ASXL2 com MLL1 em linhagens celulares de CM .....	65
Figura 5.16 - Padronização do experimento de silenciamento gênico de ASXL2 .....	67
Figura 5.17 - Silenciamento de ASXL2 após a transfecção com siRNA .....	68
Figura 5.18 - Perfil de expressão gênica global das linhagens MCF-7 e MDA-MB-231 silenciadas para ASXL2 .....	69
Figura 5.19 - Análise comparativa <i>in silico</i> por Diagrama de Venn das linhagens MCF-7 e MDA-MB-231 silenciadas para ASXL2 .....	70
Figura 5.20 - Classificação funcional de genes diferencialmente expressos após o silenciamento de ASXL2 em MCF-7 e MDA-MB-231 .....	70

Figura 5.21 - Mapa de interação dos genes encontrados diminuídos em MCF-7 após o silenciamento de <i>ASXL2</i> .....	72
Figura 5.22 - Mapa de interação dos genes encontrados aumentados em MCF-7 após o silenciamento de <i>ASXL2</i> .....	73
Figura 5.23 - Mapa de interação dos genes encontrados diminuídos em MDA-MB-231 após o silenciamento de <i>ASXL2</i> .....	74
Figura 5.24 - Mapa de interação dos genes encontrados aumentados em MDA-MB-231 após o silenciamento de <i>ASXL2</i> .....	75
Figura 5.25 - Via canônica de Wnt em MCF-7 .....	77
Figura 5.26 - Via de sinalização de Wnt em MDA-MB-231 .....	78
Figura 5.27 - Via de sinalização de TGF-beta nos modelos celulares silenciados .....	79
Figura 5.28 - Expressão gênica relativa de genes envolvidos na sinalização de Wnt após o silenciamento gênico de <i>ASXL2</i> em MCF-7 e MDA-MB-231 .....	81
Figura 5.29 - Expressão gênica relativa de genes envolvidos em vias de sinalização após o silenciamento gênico de <i>ASXL2</i> em MCF-7 e MDA-MB-231 .....	82
Figura 5.30 - Ensaio de <i>wound healing</i> após silenciamento de <i>ASXL2</i> em MCF-7 e MDA-MB-231 .....	84
Figura 5.31 - Avaliação da contagem de microvesículas no sobrenadante das culturas celulares de MCF-7 e MDA-MB-231 após silenciamento de <i>ASXL2</i> .....	85
Figura 5.32 - Avaliação do ciclo celular após silenciamento de <i>ASXL2</i> em MCF-7 e MDA-MB-231 .....	86
Figura 5.33 - Parceiros de <i>ASXL2</i> identificados por abordagem proteômica .....	87
Figura 5.34 - Classificação funcional de parceiros de <i>ASXL2</i> identificados por abordagem proteômica .....	87
Figura 5.35 - Classificação funcional de parceiros de <i>ASXL2</i> associados a DNA e RNA identificados por abordagem proteômica .....	89
Figura 5.36 - Expressão gênica relativa de <i>JADE1</i> e <i>RSBNIL</i> nas linhagens celulares MCF-7 e MDA-MB-231 .....	92
Figura 5.37 - Colocalização entre H3K27me3 e <i>ASXL2</i> em cromatina da linhagem MCF-7 .....	93

## LISTA DE TABELAS

Tabela 1.1 - Tipos de modificação covalente de histonas e exemplos de resíduos modificados .	12
Tabela 1.2 - Terapias com alvo em modificadores de histona .....	26
Tabela 4.1 - Sequências dos <i>primers</i> senso (fw) e antissenso (rv) utilizados neste estudo .....	33
Tabela 5.1 - Fatores de transcrição relacionados com os genes DE nos modelos celulares MCF-7 e MDA-MB-231 após silenciamento de <i>ASXL2</i> .....	76
Tabela 5.2 - Interleucinas identificadas DE nas linhagens celulares MCF-7 e MDA-MB-231 após o silenciamento gênico de <i>ASXL2</i> .....	80
Tabela 5.3 - Parceiros de <i>ASXL2</i> envolvidos em modificação de histonas e remodelamento de cromatina identificados por abordagem proteômica .....	91

## SUMÁRIO

1. INTRODUÇÃO.....	1
1.1 Câncer de mama .....	1
1.1.1 Epidemiologia.....	1
1.1.2 Classificação.....	2
1.1.2.1 Perfil histopatológico.....	2
1.1.2.2 Perfil molecular .....	2
1.1.3 Tratamento.....	5
1.1.4 Etiologia .....	7
1.2 Mecanismos epigenéticos de alteração da expressão gênica.....	9
1.2.1 Metilação do DNA .....	9
1.2.2 RNAs não-codificantes.....	10
1.2.3 Modificações de histona .....	11
1.3 Famílias Polycomb e Trithorax .....	15
1.4 Famílias SWI/SNF e NURF .....	19
1.5 <i>Enhancers</i> de Trithorax e Polycomb .....	21
1.6 Proteínas ASXL.....	22
1.6.1 Família ASXL e câncer .....	24
1.7 Terapias com alvo em modificadores de histona.....	25
2. JUSTIFICATIVA .....	27
3. OBJETIVOS.....	28
3.1 Objetivo Geral .....	28
3.2 Objetivos Específicos .....	28
4. METODOLOGIA.....	29
4.1 Cultivo de linhagens celulares.....	29
4.2 Lista de anticorpos.....	29
4.3 Análise de expressão gênica .....	30
4.3.1 Extração de RNA em coluna .....	30
4.3.2 Reação de digestão com DNase .....	31
4.3.3 Transcrição reversa (RT).....	31
4.3.4 Elaboração dos iniciadores .....	32
4.3.5 Análise da expressão gênica diferencial por PCR quantitativo em Tempo Real .....	33
4.4 Análise de expressão proteica.....	34
4.4.1 Lise celular e extração de proteínas totais .....	34
4.4.2 Lise celular e extração de proteínas de núcleo e citoplasma .....	35
4.4.3 Quantificação de proteínas .....	35
4.4.4 Western Blot.....	36
4.5 Avaliação de localização celular .....	37
4.5.1 Imunofluorescência (IF) .....	37
4.6 Silenciamento gênico de <i>ASXL2</i> .....	38
4.7 Análise do perfil de transcritos por microarranjo.....	39
4.8 Ensaio de migração celular.....	41
4.9 Contagem de microvesículas .....	41
4.10 Ensaio de ciclo celular.....	41
4.11 Imunoprecipitação de proteína (IP).....	42
4.11.1 Lise celular e extração de proteínas totais para imunoprecipitação .....	42

4.11.2	Imunoprecipitação .....	42
4.12	Análise proteômica .....	43
4.12.1	Troca de tampão e concentração das amostras em coluna Amicon .....	43
4.12.2	Digestão enzimática das proteínas .....	44
4.12.3	Cromatografia líquida em nano escala .....	44
4.12.4	Espectrometria de massas .....	45
4.12.5	Identificação de proteínas .....	45
4.12.6	Análise <i>in silico</i> das proteínas identificadas .....	46
4.13	Imunoprecipitação de cromatina (ChIP) .....	46
4.14	Análise estatística .....	48
5.	RESULTADOS .....	49
5.1	ASXL2 está diferencialmente expresso e localizado em modelos <i>in vitro</i> de CM .....	49
5.2	Modificadores de histona e remodeladores de cromatina encontram-se diferencialmente expressos e localizados em modelos <i>in vitro</i> de CM e podem estar colocalizados com ASXL2 .....	51
5.3	Silenciamento gênico de <i>ASXL2</i> .....	66
5.4	ASXL2 tem impacto em vias de pluripotência e nas sinalizações de Wnt e TGF-beta em modelos <i>in vitro</i> de CM .....	68
5.5	ASXL2 associa-se a proteínas envolvidas com modificações de histona e remodelamento de cromatina em modelos <i>in vitro</i> de CM .....	86
5.6	ASXL2 colocaliza-se com uma marca de silenciamento da expressão gênica na linhagem MCF-7 .....	92
6.	DISCUSSÃO .....	94
7.	CONCLUSÕES .....	111
8.	REFERÊNCIAS BIBLIOGRÁFICAS .....	112
9.	ANEXOS .....	129
	Anexo A - Tabela dos genes diferencialmente expressos após silenciamento de <i>ASXL2</i> na linhagem MCF-7 ( <i>fold change</i> 2,0) .....	129
	Anexo B – Tabela dos genes diferencialmente expressos após silenciamento de <i>ASXL2</i> na linhagem MDA-MB-231 ( <i>fold change</i> 2,0) .....	169
	Anexo C - Legenda dos símbolos apresentados nas análises <i>in silico</i> pelo software Metacore® .....	211
	Anexo D - Tabela das proteínas identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MCF-7 .....	213
	Anexo E - Tabela das proteínas identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MDA-MB-231 .....	218
	Anexo F - Tabela das proteínas associadas a DNA e RNA identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MCF-7 .....	222
	Anexo G - Tabela das proteínas associadas a DNA e RNA identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MDA-MB-231 .....	224

# 1. INTRODUÇÃO

## 1.1 Câncer de mama

O câncer de mama (CM) é uma doença complexa e heterogênea, e caracteriza-se pela proliferação anormal, desordenada e maligna de células do tecido mamário, podendo envolver a invasão de células tumorais a outros tecidos e órgãos no corpo, no caso de metástase (NATIONAL CANCER INSTITUTE, 2015).

O estudo da biologia do câncer, incluindo o CM, é de grande importância para a geração de novos conhecimentos acerca do desenvolvimento e progressão da doença. Em 2000, Hanahan e Weinberg apresentaram um conjunto de seis princípios, denominados *hallmarks* do câncer, que teriam um importante papel na transformação de células normais em malignas e no estabelecimento e manutenção do tumor. Os *hallmarks* propostos pelos autores são: 1) manutenção de sinais proliferativos na célula; 2) evasão de sinais supressores do crescimento celular; 3) resistência à morte celular; 4) imortalidade replicativa, ou seja, capacidade de se replicar indefinidamente; 5) indução de angiogênese; e 6) ativação de invasão e metástase (HANAHAN e WEINBERG, 2000).

Em 2011, dois novos *hallmarks* foram apresentados: reprogramação do metabolismo energético para indução do crescimento e divisão celular; e evasão da destruição pelo sistema imune. Além disso, os autores introduziram duas importantes características que possibilitam o surgimento dos *hallmarks* e dão suporte aos mesmos: a instabilidade genômica e a inflamação promotora de tumor (HANAHAN e WEINBERG, 2011).

### *1.1.1 Epidemiologia*

O CM é o tumor mais comum em mulheres, excluindo-se tumores de pele não-melanoma, e também é o principal responsável pelas mortes por câncer em mulheres no mundo (INCA, 2017). Em 2015, cerca de 571 mil mortes foram registradas em decorrência desta doença, segundo estimativas da Organização Mundial da Saúde (OMS, 2018). No Brasil, para o biênio 2018-2019, estimam-se 59.700 novos casos desta doença por ano (INCA, 2017).

### *1.1.2 Classificação*

O CM pode ser classificado em diversos subtipos, de acordo com o seu perfil histopatológico e/ou molecular. Estes subtipos, por sua vez, tem relação direta com fatores como curso clínico, prognóstico e resposta a tratamento (MAKKI, 2015).

#### *1.1.2.1 Perfil histopatológico*

A origem celular do CM e sua capacidade de invasão são dois fatores importantes na classificação histopatológica da doença. O tumor pode se originar de células ductais ou lobulares da mama, e, quanto à invasividade, pode ser denominado carcinoma *in situ*, quando limita-se apenas ao tecido epitelial, ou carcinoma invasivo, quando ocorre a invasão do tecido mamário adjacente (MAKKI, 2015).

Dentre os subtipos histológicos de CM, os mais comuns são os invasivos, sendo o carcinoma ductal invasivo (carcinoma invasivo de tipo não especial, segundo nova nomenclatura da OMS) o mais predominante. O diagnóstico do carcinoma ductal invasivo é feito por exclusão, através da ausência de fatores específicos que caracterizem outros tipos de CM (SINN e KREIPE, 2013). Em seguida, figura o carcinoma lobular invasivo, que geralmente afeta mulheres em idade mais avançada e apresenta diversas variantes que diferem entre si em morfologia celular (SINN e KREIPE, 2013; MAKKI, 2015).

Os tumores invasivos da mama também são classificados em grau histológico, através da avaliação de três parâmetros: formação de túbulos, um indicativo de diferenciação glandular do tecido; pleomorfismo nuclear, ou seja, alterações na forma do núcleo das células tumorais em comparação com células normais do tecido mamário adjacente; e contagem mitótica, um importante indicador de proliferação. A partir da avaliação destas variáveis, o tumor pode ser classificado em grau 1, 2 ou 3, sendo os tumores de grau 1 mais bem diferenciados e de melhor prognóstico, e os tumores de grau 3 pouco diferenciados e geralmente de pior prognóstico (TAVASSOLI e DEVILEE, 2003; VUONG *et al.*, 2014).

#### *1.1.2.2 Perfil molecular*

Atualmente, a nível molecular, o CM é classificado em quatro subtipos (Figura 1.1), que diferem em evolução clínica, agressividade e resposta ao tratamento (LAM *et al.*, 2014; MAKKI, 2015). Os subtipos moleculares de CM foram inicialmente identificados e



caracterizados por Perou *et al.* (2000) através da análise, em larga escala, da expressão gênica de diversos tumores de mama, pela técnica de microarranjo de DNA complementar. A diversidade de padrões de expressão gênica observada nesta análise permitiu a identificação de perfis moleculares característicos de cada tipo de tumor. São estes: Luminal A, Luminal B, HER2 e Triplo Negativo (TN), classificados primariamente com base na expressão diferencial dos receptores de estrógeno (RE), progesterona (RP), e receptor do fator epidermal humano 2 (HER2/neu) (LAM *et al.*, 2014; MAKKI, 2015).

Marcadores acessórios também são utilizados para auxiliar na classificação dos subtipos moleculares, dentre eles Ki-67, um marcador de proliferação, e citoqueratinas, importantes componentes do citoesqueleto. A expressão de citoqueratinas pode variar dentre os subtipos de CM, sendo as citoqueratinas de baixo peso molecular (7, 8, 18 e 19) características de células de fenótipo luminal, e as citoqueratinas de alto peso molecular (5, 6, 14 e 17) características de células de fenótipo basal (ZAHA, 2014).

❖ **Luminal A.** Tumores do tipo Luminal A são positivos para RE e/ou RP e negativos para HER2 (MAKKI, 2015). Apresentam alguma expressão de citoqueratinas de baixo peso molecular e baixa expressão de Ki-67 (<14%), e estão relacionados com bom prognóstico e pouco envolvimento de linfonodos (LAM *et al.*, 2014).

❖ **Luminal B.** O subtipo Luminal B é caracterizado pela expressão de RE e/ou RP e expressão de Ki-67 maior ou igual a 14%, e apresenta pior prognóstico quando comparado com os tumores do tipo A (MAKKI, 2015). Dados do SUS mostram que, no Brasil, de forma semelhante a outras regiões do mundo, há prevalência dos subtipos Luminal (CECILIO *et al.*, 2015).

❖ **HER2.** O subtipo HER2, por sua vez, tem como característica a superexpressão ou amplificação de HER2, associada à não-expressão de RE e RP (MAKKI, 2015). Neste subtipo, a expressão de Ki-67 também é aumentada (>14%) (MAKKI, 2015), indicando alto índice de proliferação, baixa diferenciação do tumor, e, por conseguinte, pior prognóstico (EISENBERG e KOIFMAN, 2001). Pacientes com tumores HER2 apresentam maior risco de reincidência, morte e metástase quando comparados com pacientes com tumores Luminal (LAM *et al.*, 2014).

❖ **Triplo Negativo.** O subtipo Triplo Negativo, também conhecido como Basal, não apresenta expressão de RE, RP e HER2. Por outro lado, há alta expressão de Ki-67 (>14%) e de genes relacionados com células basais do tecido mamário, incluindo diversas citoqueratinas (LAM *et al.*, 2014). O prognóstico é geralmente ruim, podendo haver desenvolvimento de metástases até mesmo nos estágios iniciais da doença (LAM *et al.*, 2014).

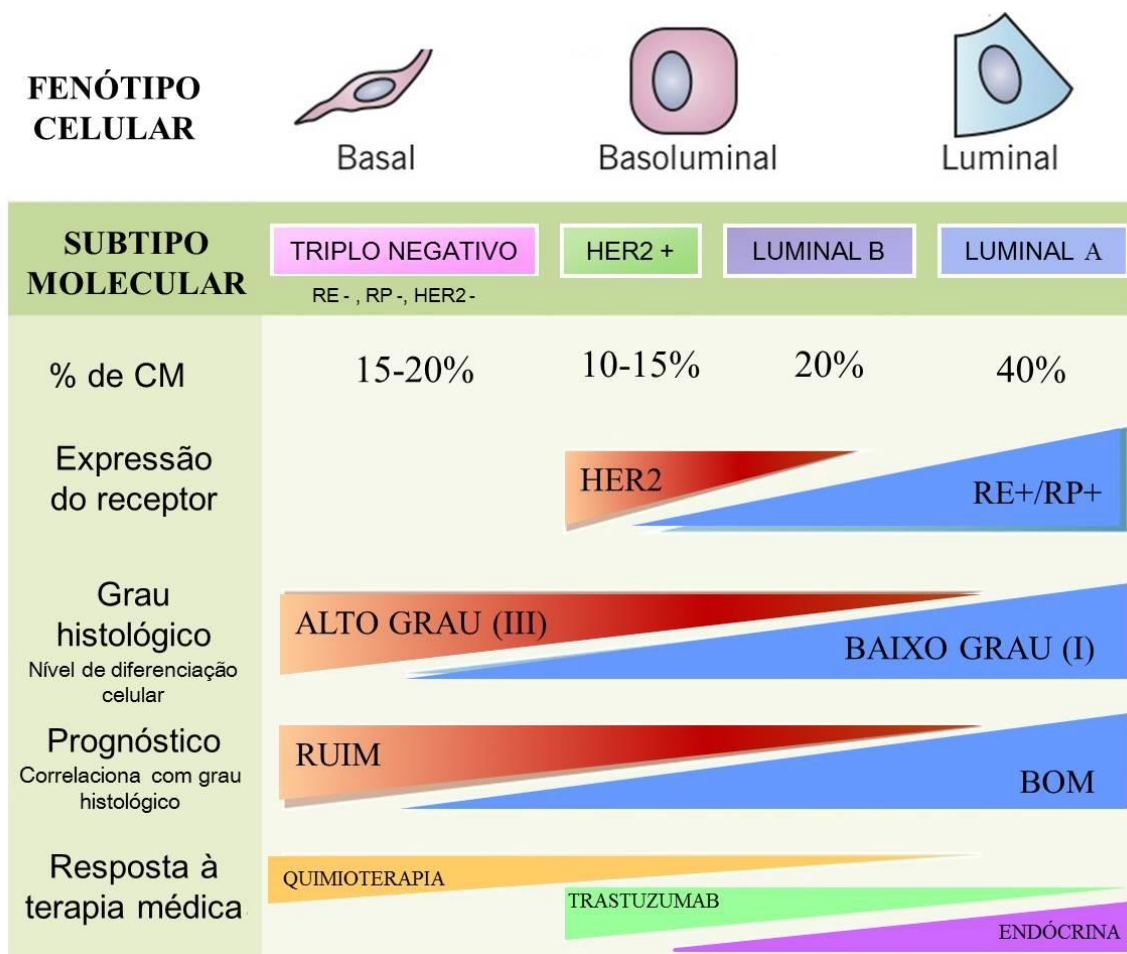


Figura 1.1- Classificação dos subtipos moleculares de CM. Adaptado de Wong e Rebelo, 2012.

Aponta-se ainda um quinto subtipo, positivo para os receptores hormonais RE e/ou RP, e também com superexpressão ou amplificação de HER2, conhecido como Luminal HER. Alguns autores identificam este subtipo como Luminal B com expressão positiva de HER2 (GOLDHIRSCH, 2011).

O diagnóstico diferencial dos subtipos moleculares de CM pode ser feito com o auxílio de técnicas laboratoriais, sendo a imunohistoquímica o método de escolha para a identificação da expressão de RE, RP, Ki-67 e citoqueratinas (ZAHA, 2014; BERSE e LYNCH, 2015). A detecção de HER2 também pode ser realizada por imunohistoquímica, que permite avaliar a superexpressão do receptor na superfície celular, através de classificação em *scores*. Tumores com *scores* 0 e 1+ são considerados negativos para HER2, enquanto tumores 3+ são positivos. No caso de um *score* 2+, faz-se necessária a realização do método de FISH (hibridização *in situ* por fluorescência), que avalia a amplificação do gene *HER2* e permite um diagnóstico conclusivo (ZAHA, 2014; NOUNOU *et al.*, 2015).

O diagnóstico específico, através da avaliação dos marcadores expressos no tumor, permite que o tratamento seja direcionado para o subtipo de CM apresentado, contribuindo para o prognóstico e a eficácia terapêutica (BERSE e LYNCH, 2015).

### 1.1.3 Tratamento

O tratamento do CM pode envolver diversas abordagens, tais como a cirurgia, radioterapia, quimioterapia, e terapias específicas dependendo do subtipo molecular (GOLDHIRSCH, 2011).

A cirurgia é uma forma de intervenção primária no tratamento do CM, permitindo a remoção de células tumorais. Pode consistir na remoção do tumor com preservação do restante da estrutura mamária, em caso de tumores bem delimitados, ou em mastectomia radical, indicada para tumores maiores ou na presença de células invasivas em diferentes quadrantes da mama. Concomitantemente, pode-se realizar cirurgia nos linfonodos axilares, caso haja comprometimento dos mesmos (LESTER, 2015).

As demais terapias podem ser classificadas como neo-adjuvantes, quando utilizadas de forma pré-operatória para diminuição do tumor e, conseqüentemente, da extensão da cirurgia, ou como adjuvantes, sendo realizadas após a cirurgia (NOUNOU *et al.*, 2015).

A radioterapia, assim como a cirurgia, é uma opção de tratamento local do CM, no qual as células são submetidas diretamente a altos níveis de radiação. A radioterapia pode ser utilizada após a cirurgia para reduzir o risco de recorrência da doença, e também pode ser combinada com quimioterapia (LESTER, 2015; NOUNOU *et al.*, 2015).

A quimioterapia, também chamada de terapia citotóxica, promove a destruição de células tumorais, impedindo seu crescimento e multiplicação (NATIONAL CANCER INSTITUTE, 2016). É recomendada principalmente para tumores TN, HER2+ e tumores luminais considerados de alto risco. O tratamento por quimioterapia geralmente envolve a combinação de fármacos, podendo permitir uma melhor resposta e maior sobrevida. Porém, para o estabelecimento do regime de tratamento, deve-se levar em conta fatores como a toxicidade dos fármacos e seus efeitos colaterais. Alguns exemplos de quimioterápicos utilizados para o tratamento do CM são a ciclofosfamida, as antraciclinas (como a doxorrubicina) e os taxanos (como o docetaxel) (CARRICK *et al.*, 2009; BOURDEANU e LIU, 2015; NOUNOU *et al.*, 2015).

A terapia endócrina (hormonoterapia), por sua vez, é indicada para o tratamento dos subtipos nos quais há expressão de receptores hormonais (Luminal A, B e HER, quando

possuir expressão de RE/RP), uma vez que atua inibindo a produção de estrógenos ou modulando seus receptores para bloquear seus efeitos (REINBOLT *et al.*, 2015). O Tamoxifeno (Nolvadex®, AstraZeneca), por exemplo, é um modulador seletivo de receptores de estrógeno prescrito em diversos casos de CM com expressão de RE (GOLDHIRSCH, 2011). Outros exemplos de terapias hormonais são os inibidores de aromatase, como o Letrozol e o Anastrozol, que bloqueiam a síntese de estrógeno, e o Fulvestrant, que diminui a expressão de RE e RP (REINBOLT *et al.*, 2015; MEISEL *et al.*, 2018).

A terapia anti-HER2 consiste principalmente no uso de Trastuzumabe (Herceptin®, Roche), um anticorpo monoclonal que se liga à porção extracelular do receptor HER2, impedindo sua ativação e sinalização (HUDIS, 2007). Esta terapia é indicada em pacientes diagnosticadas com os subtipos Luminal HER e HER2 (GOLDHIRSCH, 2011; NOUNOU *et al.*, 2015), sendo o subtipo HER2 mais sensível (MAKKI, 2015). Outro anticorpo monoclonal específico, o Pertuzumabe, é uma opção mais recente para o tratamento de CM HER2+. Este anticorpo liga-se ao domínio de dimerização de HER2, impedindo a formação de homodímeros e heterodímeros capazes de ativar diversas vias de sinalização relacionadas ao câncer, como as de proliferação e sobrevivência. Além de anticorpos, outras terapias específicas para HER2 podem ser utilizadas, dentre elas os inibidores de tirosina quinase, que atuam inibindo a sinalização por HER2, como o Lapatinibe (Tykerb®, GlaxoSmithKline) (ESCRIVÁ-DE-ROMANÍ *et al.*, 2018; MEISEL *et al.*, 2018; PERNAS, BARROSO-SOUSA, TOLANEY, 2018).

A combinação de terapias é uma abordagem muito utilizada no tratamento do CM, a fim de aumentar a sobrevida das pacientes. Entretanto, deve-se levar em conta a individualidade de cada paciente, bem como os possíveis efeitos adversos e a toxicidade associada ao uso de combinações, avaliando-se seus riscos e benefícios de acordo com cada caso (MEISEL *et al.*, 2018; PERNAS, BARROSO-SOUSA, TOLANEY, 2018).

Dentre os desafios no tratamento do CM, figuram a grande heterogeneidade desta doença e o surgimento de resistência às terapias aplicadas atualmente. Um dos focos no estudo do CM é a busca por novos alvos específicos, através do estudo de diversas vias e processos celulares que possam contribuir para o estabelecimento e progressão do CM (MASOUD e PAGES, 2017).

### *1.1.4 Etiologia*

O conhecimento sobre as possíveis causas e fatores de risco associados ao CM pode contribuir para a prevenção, o diagnóstico precoce, e conseqüentemente, para um melhor prognóstico e direcionamento ao tratamento adequado da doença (INCA, 2015). O desenvolvimento de CM pode estar relacionado com fatores ambientais, genéticos e epigenéticos.

#### **❖ Fatores de risco ambientais**

Dentre os fatores de risco ambientais que podem contribuir para o surgimento da doença, estão a exposição à radiação ionizante, o consumo de álcool, a obesidade e o sedentarismo (INCA, 2017; SINGLETARY, 2003). Hábitos como alimentação saudável, prática de exercícios físicos e amamentação contribuem para um menor risco de desenvolvimento da doença (INCA, 2015).

Um dos fatores ambientais mais bem documentados para o desenvolvimento de CM é a idade. A incidência da doença é baixa antes dos 30 anos (SINGLETARY, 2003) e aumenta rapidamente até os 50 (INCA, 2015). Após os 50, o aumento na incidência é mais lento, o que está relacionado com outro fator de risco para o desenvolvimento de CM, a exposição ao estrógeno e/ou progesterona, que diminui conforme se atinge a menopausa (INCA, 2015). O risco de desenvolver CM é maior em mulheres que, no decorrer de sua vida, são mais expostas aos hormônios femininos, ou seja, naquelas que tiveram sua primeira menstruação mais cedo, entraram na menopausa mais tardiamente, ou em mulheres que façam terapia de reposição hormonal para prevenir a osteoporose ou diminuir os sintomas da menopausa (SINGLETARY, 2003).

#### **❖ Fatores de risco genéticos**

Além dos fatores ambientais, sabe-se que o desenvolvimento de CM pode estar relacionado com componente genético, que pode ser de origem familiar (hereditária) ou esporádica.

Cerca de 10% dos casos de CM estão relacionados com origem hereditária, e o histórico familiar é de grande importância em seu surgimento precoce (INCA, 2015; SKOL, SASAKI e ONEL, 2016). O CM hereditário é caracterizado por mutações germinativas em genes de susceptibilidade, conferindo um maior risco de desenvolvimento da doença

(KENEMANS, VERSTRAETEN e VERHEIJEN, 2008). Diversas alterações genéticas que desempenham um papel no CM familiar já são conhecidas, como nos genes *BRCA*.

As mutações nos genes *BRCA1* e *BRCA2*, supressores de tumor, foram as primeiras a serem descobertas e associadas com o CM hereditário, e famílias que apresentam tais alterações são geralmente caracterizadas por diversos indivíduos afetados. Apesar de serem pouco comuns na população (prevalência de cerca de 0,1%), as mutações em *BRCA1/2*, herdadas de forma autossômica dominante, são de alta penetrância (ROUSSET-JABLONSKI e GOMPEL, 2017). A mutação no gene *BRCA1* implica em 40 a 87% de chance de desenvolvimento de CM, enquanto para mutações no gene *BRCA2* este risco está calculado em 18 a 88% (ENGEL e FISHER, 2015).

Portadores de mutação em *BRCA1/2* também possuem risco aumentado para o desenvolvimento de outros tipos de câncer, como o câncer de ovário (22 a 65% de risco para mutações em *BRCA1* e 10 a 35% para *BRCA2*) e o câncer pancreático. Homens portadores de alterações genéticas em *BRCA1/2* tem risco aumentado de desenvolver câncer de próstata e câncer de mama masculino (ENGEL e FISHER, 2015; ROUSSET-JABLONSKI e GOMPEL, 2017).

Além dos genes *BRCA*, mutações germinativas em outros genes também estão associadas ao CM hereditário. Mutações no supressor de tumor *TP53*, por exemplo, levam a uma susceptibilidade aumentada a vários tipos de câncer, incluindo CM, podendo aumentar em até 60 vezes a chance de desenvolvimento da doença (ALORAIFI *et al.*, 2015). Também estão descritas mutações em *PALB2*, gene que codifica uma proteína parceira de *BRCA2*; *CDH1*, que possui função de supressor de tumor; bem como em genes que codificam quinases envolvidas no reparo ao dano de DNA, como *ATM*, *CHEK2* e *STK11* (ALORAIFI *et al.*, 2015; SCALIA-WILBUR *et al.*, 2016; WINSHIP e SOUTHEY, 2016).

O conhecimento do histórico familiar e da predisposição genética para o desenvolvimento de CM pode contribuir para que diversas medidas de prevenção e controle sejam tomadas, incluindo o aumento da frequência de realização de exames de imagem e até mesmo cirurgia profilática de remoção das mamas (ENGEL e FISHER, 2015).

O CM de origem esporádica, por sua vez, surge através de mutações adquiridas em genes somáticos, e representa a maioria dos casos de CM (ROMAGNOLO *et al.*, 2016). Dentre as possíveis alterações genéticas envolvidas no surgimento de CM esporádico, figuram as mutações esporádicas em oncogenes como *MYC*, *CCND1* e *HER2*. Estas mutações podem levar à ativação destes genes e contribuir para o desenvolvimento da doença, podendo também estar associadas à inativação de genes supressores de tumor e a outras mutações

independentes (KENEMANS, VERSTRAETEN e VERHEIJEN, 2008; LEE e MULLER, 2010).

O CM esporádico resultante da inativação de *BRCA1* e/ou *BRCA2* também é possível quando ocorrem mutações nas duas cópias do gene, porém é pouco comum (KENEMANS, VERSTRAETEN e VERHEIJEN, 2008; LEE e MULLER, 2010).

#### ❖ Fatores de risco epigenéticos

A complexidade do CM vai além dos fatores de risco genéticos e ambientais, podendo incluir também o acúmulo de alterações epigenéticas que interferem na manutenção das atividades normais da célula (ZIOGAS, 2012). As mudanças epigenéticas diferem das genéticas por serem capazes de promover alterações na expressão de genes sem envolver modificações na sequência de DNA. As alterações epigenéticas também podem ser herdadas, contribuindo para a manutenção da memória epigenética, e, na patogênica do câncer, podem levar ao silenciamento de genes supressores de tumor e à ativação de oncogenes, o que as caracteriza como variáveis importantes no desenvolvimento e progressão da doença (ALLIS *et al.*, 2007; BASSE e AROCK, 2014).

### 1.2 Mecanismos epigenéticos de alteração da expressão gênica

Três grandes mecanismos de alteração epigenética são conhecidos: metilação do DNA, ação de RNAs não-codificantes e modificação de histonas/remodelamento de cromatina.

#### *1.2.1 Metilação do DNA*

A metilação do DNA ocorre através da adição de um grupamento metil em resíduos de citosina no DNA, e sua consequência mais conhecida é o silenciamento da expressão gênica. Os mecanismos pelos quais a metilação do DNA leva ao silenciamento gênico são a formação de uma barreira física que impede a ligação de proteínas promotoras da transcrição, o recrutamento de complexos inibitórios que se ligam à modificação, ou até mesmo a interação com modificações de histona (ALLIS *et al.*, 2007; PASKA e HUDLER, 2015).

Esta modificação ocorre em dinucleotídeos de citosina seguido de guanina, presentes em abundância nas chamadas ilhas CpG. As ilhas CpG são caracterizadas por possuir uma alta densidade destes dinucleotídeos, e são encontradas principalmente em regiões promotoras

de genes, podendo também estar presentes no corpo dos genes e em regiões intergênicas (JONES, 2012). A metilação de citosinas é promovida pelas enzimas DNA metil-transferases (DNMTs), e a remoção desta marca de silenciamento é catalisada principalmente pelas enzimas TET (JONES, 2012; DOR e CEDAR, 2018).

No contexto do câncer, genes como os supressores de tumor e metástase ou os de reparo de DNA podem estar anormalmente hipermetilados, e portanto silenciados, enquanto outros, como os oncogenes, podem se apresentar hipometilados, ou seja, ativados (JOVANOVIC *et al.*, 2010; BASSE e AROCK, 2014). Além disso, a hipometilação global de seqüências repetitivas do DNA pode possuir um papel na tumorigênese, pois pode levar à instabilidade genômica, uma importante característica de diversos tipos de câncer (ALLIS *et al.*, 2007; EHRLICH e LACEY, 2013).

No CM, alguns genes diferencialmente metilados já foram identificados. Genes de reparo, como o *BRCA1*, genes pró-apoptóticos, como *HOXA5*, e genes de inibição do ciclo celular, como p16 e *RASSF1A*, podem estar hipermetilados na doença (BASSE e AROCK, 2014). Entretanto, genes de invasão e metástase, tais como *BCSG1*, *CDH1* e *CHD3*, podem ser encontrados hipometilados em CM (JOVANOVIC *et al.*, 2010). De acordo com Basse e Arock (2014), a hipermetilação de *BRCA1* e a mutação deste gene levam ao mesmo padrão de expressão gênica, indicando que uma alteração genética não precisa estar necessariamente presente para que a doença se desenvolva.

Além disso, a hipometilação de regiões satélite centroméricas e pericentroméricas já foi identificada em CM, também contribuindo para a instabilidade genômica e para o desenvolvimento da doença (EHRLICH e LACEY, 2013).

### 1.2.2 RNAs não-codificantes

Os RNAs não-codificantes (ncRNAs) são um grupo de moléculas não traduzíveis de RNA que atuam na regulação da expressão de genes em diversos níveis, que vão da transcrição ao *splicing*, tradução e manutenção da organização e estabilidade genômica (CECH e STEITZ, 2014). Estas moléculas podem ser classificadas de acordo com seu tamanho em pequenos ncRNAs (19 a 31 nucleotídeos), ncRNAs de tamanho médio (20 a 200 nucleotídeos), e ncRNAs longos (>200 nucleotídeos) (LIN e HE, 2017).

Dentre os pequenos ncRNAs, figuram os pequenos RNAs de interferência (siRNAs, cerca de 22 nucleotídeos), capazes de induzir a clivagem de alvos complementares de RNA; os microRNAs (miRNAs, cerca de 22 nucleotídeos), capazes de modular a expressão de RNA



mensageiro (RNAm); dentre outros. Um exemplo de ncRNA de tamanho médio são os pequenos RNAs nucleolares (snoRNAs, cerca de 70 nucleotídeos), que desempenham função no processamento de RNA ribossomal (RNAr). Os RNAs longos não codificantes (lncRNAs, >200 nucleotídeos), por sua vez, podem desempenhar diversas funções nas células, dentre elas transporte núcleo-citoplasma, regulação pós-transcricional, organização da cromatina, modificações epigenéticas, entre outras (CECH e STEITZ, 2014; LIN e HE, 2017). Atualmente, os ncRNAs mais estudados são os miRNAs e os lncRNAs.

Os miRNAs atuam, principalmente, reprimindo a expressão gênica através da ligação a um RNAm alvo, seguida de degradação ou interrupção do processo de tradução do mesmo (BASSE e AROCK, 2014; LIN e HE, 2017). A expressão de miRNAs pode estar aumentada ou diminuída em diversos tipos de câncer, e, dependendo de seus alvos, os miRNAs podem agir como oncogenes (oncomiRs) ou supressores de tumor (ts-miRNAs). A expressão diferencial de miRNAs no câncer tem sido alvo de estudos, com possível utilização como biomarcadores de diagnóstico e prognóstico ou alvos terapêuticos (LIN e HE, 2017; KHORDADMEHR *et al.*, 2018).

Os lncRNAs podem ter um papel de ativação ou repressão de genes alvo, e agem por meio de diversos mecanismos, como co-ativação de fatores de transcrição, interferência nas fases de iniciação, alongamento e terminação da transcrição, entre outros (DE LEENEER e CLAES, 2015). Alterações na expressão de lncRNAs também podem estar relacionadas com o câncer, uma vez que estas moléculas podem ter um papel em diversos processos celulares envolvidos com a doença, como regulação de ciclo celular, resposta ao dano de DNA, apoptose, metástase, entre outros (DE LEENEER e CLAES, 2015; LIN e HE, 2017).

### *1.2.3 Modificações de histona*

As histonas, proteínas que promovem a estruturação e compactação da molécula de DNA, podem sofrer modificações pós-traducionais que levam a uma alteração na acessibilidade do maquinário de transcrição ao DNA. Dependendo da modificação sofrida, a cromatina pode ser compactada ou descompactada, resultando respectivamente em repressão ou ativação transcricional (ALLIS *et al.*, 2007; JOVANOVIC *et al.*, 2010; AUDIA e CAMPBELL, 2016).

Para promover a organização da cromatina, as proteínas histonas, carregadas positivamente, se reúnem em octâmeros (compostos por um tetrâmero H3-H4 e dois dímeros H2A-H2B), formando o nucleossomo. A esta estrutura, uma dupla fita de DNA de cerca de

147 pares de bases se enovela, e, a partir desta associação DNA-nucleossomo, ocorre a compactação inicial do material genético (ALLIS *et al.*, 2007).

Modificações nas caudas das histonas, promovidas por enzimas que podem agir individualmente ou em complexos, são covalentes e reversíveis, tornando o processo de ativação/repressão gênica dinâmico, e podem envolver a adição ou remoção de grupamentos ou peptídeos nos resíduos de aminoácidos das caudas. Dentre as modificações de histonas mais descritas, estão a acetilação, metilação, fosforilação e ubiquitinação, dentre outras (Tabela 1.1). O tipo de modificação sofrido e o resíduo de aminoácido atingido determinarão a ativação ou a repressão da expressão gênica do alvo. (ALLIS *et al.*, 2007; BANNISTER e KOUZARIDES, 2011).

**Tabela 1.1- Tipos de modificação covalente de histonas e exemplos de resíduos modificados.**

Papel na transcrição		Sítios de modificação nas histonas
Acetilação	ativação	H3 (K9, K14, K18, K56) H4 (K5, K8, K12, K16) H2A H2B (K6, K7, K16, K17)
Fosforilação	ativação	H3 (S10)
Metilação	ativação	H3 (K4, K36, K79)
	repressão	H3 (K9, K27) H4 (K20)
Ubiquitinação	ativação	H2B (K123)
	repressão	H2A (K119)

Adaptado de Allis *et al.*, 2007.

**Acetilação.** A acetilação de histonas ocorre em resíduos de lisina, e resulta na descompactação da cromatina, uma vez que promove a diminuição da basicidade das histonas e, conseqüentemente, enfraquece sua interação com o DNA. A desacetilação, por sua vez, promove o efeito contrário, compactando a cromatina e promovendo silenciamento da expressão gênica (BASSE e AROCK, 2014). As enzimas que modificam o estado de acetilação das histonas são as histona acetiltransferases (HATs), que adicionam grupamentos acetil aos resíduos de lisina, e as histona deacetilases (HDACs), que removem estes grupamentos (BANNISTER e KOUZARIDES, 2011; AUDIA e CAMPBELL, 2016).

**Metilação.** A metilação, por sua vez, pode ocorrer tanto em resíduos de lisina quanto em resíduos de arginina, sendo catalisada por histona metiltransferases (HMTs) e removida por histona demetilases (HDMs). Esta modificação está mais frequentemente presente nas histonas H3 e H4. Diferente da acetilação, que está relacionada com aumento da expressão de

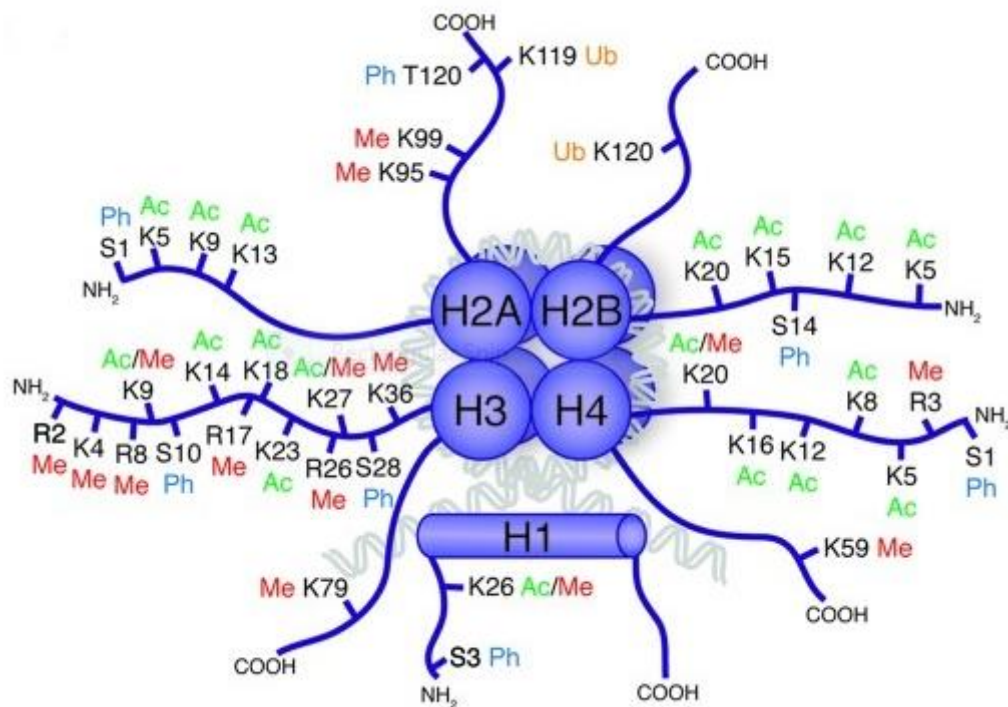
genes, a metilação pode promover tanto ativação quanto repressão da expressão gênica, dependendo da posição do resíduo modificado na cauda. Outro fator complicador reside na possibilidade de adição de até três grupos metil em um mesmo resíduo de lisina (mono, di ou trimetilação), e de até dois grupos metil nos resíduos de arginina. Atualmente, diversos sítios de metilação são conhecidos quanto ao número de modificações sofridas e quanto à sua função na alteração da expressão gênica (BANNISTER e KOUZARIDES, 2011; D'OTO *et al.*, 2016).

**Fosforilação.** A fosforilação de histonas é catalisada pelas enzimas quinases e revertida pelas fosfatases, e está geralmente relacionada com ativação da expressão gênica, mas pode promover condensação cromossomal quando ocorre em resíduos específicos. Os resíduos de serina e treonina são os alvos deste tipo de modificação pós-traducional (BANNISTER e KOUZARIDES, 2011).

**Ubiquitinação.** A ubiquitinação consiste na adição da proteína ubiquitina às caudas de histonas, e difere do mecanismo de ubiquitinação para degradação de proteínas. Em histonas, esta alteração pode promover ativação ou repressão da expressão gênica, dependendo do resíduo modificado, tal como a metilação (ALLIS *et al.*, 2007).

As modificações de histonas podem ter um papel de sinalização para proteínas capazes de regular a expressão gênica e/ou modificar a cromatina. Um exemplo é a proteína HP1, capaz de se ligar à histona H3 metilada no resíduo 9 de lisina (H3K9me) e promover a formação de heterocromatina (CAMPBELL e TURNER, 2013).

O conjunto destas possíveis modificações compõe o **código de histonas**, que pode contribuir para determinar o estado epigenético da cromatina e sua atividade em genes específicos, através das modificações em si e da ligação de proteínas efetoras às mesmas (JENUWEIN e ALLIS, 2001; ALLIS, 2007). Exemplos de modificações covalentes das caudas de histonas podem ser observados na Figura 1.2.



**Figura 1.2 - Modificações pós-traducionais covalentes das caudas de histonas:** metilação (Me), fosforilação (Ph), acetilação (Ac) e ubiquitinação (Ub). Adaptado de Tollervy e Lunnyak, 2012.

Em câncer, diversas modificações de histonas já foram descritas, como a trimetilação da lisina 27 na histona H3 (H3K27me3), que promove repressão da expressão gênica e está associada com mau prognóstico em câncer de esôfago e um melhor prognóstico em câncer de próstata. Outra modificação presente globalmente em células tumorais é a perda da acetilação na lisina 16 da histona 4 (H4K16), causada pela diminuição da expressão de HATs, juntamente com o aumento da expressão de deacetilases (CAMPBELL e TURNER, 2013).

Em CM, por exemplo, o aumento da ação de HDACs na região promotora do gene de RE (receptor de estrogênio), associado à metilação do DNA na mesma área, determina o silenciamento deste gene e caracteriza o tumor como RE negativo (BASSE e AROCK, 2014).

Segundo Campbell e Turner (2013), a compreensão do papel das modificações de histonas no desenvolvimento de câncer depende do estudo do recrutamento e atividade das proteínas promotoras das alterações. Dentre os diversos grupos de modificadores de histona já descobertos e descritos, figuram as famílias de proteínas Polycomb e Trithorax, de funções antagônicas. O estudo destas e de outras proteínas capazes de alterar a compactação da cromatina pode facilitar o entendimento da relação entre alterações epigenéticas e câncer.

### 1.3 Famílias Polycomb e Trithorax

As proteínas das famílias Polycomb (Pc) e Trithorax (Trx) são exemplos de modificadores de histona e remodeladores de cromatina que atuam em complexos para promover mudanças estruturais na cromatina (STEFFEN e RINGROSE, 2014). As famílias Pc e Trx possuem funções antagônicas de silenciamento e ativação da expressão gênica, respectivamente, e foram inicialmente descobertas em *Drosophila melanogaster*, na regulação da expressão dos genes *Homeobox* (HOX) (SCHUETTENGRUBER *et al.*, 2011).

Em humanos, proteínas com funções análogas às dessas famílias já são conhecidas, e, além de atuarem nos genes *HOX*, também são capazes de regular a expressão gênica em diversos outros *loci* do genoma (SCHUETTENGRUBER *et al.*, 2011; STEFFEN e RINGROSE, 2014).

Juntas, as famílias Pc e Trx regulam o estado da cromatina e a expressão gênica em vários tecidos, contribuem para a manutenção da memória epigenética das células através de todo o ciclo celular, e por conseguinte, promovem a preservação da memória, identidade e função celulares (MILLS, 2010; CORADINI e ORIANA, 2014; STEFFEN e RINGROSE, 2014). Por serem atuantes na manutenção da identidade de células-tronco, na proliferação celular e no desenvolvimento, disfunções na atividade destas proteínas podem levar ao câncer (SCHUETTENGRUBER *et al.*, 2007; MILLS, 2010).

#### ❖ **Família Polycomb**

As proteínas da família Pc atuam silenciando genes através da ação dos complexos PRC1 e PRC2 (*Polycomb repressive complex 1 e 2*, respectivamente), que promovem, através da alteração de resíduos nas caudas de histonas, a compactação da cromatina (Figura 1.3).

O complexo PRC2 é responsável pela mono, di ou trimetilação da lisina 27 na histona H3 (H3K27me1, H3K27me2 e H3K27me3), e seus principais componentes são uma proteína EZH (EZH1 ou EZH2), uma proteína RBBP (RBBP4 ou RBBP7), e as proteínas EED e SUZ12, podendo haver também a associação de outras proteínas. A atividade de metiltransferase de PRC2 é garantida pelas proteínas efetoras EZH (*Enhancer of Zeste*), que, segundo Blackledge, Rose e Klose (2015), possuem baixa atividade enzimática individual *in vitro*, mas são altamente ativas quando presentes no complexo. As outras proteínas deste complexo garantem a estabilidade e integridade do mesmo, bem como sua ligação à cromatina. As subunidades EED e SUZ12, por exemplo, atuam como *enhancers* da atividade

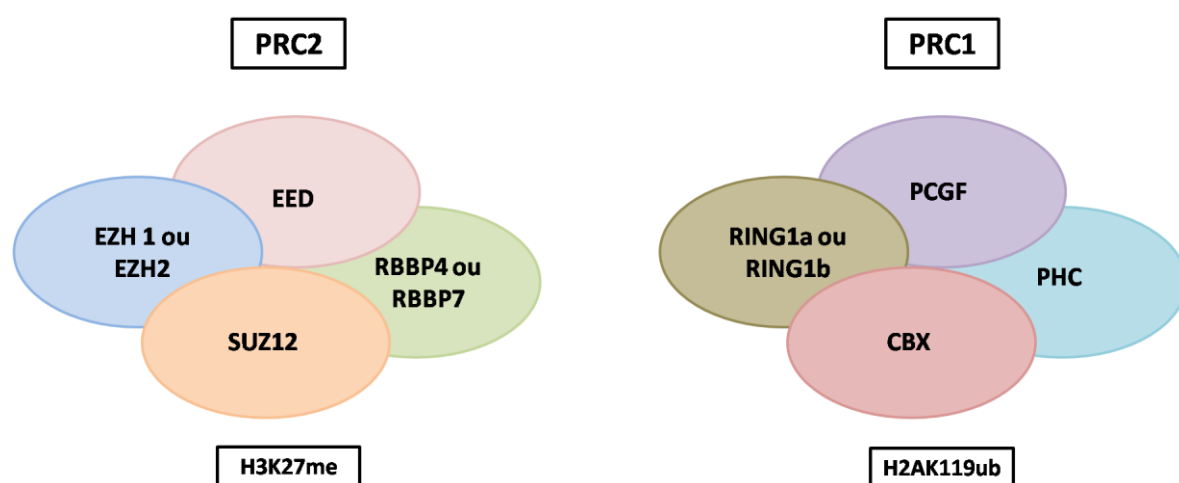
enzimática das proteínas efetoras (BLACKLEDGE, ROSE e KLOSE, 2015; STEFFEN e RINGROSE, 2014; SCELFO, PIUNTI e PASINI, 2015).

O complexo PRC2 pode agir no recrutamento de HDACs, que removem grupamentos acetil das caudas de histonas, reforçando o silenciamento gênico promovido pelo mesmo. Dentre as HDACs recrutadas estão as proteínas SIRT1 e HDAC2 (MILLS, 2010).

O complexo PRC1, por sua vez, promove a ubiquitinação do resíduo de lisina 119 na histona H2A (H2AK119ub). É composto por uma subunidade RING1 (RING1A) ou RNF2 (RING1B), responsável por catalisar a reação de ubiquitinação, uma proteína PCGF (PCGF 1-6), proteínas CBX (capazes de reconhecer e se ligar a H3K27me3), proteínas PHC (PHC 1-3), de função pouco conhecida, e outras proteínas acessórias (como a YY1) (STEFFEN e RINGROSE, 2014; SCELFO, PIUNTI E PASINI, 2015; POYNTER e KADOCH, 2016).

As proteínas PCGF participam da regulação da atividade enzimática do complexo e da sua ligação à cromatina (BLACKLEDGE, ROSE e KLOSE, 2015). Uma destas proteínas, a PCGF4 (também chamada BMI1), é capaz de se ligar ao DNA e deve estar presente no complexo para que a reação de ubiquitinação ocorra (POYNTER e KADOCH, 2016).

Os mecanismos exatos de repressão da expressão gênica pela família Polycomb ainda não foram totalmente elucidados, mas sabe-se que os complexos PRC1 e PRC2 podem estar co-localizados em alvos no genoma e possivelmente recrutam e reconhecem um ao outro para promover silenciamento gênico (BLACKLEDGE, ROSE e KLOSE, 2015). A interação com os complexos Polycomb torna a cromatina menos acessível a fatores de transcrição, e reduz a atividade da enzima RNA polimerase II, que catalisa a transcrição (MILLS, 2010; SCELFO, PIUNTI e PASINI, 2015).



**Figura 1.3 - Complexos PRC2 e PRC1, da família Polycomb.** Adaptado de Blackledge, Rose e Klose, 2015; e Onodera e Nakayama, 2015.

### ❖ Família Trithorax

A família Trx tem função oposta a Pc, promovendo a abertura da estrutura da cromatina e levando à ativação da expressão gênica (SCHUETTENGRUBER *et al.*, 2011). As proteínas Trx são menos conhecidas e caracterizadas quando comparadas com as da família Pc, mas, de forma semelhante aos seus antagonistas, atuam em complexos compostos por diferentes proteínas, promovendo modificações nas histonas e o remodelamento da cromatina (MILLS, 2010).

O complexo MLL (Figura 1.4) promove a trimetilação da lisina 4 na histona H3 (H3K4me3), uma modificação que leva à ativação da expressão de genes. Este complexo pode ser formado, em humanos, por diversas proteínas, como a metiltransferase MLL1 (também chamada KMT2A), de função efetora; ASH2L, que contribui para a ligação do complexo MLL aos seus alvos na cromatina; WDR5, capaz de reconhecer e se ligar a H3K4 metilada ou não metilada; RBBP5, dentre outras (SCHUETTENGRUBER *et al.*, 2011).

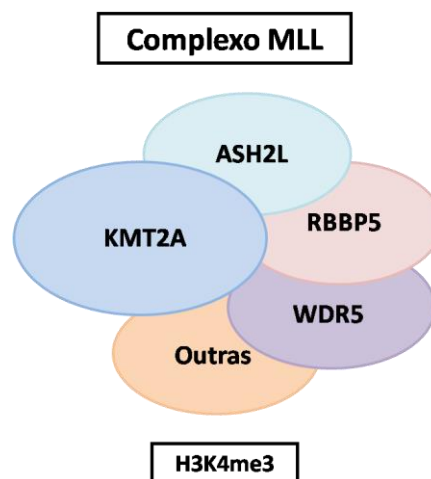


Figura 1.4 - Complexo MLL, da família Trithorax. Adaptado de Onodera e Nakayama, 2015.

### ❖ Famílias Pc / Trx e câncer

A regulação da transcrição pelas famílias Pc e Trx afeta funções essenciais da célula, tais como o ciclo celular, senescência e apoptose (MILLS, 2010). Em diversas doenças, incluindo o câncer, é possível observar uma desregulação na expressão e/ou função destas famílias. No câncer, os membros de Pc e Trx podem agir como oncogenes ou supressores de tumor, dependendo da região na qual ocorre a modificação (KINGSTON e TAMKUN, 2014; POYNTER e KADOCH, 2016).

Um aspecto de antagonismo entre as proteínas das famílias Pc e Trx no câncer reside no fato de que os membros de Polycomb podem atuar inibindo a apoptose, enquanto os

membros de Trithorax podem induzir a mesma. Um exemplo é a proteína BMI1, membro do complexo PRC1, que pode atuar junto com outros fatores inibindo a apoptose de células tumorais da mama, podendo resultar em metástase para o cérebro (MILLS, 2010).

Também foi descrito que a proteína BMI1 pode promover o silenciamento de genes supressores de tumor que induzem senescência, culminando em um aumento do potencial proliferativo das células (MILLS, 2010). Esta proteína encontra-se superexpressa em diversos tipos de câncer, incluindo o de próstata, pulmão, e o CM (MILLS, 2010; POYNTER e KADOCH, 2016).

De função semelhante em câncer, CBX7, outro membro do complexo PRC1, atua diminuindo a expressão de genes supressores de tumor (POYNTER e KADOCH, 2016). A proteína acessória YY1, que também pode fazer parte do complexo PRC1, já foi encontrada como superexpressa em diversos tipos de câncer, incluindo o CM (WAN *et al.*, 2012).

Quanto aos membros do complexo PRC2, da família Pc, sabe-se que podem ocorrer diversos tipos de mutações relacionadas ao câncer. Em linfomas, por exemplo, foram descritas mutações de ganho de função em *EZH2*, resultando em um aumento na marca de metilação promovida por esta proteína em histonas. Mutações de perda de função em membros de PRC2, por outro lado, estão relacionadas com o desenvolvimento de leucemias (SCELFO, PIUNTI e PASINI, 2015).

Não apenas mutações, mas também alterações na expressão das proteínas deste complexo podem contribuir para o desenvolvimento de neoplasias. A proteína *EZH2* é superexpressa em diversos tipos de câncer, incluindo os de pulmão, cólon e próstata (POYNTER e KADOCH, 2016), e também existem relatos de expressão aumentada e ação individual desta proteína na promoção do CM (TAN *et al.*, 2013).

Em CM, altos níveis de *EZH2* são reportados como marcadores de agressividade e pior prognóstico (YOO e HENNIGHAUSEN, 2012; WASSEF e MARGUERON, 2016). De fato, alguns estudos reportam alta expressão gênica e proteica de *EZH2* em tumores TN e linhagens celulares de CM representativas deste subtipo molecular da doença, considerado o mais agressivo (PIETERSEN *et al.*, 2008; HOLM *et al.*, 2012). Meta-análises envolvendo pacientes com diferentes subtipos de CM correlacionam alta expressão de *EZH2* com diversos fatores, tais como pior prognóstico, grau histológico mais avançado e maior índice de proliferação (Ki-67) (JIANG *et al.*, 2015; WANG *et al.*, 2015; NEUSQUEN *et al.*, 2016).

O estudo do papel das proteínas Trx em câncer é mais recente, e portanto, menos dados estão disponíveis sobre sua contribuição para a doença em comparação com a família Pc. Alguns estudos mostram que os genes *KMT2*, membros do complexo MLL, são uns



dos mais frequentemente mutados em câncer. Mutações em *KMT2A*, por exemplo, ocorrem em cerca de 10% das leucemias humanas, e translocações, fusões, duplicações e ampliações deste gene estão descritas em neoplasias hematológicas (MILLS, 2010; SCHUETTENGRUBER, 2011; POYNTER e KADOCH 2016).

Quanto aos dados de expressão gênica e proteica de membros da família Trx, sabe-se, por exemplo, que uma baixa expressão de BRM e BRG1 pode ocorrer no câncer. Isto pode estar relacionado com um potencial papel de Trx na reativação de genes supressores de tumor (MILLS, 2010). Em pacientes com CM, o aumento da expressão do gene *WDR5*, do complexo MLL, foi correlacionado com uma diminuição nos níveis de sobrevivência (DAI *et al.*, 2015).

#### 1.4 Famílias SWI/SNF e NURF

O remodelamento da cromatina também pode se dar através da ação de enzimas e complexos dependentes de ATP que não atuam a nível de modificação de histonas, mas são capazes de promover o remodelamento dos nucleossomos, tais como as famílias SWI/SNF e NURF. Membros destas famílias são capazes de mover, ejetar ou reestruturar os nucleossomos para alterar a acessibilidade ao DNA, facilitar a transcrição pela RNA polimerase II e, conseqüentemente, permitir a ativação da expressão (GANGARAJU e BARTHOLOMEW, 2007; MILLS, 2010; HO e CRABTREE, 2011; NARLIKAR, SUNDARAMOORTHY e OWEN-HUGHES, 2013).

A modificação das histonas e o remodelamento de nucleossomos dependente de ATP são mecanismos que podem agir em conjunto para promover alterações na expressão gênica. De fato, proteínas capazes de promover modificações nas caudas das histonas e enzimas com função dependente de ATP podem coexistir em um mesmo complexo (GANGARAJU e BARTHOLOMEW, 2007).

##### ❖ **Família SWI/SNF**

Complexos da família SWI/SNF podem ser compostos por diferentes proteínas em variadas conformações. As subunidades catalíticas dos complexos SWI/SNF são formadas pelas proteínas BRM e BRG1 (também chamadas de SMARCA2 e SMARCA4, respectivamente), que possuem um domínio com função de ATPase e outro capaz de reconhecer resíduos de lisina acetilados (H3K9ac, por exemplo). A capacidade de reconhecimento de tais modificações permite que os remodeladores de cromatina da família

SWI/SNF possam agir em conjunto com os modificadores de histona para promover a ativação da expressão gênica (DUTTA *et al.*, 2014; SCHUETTENGRUBER *et al.*, 2011).

Outros componentes dos complexos SWI/SNF são as proteínas ARID, que possuem um domínio de ligação ao DNA, auxiliando no recrutamento do complexo para seus alvos na cromatina (GANGARAJU e BARTHOLOMEW, 2007).

A família SWI/SNF é capaz de alterar a estrutura da cromatina em genes envolvidos no ciclo celular, na proliferação e na sinalização celular (SCHUETTENGRUBER *et al.*, 2011; DUTTA *et al.*, 2014).

#### ❖ **Família NURF**

Os membros da família NURF, também conhecido como ISWI, são capazes de reconhecer a trimetilação da lisina 4 na histona H3, promovida pelo complexo MLL. Dentre as proteínas que o compõem, estão BPTF, que reconhece H3K4me3, SMARCA1 (também chamada de ISWI e SNF2L), uma ATPase capaz de se ligar a histonas, e as proteínas RBBP4 e RBBP7, que curiosamente também foram descritas como participantes do complexo PRC2. Esta família de proteínas atua regulando a transcrição e a diferenciação celular (MILLS, 2010; ALKHATIB e LANDRY, 2011; HO e CRABTREE, 2011; SCHUETTENGRUBER *et al.*, 2011).

#### ❖ **Famílias SWI/SNF / NURF e câncer**

Mutações em componentes de SWI/SNF estão descritas em altas frequências em diversos tipos de câncer, incluindo o CM. Uma hipótese para a associação de mutações em SWI/SNF com o desenvolvimento da doença é a de que esta família de remodeladores de cromatina contribui para a estabilidade genômica, e mutações em seus membros poderiam contribuir para a tumorigênese (NARLIKAR, SUNDARAMOORTHY e OWEN-HUGHES, 2013).

Quanto à família NURF, há poucas informações sobre seu papel em câncer, porém, um estudo de 2009 mostrou que a inibição da expressão de SMARCA1 em linhagens celulares de câncer levou a dano de DNA, inibição do crescimento celular e morte por apoptose, indicando um possível papel desta proteína na doença (YE *et al.*, 2009).

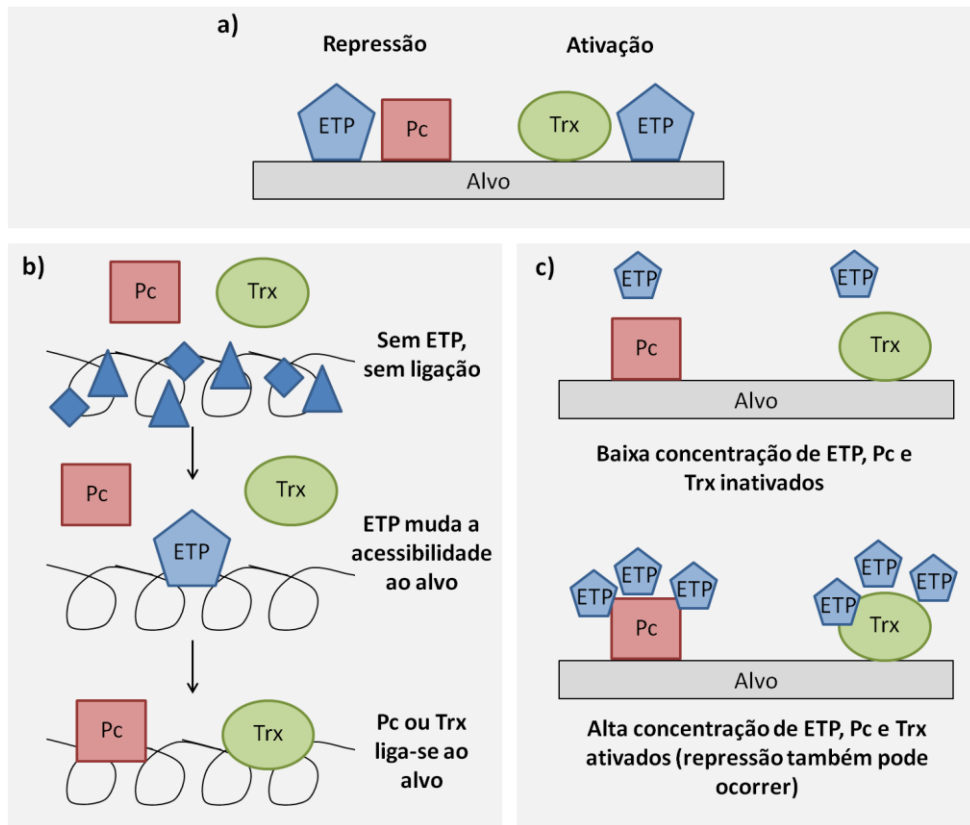
### 1.5 Enhancers de Trithorax e Polycomb

Os autores Gildea, Lopez e Shearn descreveram, em 2000, uma nova classe de proteínas em *Drosophila melanogaster*, denominada *Enhancers of Trithorax and Polycomb* (potencializadores de Polycomb e Trithorax, ETP). Os genes do grupo ETP foram inicialmente identificados como pertencentes à família Pc, porém, após a descoberta de que mutações de perda de função nestes genes são capazes de realçar tanto o fenótipo mutante de Pc quanto o fenótipo mutante de Trx, fez-se necessária sua reclassificação (GILDEA, LOPEZ e SHEARN, 2000).

De acordo com os autores, uma possível explicação para as consequências antagônicas de mutações em genes ETP seria a de que os produtos desses genes, ou seja, as proteínas ETP, sejam componentes de complexos envolvidos tanto na ativação quanto na repressão da expressão gênica (GILDEA, LOPEZ e SHEARN, 2000).

Em 2001, Brock e Van Lohuizen propuseram um modelo para explicar o papel das proteínas ETP nesta função dúbia de ativação e silenciamento. Neste modelo, alguns níveis de regulação mediados por ETPs foram propostos: 1) proteínas ETP podem atuar como subunidades de complexos Pc e Trx, 2) proteínas ETP podem alterar a acessibilidade da cromatina para permitir ou inibir a ligação de complexos Pc e Trx, 3) proteínas ETP podem agir como co-ativadores e/ou co-repressores de Pc e Trx (Figura 1.5). A interação entre proteínas ETP e membros de Pc e/ou Trx pode servir como um mecanismo que leva à alteração do estado de um alvo na cromatina, de neutro para ativado ou reprimido (BROCK e VAN LOHUIZEN, 2001).

Dentre os genes classificados como ETP em *Drosophila*, figuram *Asx* (*Additional sex combs*), *E(Pc)* (*Enhancer of Polycomb*), *Psc* (*Posterior sex combs*), *Scm* (*Sex combs on midleg*), *E(z)* (*Enhancer of zeste*) e *Su(z)2* (*Suppressor of zeste 2*). Em humanos, genes homólogos aos ETPs de *Drosophila* foram descritos, dentre eles os genes das famílias *ASXL*, *EPC*, *EZH* e *SCML* (GILDEA, LOPEZ e SHEARN, 2000; BROCK e VAN LOHUIZEN, 2001).



**Figura 1.5 – Níveis de regulação mediados por ETPs.** a) ETPs como subunidades de complexos Pc e Trx. b) alteração da acessibilidade à cromatina mediada por ETPs. c) ETPs como co-ativadores e co-repressores de Pc e Trx. Adaptado de Brock e van Lohuizen, 2001.

As funções destes genes como ativadores e repressores da expressão gênica em humanos, porém, são ainda pouco conhecidas e exploradas. Estudos recentes buscam compreender o papel destes genes e de suas proteínas correspondentes na regulação da expressão gênica, uma vez que alguns deles, como *EPC1*, *ASXL1* e *ASXL2* foram descritos como associados a câncer, principalmente os de origem hematológica (HUANG *et al.*, 2014; MICOL e ABDEL-WAHAB, 2016).

### 1.6 Proteínas ASXL

O gene *Asx* (*Additional sex combs*) foi inicialmente identificado como membro de Pc, assim como os outros membros do grupo ETP. No desenvolvimento de *Drosophila*, mutações em *Asx* resultam em dois fenótipos distintos, um característico de Pc e outro característico de Trx, decorrentes de repressão e ativação impróprias de genes *Hox*. Além disso, análises genéticas na espécie identificaram que *Asx* interage com genes de ambas as famílias,

caracterizando-o, portanto, como membro de ETP (GILDEA, LOPEZ e SHEARN, 2000; BROCK e VAN LOHUIZEN, 2001; BASKIND *et al.*, 2009; LI *et al.*, 2017).

Recentemente, também em *Drosophila*, Li *et al.* (2017) identificaram que a proteína *Asx* é capaz de interagir diretamente com E(z) (*enhancer of zeste*, homólogo de EZH) e Trx, através de ligação ao domínio SET destas proteínas, na regulação do gene *hsp70*. Estas proteínas são as histona metiltransferases responsáveis, respectivamente, pelas marcas de trimetilação em H3K27, característica de Pc, e em H3K4, característica de Trx. Além de se ligar a estas proteínas, *Asx* é capaz de regular os níveis de metilação promovidos pelas mesmas, indicando que de fato *Asx* tem um papel na regulação e função de Pc e Trx (LI *et al.*, 2017).

Em humanos, foram descritos três homólogos do gene *Asx*, os chamados *ASXL 1, 2 e 3* (*Additional sex combs-like 1, 2 e 3*). Análises filogenéticas dos membros da família ASXL mostram que as proteínas ASXL1 e ASXL2 são mais relacionadas entre si quando comparadas com a proteína ASXL3. Em termos de expressão, ASXL1 e ASXL2 estão presentes em uma ampla gama de tecidos, incluindo células do sistema hematopoiético, enquanto a expressão de ASXL3 encontra-se restrita principalmente aos olhos e cérebro (KATOH, 2013; MICOL e ABDEL-WAHAB, 2016).

As proteínas ASXL apresentam diversos domínios em comum, entre eles o domínio ASXN, predito como um domínio de ligação ao DNA, que não é encontrado no homólogo *Asx* em *Drosophila*. Além de ASXN, destacam-se os domínios ASXH, capaz de promover a interação com outras proteínas, ASXM1 e ASXM2, de ligação a receptores hormonais nucleares (incluindo receptores de estrógeno), e o domínio PHD, cuja função ainda é pouco conhecida. Acredita-se que o domínio PHD seja capaz de permitir a ligação de ASXL às proteínas histonas ou ao DNA, porém seus alvos ainda não são conhecidos (KATOH, 2013; MICOL e ABDEL-WAHAB, 2016). Análises filogenéticas sugerem que o domínio PHD de ASXL seja capaz de se ligar a H3K4me3, uma modificação promovida pelo complexo MLL da família Trx (KATOH, 2015).

Através do domínio ASXH, todas as proteínas da família ASXL são capazes de interagir com a proteína BAP1 para promover a remoção da ubiquitinação da lisina 119 na histona H2A (H2AK119ub), numa função antagônica à do complexo PRC1, da família Pc. O mesmo é observado em *Drosophila*, onde mutantes que não expressam *Asx* ou *Calypso* (homólogo de BAP1) tem um aumento global da marca de ubiquitinação em H2AK119 (ABDEL-WAHAB e DEY, 2013; MICOL e ABDEL-WAHAB, 2016).

Mutações em *ASXL1* e *ASXL2*, ou deleção desses genes, afetam a função de ubiquitinase de BAP1 e, por conseguinte, aumentam os níveis de ubiquitinação em H2AK119. Daou *et al.* (2015) observaram que esta disfunção na atividade do complexo ASXL-BAP1 pode promover a desregulação do ciclo celular e o aumento da proliferação. Isto sugere que os complexos ASXL-BAP1 podem contribuir para o controle do ciclo celular e proliferação, e que alterações na ação dos mesmos podem ter um papel no câncer.

A interação das proteínas ASXL com BAP1 também pode promover a ativação do complexo BAP1-HCF1, que leva ao recrutamento de complexos como MLL e, por conseguinte, a uma ativação transcricional mediada pela trimetilação de H3K4 (KATO, 2013; DAOU *et al.*, 2015).

Além disso, em células hematológicas, foi identificado que *ASXL1* é capaz de interagir com membros do complexo PRC2, e que sua perda leva a uma redução global da marca de trimetilação característica deste complexo, em H3K27, sugerindo que *ASXL1* tem um papel no recrutamento e/ou na estabilização do complexo em seus alvos na cromatina (ABDEL-WAHAB e DEY, 2013). De forma similar, em modelo murino, observou-se que a perda de *ASXL2* leva a uma redução global de H3K27me3 no tecido cardíaco (BASKIND *et al.*, 2009).

### 1.6.1 Família ASXL e câncer

Mutações somáticas de ganho e/ou perda de função em membros da família ASXL são observadas em alguns tipos de câncer. *ASXL1* pode ser encontrado mutado em síndrome mielodisplásica, leucemia mieloide aguda (LMA) e outras neoplasias hematológicas, câncer de fígado, de próstata e de mama, entre outros. Em malignidades mieloides, as mutações em *ASXL1* apresentam importante valor prognóstico, estando associadas com maior agressividade, menor sobrevivência, e mais rápida progressão da doença (GELSI-BOYER *et al.*, 2012). Mutações em *ASXL2* já foram descritas em neoplasias hematológicas, câncer de pâncreas, próstata e mama, e *ASXL3* pode estar mutado em melanoma (KATO, 2015).

Mutações em *ASXL1* ou *ASXL2* foram descritas em casos de câncer de mama RE-positivos, e podem estar relacionadas com a perda dos domínios ASX, de ligação a receptores nucleares como o RE (KATO, 2015).

Recentemente, foi descrito que, em murinos, *Asxl2* é essencial para a hematopoiese normal, possuindo um importante papel de supressão de tumor, e sua perda pode acarretar no desenvolvimento de malignidades mieloides. A deleção de *Asxl2* em murinos (camundongos

*knockout*) e o silenciamento de ASXL2 em células leucêmicas humanas estão associadas à desregulação de H3K27ac e H3K4me1/2 (LI *et al.*, 2017; MICOL *et al.*, 2017).

Park *et al.* (2016) observaram expressão aumentada de ASXL2 em células de CM positivas para RE, e demonstraram que, a partir de estímulo com 17 $\beta$ -estradiol, ASXL2 é capaz de interagir com este receptor, mediando sua ativação e contribuindo para a transcrição de seus genes alvo. ASXL2 pode recrutar LSD1, UTX e MLL2 para os promotores destes genes para promover, respectivamente, a remoção de marcas de silenciamento em H3K9 e H3K27, e a adição de marcas de ativação em H3K4. Isto possivelmente leva à ativação de genes alvo de RE que contribuem para a carcinogênese da mama (PARK *et al.*, 2016).

### 1.7 Terapias com alvo em modificadores de histona

Juntamente com as recentes descobertas envolvendo as modificações de histona no câncer, há um crescente interesse no desenvolvimento e aplicação de novas terapias voltadas para este tipo de modificação epigenética.

Medicamentos e moléculas naturais cujos alvos são as modificações epigenéticas das histonas podem apresentar ação antitumoral, e alguns já são utilizados para o tratamento de alguns tipos de câncer. Estes medicamentos, seus mecanismos de ação e suas aplicações estão descritos na Tabela 1.2.

A fim de aumentar a eficácia terapêutica, diversas drogas quimioterápicas tem sido testadas em combinação com terapias epigenéticas, como a combinação entre Vorinostat e Tamoxifeno, que, em estudos clínicos, tem mostrado boa atividade em CM positivo para RE (SACHAN e KAUR, 2015). Uma melhora na eficácia também pode resultar da combinação entre duas ou mais terapias epigenéticas, ou entre elas e a radioterapia (DAWSON e KOUZARIDES, 2012; SACHAN e KAUR, 2015).

A terapia epigenética, apesar de bastante promissora, apresenta algumas limitações, como a toxicidade das drogas, a baixa especificidade das mesmas, que pode levar a efeitos globais indesejados, e os poucos resultados satisfatórios em tumores sólidos, em comparação com uma alta suscetibilidade em tumores hematopoiéticos (DAWSON e KOUZARIDES, 2012; NERVI, DE MARINIS e CODACCI-PISANELLI, 2015; SACHAN e KAUR, 2015). A especificidade dos medicamentos é um aspecto de grande importância que deve ser considerado no desenvolvimento de novas drogas, a fim de evitar a ativação de oncogenes e o silenciamento de genes supressores de tumor, contribuindo para a redução da toxicidade e dos efeitos adversos (SACHAN e KAUR, 2015).

Uma vantagem dos inibidores epigenéticos é sua capacidade de promover efeitos importantes sobre as células tumorais, enquanto as células normais se mantêm inalteradas, de forma geral. Isto ocorre porque as células tumorais são dependentes das vias epigenéticas para a manutenção da expressão gênica, ao passo que células normais são menos dependentes, uma vez que podem possuir vias compensatórias de manutenção da expressão (DAWSON e KOUZARIDES, 2012).

**Tabela 1.2 – Terapias com alvo em modificadores de histona.**

<b>Tipo de medicamento</b>	<b>Mecanismo de ação</b>	<b>Aplicação</b>	<b>Nome comercial / Princípio Ativo</b>
Inibidores de HDACs	Reversão do silenciamento de genes supressores de tumor; diferenciação e apoptose de células tumorais; inibição de metastase e angiogênese (DAWSON e KOUZARIDES, 2012; SACHAN e KAUR, 2015).	Aprovados para uso em linfoma cutâneo de células T.	Vorinostat (Zolinza®, Merck); Romidepsin (Istodax®, Celgene)
Inibidores de HATs	Interferência em cascatas de sinalização relacionadas com proliferação, invasão e metástase (SACHAN e KAUR, 2015).	Não aprovadas para uso, mas apresentam ação em diversos tipos de câncer.	Curcumina, Garcinol, Ácido Anacárdico
Inibidores de HMT	Redução da expressão de EZH2 e dos níveis de H3K27me3, com consequente ativação de genes alvo de PRC2 e indução de apoptose (CREA <i>et al.</i> , 2012; MOMPALER e CÔTÉ, 2015. SACHAN e KAUR, 2015).	Ação <i>in vitro</i> em diversas linhagens celulares de câncer.	3-Deazaneplanocin A (DZNep)



## 2. JUSTIFICATIVA

O CM é uma doença heterogênea que envolve, entre outros fatores, um componente epigenético, capaz de contribuir para o desenvolvimento e progressão da doença. Dentre os mecanismos epigenéticos de alteração da expressão gênica, a modificação de histonas ainda é o menos estudado.

A existência de famílias de proteínas com papel de regulação e controle dos efetores de modificações de histona adiciona complexidade ao estudo deste mecanismo epigenético. Dentre elas, pode-se destacar a família ASXL, capaz de regular em diversos níveis a ação de proteínas Pc e Trx.

Mutações em genes *ASXL* são bem descritas em câncer, a maioria em *ASXL1* e principalmente em tumores hematológicos, e estão relacionadas com o prognóstico da doença. Apesar disso, poucos dados estão disponíveis sobre a ação dos membros da família ASXL e os mecanismos pelos quais as proteínas ASXL podem contribuir para a regulação epigenética no câncer.

Recentemente, em CM, a proteína ASXL2 foi descrita como uma possível reguladora da expressão de genes alvo de RE, através do recrutamento e regulação de modificadores de histona para a cromatina (PARK *et al.*, 2016). Ainda assim, sua ação como ETP em CM permanece pouco explorada.

Desta forma, o estudo mais aprofundado da função e mecanismo de ação de ASXL2 no CM permitirá elucidar seu papel no controle das famílias Pc e Trx e no desenvolvimento desta doença, e poderá contribuir para o desenvolvimento de novas abordagens terapêuticas, mais específicas, para o CM no futuro.

### **3. OBJETIVOS**

#### 3.1 Objetivo Geral

Investigar o papel de ASXL2 em modelos *in vitro* do câncer de mama.

#### 3.2 Objetivos Específicos

- Avaliar a expressão de ASXL2 em linhagens celulares utilizadas como modelos *in vitro* dos subtipos de CM e na linhagem normal de mama HMEC;
- Avaliar a localização de ASXL2 em conjunto com a expressão e localização de membros das famílias Pc e Trx em linhagens celulares de CM;
- Realizar o silenciamento gênico transiente de ASXL2 em linhagens celulares de CM;
- Identificar, *in silico*, vias de sinalização, interações e processos biológicos potencialmente relacionados com ASXL2 em CM;
- Validar genes e processos biológicos de interesse;
- Identificar os parceiros de ASXL2 em linhagens celulares de CM;
- Investigar a presença de ASXL2 na cromatina em linhagens celulares de CM.

## 4. METODOLOGIA

### 4.1 Cultivo de linhagens celulares

Neste estudo, foram utilizadas linhagens celulares de CM representativas dos subtipos moleculares da doença: MCF-7 (receptor de estrógeno positiva), EVSA-T (receptor de progesterona positiva), HCC-1954 (HER2/neu positiva) e MDA-MB-231 (triplo negativa). Além disso, a linhagem HMEC (*Human Mammary Epithelial Cells* - Thermo Fisher) foi utilizada como controle de mama normal.

O cultivo das linhagens foi realizado em estufa a 37°C em 5% de concentração de CO<sub>2</sub>. As linhagens MCF-7 e MDA-MB-231 foram cultivadas com meio de cultura RPMI-1640 (Thermo Fisher) suplementado com L-glutamina 2 mM (Thermo Fisher), e soro fetal bovino (SBF, Thermo Fisher) em concentração final de 10%. A linhagem HCC-1954 foi cultivada nas mesmas condições, porém com uma concentração final de 15% de SBF. A linhagem EVSA-T, por sua vez, foi cultivada em meio  $\alpha$ MEM (Thermo Fisher), também suplementado com L-glutamina 2 mM, e com 10% de SBF. Para a linhagem HMEC, utilizou-se meio de cultura basal HuMEC (Thermo Fisher), suplementado com 25 mg de extrato de pituitária bovina (Thermo Fisher) e 1% de suplemento HuMEC (Thermo Fisher). Todas as linhagens foram cultivadas com 1% de antibiótico Penicilina/Streptomicina (Thermo Fisher).

As linhagens celulares utilizadas crescem na forma de monocamada aderente, e, ao atingirem 70% de confluência de cultivo, foram submetidas à expansão. Para tal, utilizou-se enzima tripsina (Thermo Fisher) diluída em PBS 1X estéril (Sigma-Aldrich). Para a linhagem HCC-1954, utilizou-se 0,05 U de tripsina, e para as demais linhagens, 0,025 U. Em seguida, as células foram centrifugadas a 1500 rpm por 10 min em temperatura ambiente (TA) e lavadas duas vezes com PBS 1X estéril para expansão. Este mesmo processo foi realizado a fim de se obter a densidade de células necessária para os experimentos posteriores a serem descritos.

### 4.2 Lista de anticorpos

Os seguintes anticorpos primários foram utilizados: anti-ASXL2 (IgG de coelho, ab106540, Abcam), anti-BMI1 (IgG<sub>1</sub> de camundongo, 37-5400, Thermo Fisher), anti-BPTF (IgG<sub>1</sub> de camundongo, sc-81088, Santa Cruz), anti-EED (IgG de cabra, sc-18456, Santa Cruz), anti-EZH2 (IgG de coelho, 36-6300, Thermo Fisher), anti-GAPDH (IgG<sub>1</sub> de

camundongo, AM4300, Thermo Fisher), anti-H3K27me3 (IgG de coelho, #9756, Cell Signaling), anti-MLL (IgG de cabra, sc-30729, Santa Cruz), anti-SMARCA1 (IgG de coelho, ab175226, Abcam), anti-SUZ12 (IgG de coelho, ab12073, Abcam), anti-YY1 (IgG de camundongo, sc-7341, Santa Cruz).

Os anticorpos secundários conjugados a HRP utilizados no ensaio de *Western Blot* foram: cabra anti-coelho (sc-2004, Santa Cruz), jumento anti-cabra (sc-2033, Santa Cruz) e cabra anti-camundongo (sc-2005, Santa Cruz).

Já os seguintes anticorpos secundários marcados com fluoróforo (Alexa Fluor®) foram utilizados para o ensaio de Imunofluorescência: cabra anti-camundongo 488 (A-11001, Thermo Fisher), cabra anti-coelho 546 (A-11037, Thermo Fisher), jumento anti-coelho 488 (A-21206, Thermo Fisher), jumento anti-cabra 555 (A-21432, Thermo Fisher).

### 4.3 Análise de expressão gênica

#### *4.3.1 Extração de RNA em coluna*

O RNA total foi extraído dos citocentrifugados das linhagens celulares HMEC, MCF-7, EVSA-T, HCC-1954 e MDA-MB-231, através do *kit* de extração em coluna *RNeasy* (Qiagen), seguindo o protocolo fornecido pelo fabricante. Todas as centrifugações foram realizadas em TA.

Uma densidade de  $1 \times 10^6$  células de cada linhagem foi ressuspensa em 350  $\mu$ L de tampão RLT acrescido de 1%  $\beta$ -mercaptoetanol, para lise celular. Em seguida, foram adicionados 350  $\mu$ L de etanol 70% (RNase free), e, após homogeneização, a totalidade das amostras (700  $\mu$ L) foi aplicada às colunas de purificação fornecidas no *kit*. Os lisados celulares foram então centrifugados a 8000 x g por 15 seg, e o eluído da coluna foi descartado.

As membranas das colunas foram lavadas com 700  $\mu$ L de tampão RW1, com centrifugação a 8000 x g por 15 seg seguida de descarte do eluído. Em seguida, foram lavadas novamente com 500  $\mu$ L de tampão RPE, com repetição da centrifugação. Uma terceira lavagem foi realizada com 500  $\mu$ L de tampão RPE, seguida de centrifugação a 8000 x g por 2 min. Posteriormente, foi realizada uma centrifugação à velocidade máxima (20238 x g) por um minuto, para total descarte do volume das colunas.

Foi realizada a eluição do RNA, com a adição de 20  $\mu\text{L}$  de água RNase *free* à membrana da coluna, seguida de centrifugação a 8000 x g por 1 min. Após a extração de RNA, os mesmos foram quantificados no Espectrofotômetro ND-1000 (Nanodrop Technologies), no comprimento de onda de 260 nm. Como controle de qualidade, para avaliação de integridade e qualidade dos RNAs, foram considerados satisfatórios aqueles RNAs com razão de absorbância 260/280 entre 1,8 e 2,2. Os RNAs extraídos e quantificados foram estocados em microtubos (1,5 mL) a  $-80^{\circ}\text{C}$  até a sua utilização.

#### 4.3.2 Reação de digestão com DNase

Para eliminar as possíveis contaminações com DNA, 2  $\mu\text{g}$  de RNA total obtidos de cada linhagem celular foram submetidos à reação de digestão com a enzima *DNase I* (Thermo Fisher). Aos 2  $\mu\text{g}$  de RNA total (microtubo 200  $\mu\text{L}$ ), adicionou-se água RNase *free* para um volume de 8  $\mu\text{L}$ , e em seguida foram adicionados 1  $\mu\text{L}$  da enzima 1 U/ $\mu\text{L}$  e 1  $\mu\text{L}$  de tampão *DNase I reaction buffer 10X*. As reações foram incubadas por 30 min em TA, conforme protocolo do fabricante. Em seguida, para a inativação da enzima, adicionou-se 1  $\mu\text{L}$  de solução EDTA 25  $\mu\text{M}$ , com incubação por 10 min a  $65^{\circ}\text{C}$  no termociclador *MyCycler* (BioRad).

#### 4.3.3 Transcrição reversa (RT)

A transcrição reversa para produção do DNA complementar (cDNA) foi realizada com a enzima *SuperScript II* (Thermo Fisher). À solução de RNAs tratados com DNase, adicionou-se 1  $\mu\text{L}$  de Oligo(dT)<sub>12-18</sub> 500  $\mu\text{g}/\mu\text{L}$ . Além disso, adicionou-se 1  $\mu\text{L}$  de dNTP Mix 10 mM, que fornece os desoxirribonucleotídeos fosfatados utilizados para a síntese das fitas de cDNA. Esta solução foi incubada por 15 min a  $65^{\circ}\text{C}$  no termociclador *MyCycler* (BioRad), seguida de rápido resfriamento no gelo ( $0^{\circ}\text{C}$ ).

A cada microtubo (200  $\mu\text{L}$ ), adicionou-se 4  $\mu\text{L}$  de tampão *First-Strand Buffer 5X*, 2  $\mu\text{L}$  de DTT 0,1 M e 1  $\mu\text{L}$  da enzima RNase OUT 40 U/ $\mu\text{L}$  para a inibição de RNases. Após homogeneização, as reações foram incubadas por 2 min a  $42^{\circ}\text{C}$  no termociclador.

Adicionou-se, a cada microtubo (200  $\mu\text{L}$ ), 1  $\mu\text{L}$  da enzima *SuperScript II*, transcriptase reversa, com posterior incubação no termociclador por 50 min a  $42^{\circ}\text{C}$ , para a síntese do cDNA, e por 15 min a  $70^{\circ}\text{C}$ , para a inativação da mesma. Os cDNAs foram armazenados a  $-20^{\circ}\text{C}$  até sua utilização.

#### 4.3.4 Elaboração dos iniciadores

O desenho de iniciadores (*primers*) para cada gene de interesse foi realizado a partir da sequência do RNA mensageiro (RNAm) de cada gene. As mesmas foram obtidas no banco de dados do genoma humano ([www.ncbi.nlm.nih.gov/nucleotide](http://www.ncbi.nlm.nih.gov/nucleotide)), no formato FASTA. A visualização dos éxons foi feita a partir do site Ensembl ([www.ensembl.org](http://www.ensembl.org)). Os *primers* senso (*forward* - fw) e antissenso (*reverse* - rv) foram desenhados em éxons diferentes, em alguns casos com a ajuda do software *Primer3* (NCBI).

A ferramenta online Oligo Analyzer ([www.idtdna.com/calc/analyzer](http://www.idtdna.com/calc/analyzer)), da IDT (*Integrated DNA Technologies*), foi utilizada para avaliação dos seguintes parâmetros nos *primers* desenhados: tamanho de 18 a 22 pares de bases; conteúdo de guanina e citosina próximo ou equivalente a 50%; temperatura de dissociação (*melting*) entre 55°C e 60°C; temperatura de formação de grampo (*hairpin*) de no máximo 40°C; e  $\Delta G \geq -10$  kcal/mole para *self-dimer* (dímero entre dois *primers* iguais) e *hetero-dimer* (dímero entre *primers* senso e antissenso).

As sequências dos *primers* desenhados foram submetidas à análise pela ferramenta BLAST ([www.blast.ncbi.nlm.nih.gov/Blast.cgi](http://www.blast.ncbi.nlm.nih.gov/Blast.cgi)), que permite verificar se os mesmos são complementares exclusivamente à sequência correspondente ao gene de interesse. Os genes avaliados e seus respectivos *primers* estão discriminados na Tabela 4.1.

**Tabela 4.1 – Sequências dos primers senso (fw) e antissenso (rv) utilizados neste estudo.**

Gene	Sequência (5' - 3')		Temperatura de anelamento (°C)
	fw	rv	
ARID1A	TCC CAG CAA ACT GCC TAT TC	ATC TTC TTG CCC TCC CTT AC	62
ASH2L	TCC CGA ACA CAG GAT GAA CAT C	CCA AGT CAT TTT CCC AGG TCT C	64
ASXL1	TCT TTC CAA TCC CAG GGA CAG	CAG CAT CAC CCC AGT CTT TTT C	62
ASXL2	GGC AGC CCT CAA AAC TCA AAC	TCC GGT GTC TCA ACG TCA ATG	60
BM11	CTG ATG TGT GTG CTT TGT GGA G	TGG TCT GGT CTT GTG AAC TTG G	63
BPTF	GAA GAA ACG ATG CAG CAA GC	TCC AAC CAC CAT ATC CTG TC	60
CBX8	TGA AAT GGA AGG GAT GGT CG	AGA GCT CCA TCT CTC TTT CC	64
CTNNB1	AAG ACA TCA CTG AGC CTG CCA T	CGA TTT GCG GGA CAA AGG GCA A	60
EED	TGA GCA GTG ACG AGA ACA GCA A	TCC AGG TGC ATT TGG CGT GTT T	64
EZH1	CTT CAG GCA AAT ATG GGT GC	TCC ACT CAC AGG CTT CAT TG	60
EZH2	AGA AGG GAC CAG TTT GTT GG	GTG CAC AGG CTG TAT CCT TC	62
GAPDH	GTC AAC GGA TTT GGT CGT ATT G	TGG AAG ATG GTG ATG GGA TTT	60
JADE1	TGC TCC AGA CAT GAA GAT CG	AGG GAT CTG CCA ACA CAT AG	60
KMT2A	ACG AGC TGT GTT TCC TGA TGA C	TTG GTT TGA TGG GTG GAG CAA G	64
KMT2B	AAA GCA GAA GTC TCC CGA TG	CCA TCA GCC ACA CGT TCT CTA	60
KREMEN1	CGT CTC TCT GGA CTT CGT CAT CTT	CCT GTG GCA GTT CTT CCT TGA	60
LEF1	CAG ACA TCC TCC AGC TCC TGA TAT	CGT GAT GGG ATA TAC AGG CTG AC	60
PHC3	ACA GGA AGT GTC ACA CAG CAG	GGG AAC GGC TTA TTA ACT GTG C	60
PORCN	CCC TCC TAC ATG GCT TCA GTT	CCG CTT TGA CAA GAC ACA GG	60
PRICKLE1	GGT GCT CAG CAT GTG ACG AGA TAA	TCA CAC TCA AGG CAG CAG AAG T	60
PRMT6	TGA AGC AGC ACT ATG GTG TGG	ACA ATC CCT GCA CAA CGA TCT C	62
RBBP4	AGT GTG CAG CTC CCT AAT GAT G	GGC CCT GTT TAC TTC TCC TTC	60
RBBP7	TTA CCG TTC AGT GGC TTC CTG	TCG AGC AAC CAC CAG ATG ATT C	62
RING1	CAG AAT GCC AGC AAA ACG TG	AGC AAT CTC TGT GCC ATC CA	60
RSBN1L	TGA TCC CTG TGG CTG ATA TG	TGC ATG AGC TCT TGT CCT GT	62
RUNX2	CGG AAT GCC TCT GCT GTT AT	AGT CAT CAA GCT TCT GTC TGT G	60
SIRT1	AAC AGG TTG CGG GAA TCC AAA G	ATC TCC TCG TAC AGC TTC ACA G	62
SMAD3	CGT GCG GCT CTA CTA CAT C	CCT CAA AGC CCT GGT T	62
SMAD4	GCA ATT GAA AGT TTG GTA AAG AAG C	GTG AAG ATC AGG CCA CCT C	60
SMARCA1	ACT GGC GAT ACC TGG TCA TTG	GCC CAC AGT TCA TGC AGG TTA T	63
SMARCA2	ATC CAT GAA GGG CAC TGG TAT G	AAT GGA GAT GGG TGT GGT GAC A	64
SMARCA4	TCA GGA CAA CAT GCA CCA GAT G	CGC ATC CCC ATT CCT TTC AT	63
SUZ12	GAT AAA AAC AGG CGC TTA CAG CTT	AGG TCC CTG AGA AAA TGT TTC GA	60
TCF7	ACT CTT CCC GGA CAA ACT TCC	GCA GAT TGA AGG CGG AGT AGA C	60
TGFB1	CAA CAA TTC CTG GCG ATA CC	CAG TGT GTT ATC CCT GCT G	62
WDR5	CTT GAA GAT ATG GGA CGT GAG C	CTT GAG GCA CTT CCC TGT TTT C	62
YY1	TCA GAT CCC AAA CAA CTG GCA G	ATG GCC GAG TTA TCC CTG AAC	62

#### 4.3.5 Análise da expressão gênica diferencial por PCR quantitativo em Tempo Real

Para as análises da expressão gênica diferencial (RNAm) em PCR quantitativo em tempo real (RT-qPCR), foi utilizado o intercalante de DNA *Sybr Green*, que se liga à fita dupla de DNA e emite fluorescência quando excitado, permitindo que a reação seja quantificada e acompanhada em tempo real. Inicialmente, os primers tiveram o anelamento testado entre as temperaturas de 58°C e 64°C, com a finalidade de otimizar a reação de PCR, e as temperaturas de anelamento padronizadas para cada par de primers encontram-se discriminadas na Tabela 4.1.

Em seguida, foi realizada a padronização das curvas padrão dos genes de interesse, a fim de se obter valores de eficiência acima de 70% e de R<sup>2</sup> acima de 0,99. Para tal, foram

utilizadas concentrações seriadas de cDNA (100 ng, 10 ng, 1 ng, 0,1 ng e 0,01 ng) das linhagens celulares K562 e MDA-MB-231 (controles positivos).

As reações para análise quantitativa foram realizadas nas linhagens celulares HMEC, MCF-7, EVSA-T, HCC-1954 e MDA-MB-231, em triplicatas experimentais, e consistiam de 2,5 µL do *mix* de *primers* 2µM, 2,5 µL de cDNA 6 ou 10 ng/µL, e 5µL de reagente *SYBR Green Master Mix* (Applied Biosystems). As reações de PCR foram submetidas ao termociclador *Rotor-Gene Q* (Qiagen), inicialmente por 10 min a 95°C; seguido de 60 ciclos de 20 seg a 95°C, 30 seg a 58-64°C (etapa de anelamento, temperatura variável de acordo com cada par de *primers*), e 30 seg a 72°C; e uma curva de dissociação (*melting*) ao final da reação.

Para realização da quantificação relativa, obteve-se, a partir das curvas padrão, um valor de *threshold* para cada par de *primers* e uma equação de relação entre massa e fluorescência, fornecidos pelo *software* do termociclador (*Rotor-Gene Q Series Software* - Qiagen). Aplicou-se o valor de *threshold* em cada experimento, para obtenção do valor de Ct (*threshold cycle*, ciclo no qual a fluorescência ultrapassa o valor do *threshold*), que foi então aplicado na equação de massa previamente obtida, gerando um valor de massa. A expressão gênica de cada gene de interesse foi normalizada pela massa do gene endógeno *GAPDH*.

#### 4.4 Análise de expressão proteica

##### *4.4.1 Lise celular e extração de proteínas totais*

Para a obtenção do extrato total de proteínas das linhagens celulares estudadas, utilizou-se uma densidade de  $1 \times 10^6$  células. Após duas lavagens com PBS 1X estéril, adicionou-se 40 µL de tampão de extração de proteínas totais [Hepes 100 mM, KCl 500 mM, MgCl<sub>2</sub> 5 mM, EDTA 0,5 mM, glicerol 35%, β-glicerolfosfato 0,1 mM, Ortovanadato de sódio (Na<sub>3</sub>VO<sub>4</sub>) 0,1 mM e DTT 1 mM] e 1% de mix inibidor de proteases (GE Healthcare Bio-Sciences).

Após homogeneização, os extratos celulares foram congelados em nitrogênio líquido e descongelados em TA por três vezes. Em seguida, as amostras foram mantidas por 30 min no gelo (0°C), e posteriormente centrifugadas a 12000 x g, por 30 min, a 4°C. O sobrenadante, que consiste no extrato total de proteínas, foi transferido para um microtubo (1,5 mL) identificado e armazenado à temperatura de -80°C até sua quantificação e uso.



#### 4.4.2 Lise celular e extração de proteínas de núcleo e citoplasma

Para a obtenção de extrato de proteínas das frações de núcleo e citoplasma, foi utilizada uma densidade  $4 \times 10^7$  células. Após duas lavagens com PBS 1X estéril, as células foram ressuspensas em 900  $\mu$ L tampão de extração de proteínas citoplasmáticas (Tampão A: Tris HCl 10 mM pH 8,0, KCl 60 mM, EDTA 1 mM, DTT 1 mM) e 2  $\mu$ L de mix inibidor de proteases (GE Healthcare Bio-Sciences), e incubadas por 15 min no gelo (0°C). Em seguida, foram adicionados 100  $\mu$ L de NP-40 5% aos microtubos (1,5 mL), com posterior homogeneização e centrifugação a 1800 x g, por 5 min, a 4°C. Os sobrenadantes, que correspondem aos extratos citoplasmáticos, foram coletados para novos microtubos (1,5 mL).

Os *pellets* resultantes da centrifugação foram lavados com 900  $\mu$ L do tampão A e centrifugados a 4100 x g, por 5 min, a 4°C. Os sobrenadantes foram descartados e os *pellets* ressuspensos em 70  $\mu$ L de tampão de extração de proteínas nucleares (Tampão C: Tris HCl 20 mM pH 8,0, NaCl 420 mM, MgCl<sub>2</sub> 1,5 mM, EDTA 0,2 mM, Glicerol 25%) e 1  $\mu$ L de mix inibidor de proteases (GE Healthcare Bio-Sciences). Em seguida, os microtubos (1,5 mL) foram submetidos a congelamento em nitrogênio líquido seguido de descongelamento em TA por três vezes, e incubados por 15 min no gelo (0°C).

Foi realizada uma centrifugação a 14000 x g, por 15 min, a 4°C, e os sobrenadantes, que consistem nos extratos nucleares, foram coletados para novos microtubos (1,5 mL). Os extratos nuclear e citoplasmático foram armazenados a -80°C até sua quantificação e utilização.

#### 4.4.3 Quantificação de proteínas

Os extratos de proteína obtidos foram quantificados através do método de Bradford (BRADFORD, 1976). Para tal, fez-se necessária a realização de uma curva padrão com albumina do soro bovino (BSA, *bovine serum albumin*, 1  $\mu$ g/ $\mu$ L) (Sigma-Aldrich). As concentrações utilizadas na curva foram de 0,5  $\mu$ g, 1  $\mu$ g, 2,5  $\mu$ g, 5  $\mu$ g, 7,5  $\mu$ g e 10  $\mu$ g de BSA. Para a quantificação dos extratos de proteína, utilizou-se 1  $\mu$ L dos mesmos.

As reações foram completadas com água destilada q.s.p. 100  $\mu$ L e 900  $\mu$ L do reagente de Bradford (Azul de Coomassie G-250 10%; Etanol 5%; Ácido fosfórico 85%, 10%; q.s.p. água milliQ). Após incubação de 10 min, foi realizada a leitura da absorbância no comprimento de onda de 595 nm com o espectrofotômetro BioPhotometer (Eppendorf),

através da cubeta UVette® (Eppendorf). A quantificação foi realizada através da correlação entre as absorbâncias observadas na curva padrão e aquelas observadas para cada amostra.

#### 4.4.4 Western Blot

Foram resolvidos 30 µg de extrato de proteínas de cada linhagem, adicionados de tampão de amostra 4X (SDS 4%, Glicerol 20%, 2-Mercaptoetanol 10%, Azul de Bromofenol 0,0004%, Tris 0,125 M), em gel de poliacrilamida SDS-PAGE (Acrilamida 30%, Bis-Acrilamida 0,8%; Tris HCl 1,5 M pH 8,8; SDS 10%; APS 10%; TEMED 1%). Para tal, utilizou-se tampão de corrida 2X [Tampão Tris-Glicina 10X (Tris 250 mM, Glicina 1,9 M, SDS 34,7 mM), diluído em H<sub>2</sub>O Milli-Q], a uma voltagem fixa de 90 V.

Após a resolução das proteínas, as mesmas foram transferidas para uma membrana de nitrocelulose de 0,45 µm (Biorad), sob o sistema semi-seco Trans-Blot® Turbo™ Transfer System (Biorad), por 30 min a uma voltagem fixa de 25 V. Para esta etapa, utilizou-se tampão de transferência (Tris 48 mM, Glicina 38,6 mM, SDS 1,3 mM, Metanol 20%), com incubação do gel, membrana e papéis de filtro (Munktell) nesta solução por 15 min antes da transferência.

Para confirmação da transferência, a membrana foi corada com Rouge Ponceau S (Ponceau S 0,1%, Ácido acético 5%) por 5 min, sob agitação. Após visualização do padrão de transferência, foi realizada a etapa de bloqueio de ligações inespecíficas com solução caseína 5% em TBS-T (Tris 1M pH 7,6; NaCl 0,8%, Tween-20 0,1%), com incubação por 1 h, sob agitação, em TA. Após o bloqueio da membrana, foram realizadas duas lavagens com TBS-T, a primeira de 10 min e a segunda de 5 min.

À membrana, foi então adicionada solução de anticorpo primário contra a proteína de interesse diluído em TBS-T 0,1%, em concentração variando entre 1:350 a 1:1000. A incubação com cada anticorpo primário foi feita por 16 h (*overnight*), sob agitação constante, à temperatura de 4°C.

Após a incubação com o anticorpo primário, a membrana foi lavada duas vezes com TBS-T 0,1%, por 10 e 5 min, e em seguida incubada por 2:30 h, sob agitação e em TA, com solução de anticorpo secundário conjugado a HRP (diluído 1:2500 em TBS-T 0,1%).

Foram realizadas mais duas lavagens da membrana com TBS-T 0,1%, de 5 min cada. A detecção da expressão proteica se deu por quimioluminescência, com exposição da membrana às soluções de peróxido de hidrogênio e luminol do kit *Pierce® ECL Western Blotting Substrate* (Thermo Scientific), em proporção 1:1. Estas soluções reagem com o HRP

conjugado ao anticorpo secundário e emitem luz, que foi captada pelo *software* Image Studio Digits v 3.1, através do instrumento C-Digit LI-COR (Uniscience).

#### 4.5 Avaliação de localização celular

##### *4.5.1 Imunofluorescência (IF)*

Para os ensaios de imunofluorescência, realizou-se o plaqueamento das células MCF-7 e MDA-MB-231. A linhagem celular MCF-7 foi plaqueada em concentração de  $9,5 \times 10^3$  células/cm<sup>2</sup>, em lamínulas de vidro (Electron Microscopy Sciences) tratadas por 30 min com solução de Poli-L-lisina (Sigma-Aldrich) diluída em água milliQ em concentração 1:10. A linhagem MDA-MB-231, por sua vez, foi plaqueada em concentração de  $11,29 \times 10^3$  células/cm<sup>2</sup>, em câmaras de 10 poços *Cellview™ Slide* (Greiner Bio-One). As células foram plaqueadas como descrito no subtópico 4.1, e foi realizada troca do meio em 24 h após o plaqueamento.

Decorridas 48 h do plaqueamento, as células foram lavadas com PBS 1X. Para fixação da linhagem MCF-7, utilizou-se solução de metanol 70% (metanol 100% diluído em PBS 1X) por 20 min a -20°C, e, para fixação da linhagem MDA-MB-231, utilizou-se solução de paraformaldeído 4% (paraformaldeído diluído em PBS 1X) por 20 min em TA. Após a fixação, foram realizadas três lavagens com PBS 1X por 5 min cada, sob leve agitação. As lamínulas com células MCF-7 fixadas foram armazenadas a -20°C, enquanto as câmaras com células MDA-MB-231 foram armazenadas a 4°C em PBS 1X até a realização das próximas etapas.

As células foram lavadas com PBS 1X por 10 min sob leve agitação, e permeabilizadas com solução Triton X-100 0,3% em PBS 1X por 10 min, também sob leve agitação. Em seguida, realizou-se o bloqueio por 1 h com solução BSA 1% Triton X-100 0,1% em PBS 1X, sob agitação. Após o bloqueio, foram realizadas duas lavagens de 10 min cada com Triton X-100 0,3% em PBS 1X.

Aplicou-se anticorpo primário para as proteínas de interesse, diluído em BSA 0,1% Tween-20 0,05% em PBS 1X, em concentrações variando de 1:10 a 1:50. Os anticorpos primários foram incubados por 72 h a 4°C em ambiente úmido e protegido da luz.

Após a incubação com o anticorpo primário, as células foram lavadas quatro vezes com Triton X-100 0,3% em PBS 1X por 2 min, sob leve agitação. Em seguida, adicionou-se

anticorpo secundário marcado com fluoróforo (concentração de 1:500, diluído em BSA 0,1% Tween-20 0,05% em PBS 1X), com incubação de 2 h em TA, protegido da luz.

Após, foram realizadas cinco lavagens de 2 min cada com Tween-20 0,05% em PBS 1X, sob agitação leve. As lamínulas e câmaras utilizadas tiveram o excesso de solução removido. Em seguida, as lamínulas foram invertidas sobre lâminas contendo uma gota de meio de montagem com DAPI (*Vectashield® mounting medium with DAPI*, Vector Labs). As câmaras, por sua vez, tiveram o *frame* de delimitação dos poços removido, seguido de adição de uma gota do meio de montagem com DAPI em cada poço e colocação de uma lamínula.

As imagens foram adquiridas utilizando o microscópio confocal Leica TCS SPE (Leica Microsystems). Como controle, foram realizadas incubações apenas (1) com os anticorpos secundários e (2) com meio de montagem. Os valores de co-localização foram obtidos por meio do software do microscópio confocal, sob os parâmetros *background* 30% e *threshold* 30%.

#### 4.6 Silenciamento gênico de ASXL2

Para a redução dos níveis de RNAm e proteína de ASXL2 nas linhagens celulares MCF-7 e MDA-MB-231, utilizou-se o método de silenciamento gênico por RNA de interferência (siRNA). Esta abordagem baseia-se na capacidade de pequenos RNAs de fita dupla (dsRNAs) em reconhecer sequências específicas (complementares) de RNAm, levando à sua clivagem.

As linhagens celulares foram plaqueadas em placas de 12 poços, sob concentração de  $9,5 \times 10^3$  células/cm<sup>2</sup> para MCF-7 e  $9 \times 10^3$  células/cm<sup>2</sup> para MDA-MB-231, e mantidas sob as condições descritas em 4.1.

Para avaliar efeitos de toxicidade celular, foi necessário definir a concentração ideal do reagente de transfecção Lipofectamine® 3000 (Thermo Fisher) a ser utilizada. Para tal, as células foram submetidas ao procedimento de transfecção na ausência de siRNA (*mock*). Após 24 h do plaqueamento, em confluência de 50%, as células foram cultivadas com meio de transfecção (RPMI-1640, SBF 6%, Glutamina 1%) e tratadas com curva crescente do reagente de transfecção (1,5; 2,0; 2,5; 3,0 µg). Para fins de comparação, poços foram cultivados apenas com meio de transfecção (controle).

Após 24 h do tratamento, as células foram tripsinizadas conforme descrito em 4.1 e ressuspendidas em 500 µL de PBS 1X e sua densidade foi obtida após contagem em câmara de Neubauer. A viabilidade das mesmas foi avaliada em citômetro de fluxo BD

FACScalibur™ (BD Biosciences), no canal FL3 (670LP), através da adição de 1 µL de iodeto de propídio (PI, 50 µg/µL em PBS 1X pH 7,4, Sigma-Aldrich) a cada 1x10<sup>5</sup> células. Foram avaliados 10.000 eventos de cada condição, e este experimento foi realizado em triplicata.

Posteriormente, foi necessário definir a concentração de siRNA a ser utilizada para obter uma eficiência de transfecção maior que 80%. Para isto, as células foram cultivadas nas seguintes condições: (1) apenas com meio de transfecção (controle); (2) com meio de transfecção e reagente de transfecção (*mock*); e (3) com meio de transfecção, reagente de transfecção e 50 nM de siRNA (Fluorescein Conjugate)-A (sc-36869, Santa Cruz). Este siRNA é inespecífico e conjugado a fluoresceína. A eficiência de transfecção foi avaliada 24 h após tratamento, através da detecção de fluorescência no canal FL1 (530/30) no mesmo citômetro de fluxo e em conjunto com a viabilidade celular, como descrito anteriormente.

Para o silenciamento transiente de *ASXL2*, as células foram cultivadas em três condições: (1) apenas com meio de transfecção (controle); (2) com meio de transfecção, reagente de transfecção e siRNA scramble (sc-37007, Santa Cruz); e (3) com meio de transfecção, reagente de transfecção e siRNA para *ASXL2* (Hs\_ASXL2\_6 FlexiTube siRNA, Qiagen). Após 24 h do tratamento, o meio de transfecção foi retirado e todas as células foram incubadas com meio de cultura conforme descrito em 4.1.

Para avaliação da taxa de silenciamento, os RNAs totais das condições controle, *scramble* e siRNA *ASXL2* foram extraídos nos tempos de 24, 48 e 72 h após o silenciamento. A extração foi feita com o kit *RNeasy* (Qiagen), conforme protocolo descrito em 4.3.1. Ao final da extração, foi realizada eluição com 25 µL de H<sub>2</sub>O RNase *free*. Os RNAs foram quantificados, tratados com DNase e submetidos à transcrição reversa, como descrito em 4.3. Os cDNAs resultantes foram utilizados para avaliação da expressão gênica de *ASXL2* por RT-qPCR nos diferentes tempos de avaliação do silenciamento.

Para confirmação do silenciamento a nível proteico, foi realizada a extração de proteínas totais seguida de WB, conforme descrito em 4.4.

#### 4.7 Análise do perfil de transcritos por microarranjo

Com a finalidade de verificar possíveis genes envolvidos com a atividade de *ASXL2* nas linhagens MCF-7 e MDA-MB-231, realizou-se uma avaliação do perfil global de transcritos nos modelos celulares silenciados (si*ASXL2*), em comparação com os modelos não silenciados (*scramble*). Esta análise se deu pela técnica de microarranjo, na qual transcritos

podem ser identificados e quantificados por complementaridade de bases através da interação com sondas presentes em uma lâmina ou *chip*. Este experimento foi realizado com n=1.

Previamente à realização do experimento de microarranjo, a integridade e qualidade dos RNAs provenientes dos modelos celulares em ambas as condições (*siASXL2* e *scramble*) foram avaliadas com o *kit* RNA 6000 Nano Chip (Agilent Technologies), seguindo o protocolo do fabricante, e analisadas com o equipamento BioAnalyzer 2100 (Agilent Technologies). Neste *kit*, as amostras de RNA são separadas por eletroforese e detectadas por fluorescência, permitindo a separação e visualização das bandas ribossomais 18S e 28S. Um valor de RIN (*RNA Integrity Number*) é gerado, levando em conta a razão entre bandas ribossomais, e indica a integridade do RNA analisado. Os valores de RIN variam de 1 a 10, sendo 10 o valor de referência para amostras com alta integridade.

Neste trabalho, foi utilizado o *kit* de microarranjo Two-Color Microarray-Based Gene Expression Analysis (Agilent Technologies), em protocolo conforme instruções do fabricante. Para a realização do microarranjo, utilizou-se, além de 200 ng de RNA das linhagens celulares nas condições de interesse, um RNA de referência, fornecido no *kit*. Com o auxílio do *kit* Low Input Quick Amp Labeling (Agilent Technologies), os RNAs foram marcados (RNA referência marcado com Cy3, RNAs das amostras marcados com Cy5) e utilizados em uma etapa de síntese de cDNA, seguida de transcrição e amplificação, a fim de se obter o cRNA. Os cRNAs foram purificados em colunas RNeasy (Qiagen) e quantificados no Espectrofotômetro ND-1000 (Nanodrop Technologies), no comprimento de onda de 260 nm. Em seguida, foram submetidos à hibridização em lâmina HD Human GE 4x44K V2 (Agilent Technologies), por 17 h a 65°C à rotação de 60 rpm, no forno de hibridização SHO1 (Shel Lab). A lâmina de microarranjo foi então lavada com soluções presentes no *kit* e digitalizada no SureScan Microarray Scanner (Agilent Technologies). Para a obtenção das listas brutas de genes identificados e seus respectivos valores de expressão, foi utilizado o *software* Agilent Feature Extraction (Agilent Technologies), e em seguida os dados foram analisados no *software* GeneSpring GX 11.5 (Agilent Technologies). As listas de genes geradas foram avaliadas com um critério de *fold change* de 2,0 com o auxílio dos *softwares online* Metacore® (Clarivate Analytics, disponível em [portal.genego.com](http://portal.genego.com)), String (String Consortium, versão 10.5, disponível em [string.db.org](http://string.db.org)), em intervalos de média (0.400) e alta (0.900) confiança, e PANTHER (versão 14.0, disponível em [pantherdb.org](http://pantherdb.org)).

#### 4.8 Ensaio de migração celular

Para a avaliação da contribuição de ASXL2 na migração celular, realizou-se um ensaio de *wound healing*, no qual as linhagens celulares MCF-7 e MDA-MB-231 foram plaqueadas em placas de 12 poços e silenciadas para ASXL2 por 72 h.

Após o silenciamento, fez-se um risco em cada poço com uma ponteira de 200 µL, a fim de criar uma "fenda" entre as células. Em seguida, as células foram lavadas com PBS 1X e mantidas em meio privado de SBF (RPMI SBF 1%) por 24 h, a fim de inibir a proliferação celular. As células foram fotografadas no momento do risco (0 h) e 24 h após com o auxílio do microscópio Axio Observe.Z1, equipado com uma câmera AxioCam HRc (ZEISS), e a distância entre as bordas foi calculada para mensurar a migração celular, sendo a distância em 0 h considerada como 100%. Este experimento foi realizado em triplicata, e como controle, foram utilizadas células *scramble*.

#### 4.9 Contagem de microvesículas

A fim de avaliar a contagem de microvesículas extracelulares presentes nos sobrenadantes das linhagens silenciadas e *scramble*, realizou-se o silenciamento de ASXL2 por 72 h em placas de 24 poços, conforme descrito anteriormente em 4.6. Após as 72 h de silenciamento, o meio de cultura foi removido e substituído por meio RPMI 0% SBF. As células silenciadas permaneceram por 24 h em cultura, e seus sobrenadantes foram coletados e centrifugados a 6000 x g por 20 min a 4 °C, e em seguida congelados a -80 °C.

A contagem das microvesículas foi realizada em alíquotas de 1 mL do sobrenadante com o auxílio do instrumento ZetaVIEW® S/N 205, e do *software* ZetaView 8.02.31 (Particle Metrix), sob os seguintes parâmetros de análise: tamanho máximo 500 nm, tamanho mínimo 5 nm, brilho mínimo 80.

#### 4.10 Ensaio de ciclo celular

A fim de avaliar o efeito de ASXL2 no ciclo celular, as linhagens celulares MCF-7 e MDA-MB-231 foram submetidas ao silenciamento gênico de ASXL2, e após 72 h, foram tripsinizadas conforme descrito em 4.1. Uma densidade de  $7,5 \times 10^5$  células foi marcada com solução contendo PI (Citrato de Sódio 0,1%, Triton X-100 0,3%, Iodeto de Propídio 0,01%, Ribonuclease A 0,002%) e incubada por 30 min a 4°C. Em seguida, a marcação por PI foi

medida em citômetro de fluxo BD FACScalibur™ (BD Biosciences), no canal FL3 (670LP), para avaliação do ciclo celular. Foram avaliados 10.000 eventos de cada condição, incluindo células controle e *scramble*.

#### 4.11 Imunoprecipitação de proteína (IP)

##### *4.11.1 Lise celular e extração de proteínas totais para imunoprecipitação*

A extração de proteínas para o experimento de imunoprecipitação foi feita a partir dos citocentrifugados das linhagens celulares MCF-7 e MDA-MB-231. Uma densidade de  $1 \times 10^7$  células foi ressuspensa, após duas lavagens com PBS 1X estéril, em 250  $\mu$ L de tampão de lise *high salt* (NaCl 500 mM, NP40 1%, Tris 50 mM pH 8,0, mix inibidor de proteases 1% - GE Healthcare Bio-Sciences) em um microtubo (1,5 mL). Após homogeneização, os microtubos (1,5 mL) contendo os lisados celulares foram submersos em nitrogênio líquido, descongelados em TA, e em seguida centrifugados a 12000 x g, por 15 min, em temperatura de 4°C. Os sobrenadantes, correspondentes aos extratos de proteína em tampão *high salt* (HS), foram transferidos para novos microtubos (1,5 mL), identificados e armazenados à temperatura de -80°C, até sua quantificação.

##### *4.11.2 Imunoprecipitação*

Foi realizado o *pre-clearing* dos extratos HS. Para tal, adicionou-se 1  $\mu$ g de IgG normal de coelho para cada 1 mg de extrato de proteína HS em um microtubo (1,5 mL), com incubação por 1h a 4°C, sob agitação orbital. Em seguida, adicionou-se 50  $\mu$ L de proteína A-Sepharose® (GE Healthcare) diluída em tampão *high salt*, e os extratos HS foram incubados sob as mesmas condições. Os extratos HS foram centrifugados a 14000 x g, por 15 min, em temperatura de 4°C, e os sobrenadantes (extratos HS *pós-clearing*) foram coletados para outros microtubos (1,5 mL). Para aumentar a recuperação, as *beads* de proteína A foram lavadas com 60  $\mu$ L de tampão *high salt*, seguida de nova centrifugação e coleta dos sobrenadantes para o mesmo microtubo (extratos HS *pós-clearing*).

Aos extratos HS *pós-clearing*, adicionou-se 2  $\mu$ g de anticorpo anti-ASXL2, e para o controle, 2  $\mu$ g de anticorpo IgG de coelho. Os anticorpos foram incubados por aproximadamente 16 h a 4°C, sob agitação orbital. Em seguida, a cada microtubo, foram



adicionados 50  $\mu$ L de proteína A diluída em tampão *high salt*, com incubação de 4h a 4°C, sob agitação orbital.

Após a incubação com a proteína A, foi realizada uma centrifugação a 14000 x g, por 1 min, a 4°C, seguida do descarte do sobrenadante. Foram feitas duas lavagens com 1 mL de tampão *high salt*, com centrifugação a 14000 x g, por 20 seg, a 4°C; e uma lavagem com Bicarbonato de Amônio 50 mM pH 8,0, com centrifugação a 14000 x g, por 10 seg, a 4°C.

Os imunoprecipitados foram eluídos das *beads* de proteína A com 250  $\mu$ L de Hidróxido de Amônio 0.5 M pH 11, sob agitação a 1300 rpm, 4°C, por 15 min, seguida de centrifugação 14000 x g, por 30 seg, a 4°C. Os sobrenadantes foram coletados e transferidos para microtubos (1,5 mL) identificados. A etapa de eluição foi realizada 3 vezes, totalizando um volume final de 750  $\mu$ L de imunoprecipitado para cada condição. Foram realizados três experimentos independentes de imunoprecipitação (n=3) para cada linhagem celular.

#### 4.12 Análise proteômica

##### *4.12.1 Troca de tampão e concentração das amostras em coluna Amicon*

Os extratos de proteína correspondentes aos imunoprecipitados foram agrupados e submetidos à troca de tampão e concentração por coluna de exclusão. Para tal, foram utilizadas as colunas filtrantes Amicon® Ultra-2 3K (Millipore).

Inicialmente, foi realizada a hidratação da membrana de cada coluna com a adição de 1 mL de água LC/MS (Merck), descartada por inversão do tubo. Em seguida, realizou-se uma etapa de ativação da membrana com 1 mL de Bicarbonato de Amônio 200 mM pH 8,0 e centrifugação a 4000 x g por 15 min, em temperatura de 16°C, com posterior descarte do filtrado.

As amostras de imunoprecipitados, em sua totalidade, foram adicionadas às colunas, ajustando-se o volume para 2 mL com Bicarbonato de Amônio 50 mM pH 8,0, e submetidas a centrifugação a 4000 x g, por 1 h a 16°C. Os filtrados foram descartados e as colunas foram invertidas, conforme protocolo, e centrifugadas a 1000 x g por 2 min a 16°C para a recuperação dos concentrados proteicos. As amostras concentradas foram transferidas para um microtubo (1,5 mL) de baixa retenção (Axygen) e em seguida quantificadas pelo método de Bradford, conforme item 4.3.3, utilizando-se Bicarbonato de Amônio 50 mM pH 8,0 como diluente no lugar de água destilada, tanto para a curva padrão quanto para as amostras, e com um tempo de incubação de 20 min com o reagente de Bradford.

#### 4.12.2 Digestão enzimática das proteínas

Após a etapa de concentração das proteínas, fez-se necessária uma etapa de digestão enzimática das mesmas. Para isso, as amostras concentradas tiveram seu volume ajustado para 104  $\mu\text{L}$  com Bicarbonato de Amônio 50 mM pH 8,0. Em seguida, as amostras foram incubadas com 25  $\mu\text{L}$  de solução surfactante *RapiGest SF* (Waters) a 80°C por 15 min, em uma etapa de solubilização das proteínas.

Posteriormente, as proteínas foram reduzidas com 2,5  $\mu\text{L}$  do agente redutor de pontes dissulfeto Ditioneitol (DTT) 100mM (GE), a 60°C por 30 min, e em seguida alquiladas com 2,5  $\mu\text{L}$  de Iodoacetamida (IAA) 300mM, por 30 min em TA, protegido de luz.

Para a digestão das proteínas em peptídeos, as amostras foram incubadas com 16  $\mu\text{L}$  de solução tripsina 0,25  $\mu\text{g}/\mu\text{L}$  (*Sequencing Grade Modified Trypsin, Porcine*, Promega), diluída em Bicarbonato de Amônio 50 mM pH 8,0, a 37°C sob agitação (600 rpm), por um período de aproximadamente 16 h (*overnight*). Após esta incubação, o *RapiGest* foi precipitado através da adição de 10  $\mu\text{L}$  de solução Ácido Trifluoroacético 5%, seguida de incubação a 37°C por 1:30 h (600 rpm). As amostras foram então centrifugadas a 16000 x g por 30 min, à temperatura de 6°C, e os sobrenadantes foram transferidos para microtubos (1,5 mL, Axygen) identificados, sendo armazenados a -80°C até a etapa seguinte.

#### 4.12.3 Cromatografia líquida em nano escala

Às amostras, adicionou-se solução 5 mM Formiato de Amônio ( $\text{NH}_4\text{FA}$ ), 5% Acetonitrila (ACN) para um volume final de 160  $\mu\text{L}$ , e 40  $\mu\text{L}$  de padrão interno ADH (*yeast alcohol dehydrogenase*; P00330, Waters) 1 pmol/ $\mu\text{L}$ , para uma concentração final de 0,2 pmol/ $\mu\text{L}$ . Em seguida, as amostras foram submetidas a separação por cromatografia líquida em nano escala (2DNanoLC), utilizando o equipamento NanoAcquity (Waters).

Para a primeira dimensão, foram utilizadas colunas de troca catiônica forte de 300 Å (*strong cation exchange - SCX*) (Waters) empacotadas com 5  $\mu\text{m}$  de PolySULFOETHYL Aspartamide (PolyLC, Columbia, MD). 5  $\mu\text{g}$  de amostra (peptídeos) foram carregados à coluna, e cada amostra foi injetada em quintuplicata.

As amostras foram eluídas em 10 frações a partir da coluna SCX utilizando um gradiente de sal (3  $\mu\text{L}$  de plugs de sal contendo 50, 100, 150, e 200 mM de  $\text{NH}_4\text{FA}$  com ACN 5%), seguido por um gradiente de 200 mM de  $\text{NH}_4\text{FA}$  (3  $\mu\text{L}$  de plugs orgânicos de 10, 20, e 30% de ACN), e para lavagem, 2 plugs de  $\text{NH}_4\text{FA}$  350 mM e de ACN 50%. Os peptídeos

liberados foram capturados por uma coluna *trap* (180  $\mu\text{m}$   $\times$  20 mm, 5  $\mu\text{m}$  Symmetry C18) (Waters). Após a captura dos peptídeos, a coluna *trap* foi ligada, *online*, a uma coluna analítica de fase reversa (75  $\mu\text{m}$   $\times$  150 mm, 1.8  $\mu\text{m}$  C18, nanoACQUITY UPLC HSS T3) (Waters). Como segunda dimensão, foi usado um gradiente de fase reversa de 5–40% de ACN (contendo 0.1% v/v de Ácido Fórmico) em 80 min, com uma taxa de fluxo de 600 nl/min.

#### 4.12.4 Espectrometria de massas

Após serem eluídos da coluna, os peptídeos foram ionizados utilizando *nanoelectrospray* em modo positivo (nanoESI+) em uma fonte de ionização NanoLockSpray (Waters), acoplada ao cromatógrafo NanoAcquity. Como referência, utilizou-se a solução de GFP ([Glu1]- fibrinopeptídeo B humano, Sigma-Aldrich), entregue pelo NanoLockSpray a cada 30 seg. O íon duplamente carregado ( $[M + 2H]^{2+} = 785.8426$ ), gerado por MS/MS, foi usado como referência para monitorar a intensidade do sinal.

A aquisição dos dados foi realizada no modo MSE de análise independente de dados (Data-independent analyses - DIA), sendo o escaneamento realizado com especificidade e seletividade em modo de mobilidade iônica "T-wave" não linear (HDMSE).

Os experimentos foram realizados utilizando o espectrômetro de massas de alta definição Synapt HDMS (Waters), em analisador do tipo tempo de voo (*Time-of-Flight*) com sistema de aceleração ortogonal (oa-TOF). Os espectros foram adquiridos em intervalo de massa/carga ( $m/z$ ) entre 50 a 2000. O perfil MS foi ajustado para garantir que os dados LC-MS gerados em baixa energia fossem efetivamente adquiridos num intervalo  $m/z$  de 400 a 2000, assegurando que todas as massas inferiores a  $m/z$  400 no LC/MSE surgiram de dissociações na célula de colisão.

#### 4.12.5 Identificação de proteínas

Para o processamento dos espectros, busca em banco de dados e identificação, utilizou-se o *software* ProteinLynx Global Server v.3.0.2 (PLGS), com a ferramenta ExpressionE informatics. O banco de dados UniProtKB (release 2018\_07) foi utilizado com proteínas anotadas.

A identificação de proteínas foi gerada por algoritmos específicos para procura em banco de dados humano. Os bancos de dados utilizados foram randomizados *on the fly*

durante a pesquisa para acessar a taxa de falsos positivos (FDR). Foram considerados aceitáveis os valores de FDR de 1% para proteínas e de 0,1% para peptídeos. As condições de busca foram baseadas em taxonomia (*Homo sapiens* [humano]), obedecendo aos critérios: enzima utilizada (tripsina) com no máximo 1 sítio de clivagem perdido; carbamidometilação de cisteínas (Cys) e oxidação da metionina (Met) como modificações variáveis.

Para a análise das proteínas identificadas e posterior comparação entre as condições (ASXL2 vs. IgG), foram avaliadas três replicatas experimentais.

#### 4.12.6 Análise *in silico* das proteínas identificadas

As proteínas identificadas em 2/3 ou 3/3 replicatas, em cada condição, foram organizadas em forma de listas, que por sua vez foram submetidas à comparação por diagrama de Venn, através do *software* da Bioinformatics Evolutionary Genomics (disponível em [bioinformatics.psb.ugent.be/webtools/Venn/](http://bioinformatics.psb.ugent.be/webtools/Venn/)). Para as análises posteriores, foram consideradas as proteínas identificadas exclusivamente com a imunoprecipitação de ASXL2, ou seja, foram excluídas quaisquer redundâncias ou similaridades com o controle (IgG).

Em seguida, a lista de proteínas imunoprecipitadas com o anticorpo ASXL2 foi convertida em uma lista de genes, com o auxílio do *software* Retrieve/ID mapping, da UniProt (disponível em [uniprot.org/uploadlists/](http://uniprot.org/uploadlists/)). Este *software* gera, a partir dos códigos de acesso das proteínas identificadas, uma lista de genes correspondentes. Genes identificados a partir de proteínas não revisadas foram desconsiderados em nossa análise.

Para a análise de cada gene identificado quanto à sua função, localização e expressão, dentre outros, foram utilizados o banco de dados GeneCards® (Weizman Institute of Science, versão 4.7.1, disponível em [genecards.org](http://genecards.org)) e o *software* String (String Consortium, versão 10.5, disponível em [string.db.org](http://string.db.org)).

#### 4.13 Imunoprecipitação de cromatina (ChIP)

A imunoprecipitação de cromatina (*chromatin immunoprecipitation*, ChIP) foi realizada utilizando o kit *SimpleChIP*® *Enzymatic Chromatin IP kit (Magnetic Beads)* (Cell Signaling Technology, EUA), conforme especificação do fabricante.

Uma quantidade equivalente a  $4 \times 10^6$  células da linhagem MCF-7 foi cultivada, fixada com 270  $\mu$ L de formaldeído 37% (Sigma Aldrich) por 10 min em TA, e em seguida incubada

com 1 mL de Glicina 10X (concentração final de 1X) por 5 min em TA. Após fixação, as células foram lavadas duas vezes com PBS 1X gelado.

Às células, adicionou-se 2 mL de PBS 1X gelado e 10  $\mu$ L de coquetel inibidor de protease 200X (PIC 200X). As células foram raspadas da garrafa com uma haste de plástico (*cell scraper*), transferidas para um tubo de propileno de 15 mL e centrifugadas a 1500 rpm por 5 min a 4°C. O sobrenadante foi removido e o citocentrifugado foi ressuscitado em 1 mL de Tampão A 1X gelado, acrescido de 0,5  $\mu$ L DTT 1M e 5  $\mu$ L PIC 200X, e incubado por 10 min no gelo (0°C), com leve homogeneização da suspensão celular a cada 3 min.

Após incubação, a suspensão foi centrifugada a 3000 rpm por 5 min, a 4°C, para precipitação dos núcleos. O sobrenadante foi descartado e o *pellet* resultante foi ressuscitado em 1 mL de tampão B 1X gelado, acrescido de 0,5  $\mu$ L de DTT 1M. Em seguida, repetiu-se a centrifugação, com o descarte do sobrenadante e a ressuspensão do precipitado em 100  $\mu$ L de tampão B 1X gelado, acrescido de 0,05  $\mu$ L de DTT.

Para a etapa de digestão da cromatina, foram adicionados 5  $\mu$ L da *Micrococcal nuclease* (2000 units/ $\mu$ L) diluída 1:10 em tampão B a cada 100  $\mu$ L da cromatina preparada, com incubação de 10 min à 37°C sob agitação de 1200 rpm. Para a inativação da enzima de digestão, adicionou-se 10  $\mu$ L de EDTA 0,5 M, seguido de centrifugação a 13000 rpm por 1 min a 4°C. O sobrenadante foi removido e o *pellet* ressuscitado em 100  $\mu$ L do tampão de eluição CHIP 1X, acrescido de 0,5  $\mu$ L de PIC 200X, permanecendo por 10 min no gelo (0°C). Para verificação do padrão da digestão da cromatina, foi conduzida uma corrida em gel de agarose 1%, e comparação com padrão de 100 pb.

Foram conduzidas as imunoprecipitações de ASXL2, H3K27me3, e como controle, IgG de coelho (anticorpo presente no *kit* de imunoprecipitação). As imunoprecipitações foram realizadas com 5  $\mu$ g da cromatina digerida, completando-se o volume para 500  $\mu$ L com tampão CHIP 1X + PIC 200X. Um volume de 20  $\mu$ L (4% da amostra) foi retirado de cada microtubo (1,5 mL) e armazenado a -80°C para uso como controle do DNA não imunoprecipitado (*input*). Aos volumes restantes, foram adicionados 2  $\mu$ g de anticorpo, em cada condição. As amostras foram então incubadas por 16 h (*overnight*) a 4°C, sob agitação orbital.

Após a incubação com os anticorpos, foram adicionados a cada reação de imunoprecipitação um volume de 30  $\mu$ L de *beads* magnéticas de proteína G (ChIP-Grade Protein G Magnetic Beads), que se ligam aos anticorpos. As *beads* magnéticas foram incubadas por 2 h a 4°C, sob agitação orbital, e em seguidas lavadas três vezes com 1 mL de solução de lavagem com baixo teor de sal (100  $\mu$ L tampão CHIP 10X e 900  $\mu$ L de água). Cada

incubação foi realizada por 5 min a 4°C em agitação orbital, e para a recuperação das *beads* entre lavagens, foi utilizada uma estante magnética (Magna Step, Thermo Fisher).

A seguir, foi realizada uma lavagem com 1 mL de solução de alto teor de sal (100 µL tampão ChIP 100 X, 900 µL de água e 70 µL NaCl 5 M), também por 5 min, a 4°C, sob agitação orbital. Para eluir os imunoprecipitados das *beads* magnéticas, foram acrescentados 150 µL de tampão de eluição ChIP (75 µL ChIP elution buffer 2X e 75µL água) a cada microtubo, com posterior incubação por 30 min a 65°C, sob agitação de 1200 rpm. O mesmo procedimento foi realizado com a amostra de *input* 4% que havia sido previamente armazenada.

Ao fim da incubação, as *beads* magnéticas foram coletadas com o auxílio da estante magnética, e o sobrenadante, contendo os complexos cromatina + anticorpo, foi transferido cuidadosamente para um microtubo (1,5 mL) identificado.

Os produtos da imunoprecipitação foram concentrados a vácuo com o auxílio do instrumento *Concentrator 5301* (Eppendorf), para um volume final de 20 µL. A estes 20 µL, foram adicionados 10 µL de tampão de amostra 4X (descrito em 4.4.4), e em seguida os imunoprecipitados foram resolvidos em gel SDS-PAGE 15% para a realização de Western Blot. Junto a essas amostras, também foi resolvido em gel 30 µg de extrato de proteína nuclear da linhagem MCF-7.

O experimento de Western Blot dos imunoprecipitados de cromatina foi realizado como descrito em 4.4.4, com o anticorpo anti-H3K27me3, para confirmar a imunoprecipitação e avaliar a interação entre H3K27me3 e ASXL2 na linhagem celular de interesse.

#### 4.14 Análise estatística

As análises estatísticas das replicatas experimentais foram realizadas com o auxílio do software *GraphPad Prism 6*, utilizando o teste *t* de *Student*. Valores de  $P < 0,05$  foram considerados estatisticamente significativos. \* =  $p < 0,05$ . \*\* =  $p < 0,01$ . \*\*\* =  $p < 0,001$ . \*\*\*\*= $p < 0,0001$ . Barras de erro nos gráficos são referentes ao desvio padrão.

## 5. RESULTADOS

Uma vez que os subtipos moleculares de CM diferem entre si em expressão de receptores, prognóstico e agressividade, o estudo e descoberta de quaisquer outros aspectos genéticos e epigenéticos envolvidos podem trazer uma melhor classificação e entendimento desta doença, que representa um grande impacto na saúde pública. Assim, torna-se importante explorar novos mecanismos, ainda pouco conhecidos, que podem contribuir para o CM, dentre eles a função dos ETPs. Por esta razão, escolheu-se por avaliar, neste trabalho, o potencial papel de ASXL2 na regulação epigenética do CM.

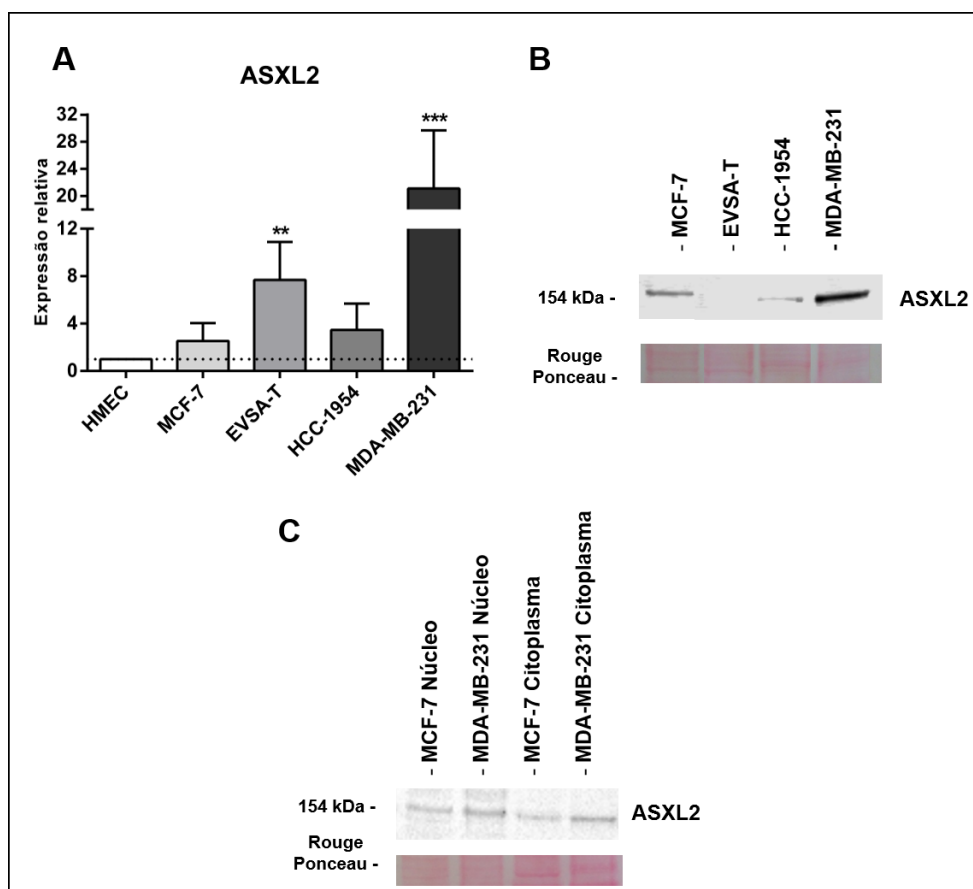
### 5.1 ASXL2 está diferencialmente expresso e localizado em modelos *in vitro* de CM

Uma análise inicial envolveu a avaliação da expressão de ASXL2 em linhagens celulares representativas dos subtipos moleculares da doença. Em análise por PCR quantitativo em tempo real, foi possível observar um aumento na expressão relativa de ASXL2 em todas as linhagens celulares de CM avaliadas, quando comparadas com a linhagem normal HMEC (Figura 5.1A). Em MCF-7 e HCC-1954 esta expressão é similar, com *fold change* próximo a 4, e em EVSA-T o *fold change* é de aproximadamente 9 ( $p < 0,01$ ). Na linhagem Triplo Negativa (MDA-MB-231), há um aumento mais expressivo na expressão de ASXL2, com um *fold change* de aproximadamente 20 ( $p < 0,001$ ).

Em nível de proteína, a avaliação por Western Blot permitiu verificar que todas as linhagens apresentam expressão de ASXL2, com exceção de EVSA-T (Figura 5.1B). As linhagens MCF-7 e MDA-MB-231 possuem a maior expressão de ASXL2 em níveis proteicos. A expressão de ASXL2 no modelo Luminal A corrobora com o observado por Park *et al.* (2016), e a expressão mais acentuada em MDA-MB-231 corrobora com os resultados de transcritos observados neste modelo.

O fato de ASXL2 estar expresso principalmente nos modelos Luminal A e TN chamou a atenção, pois estes modelos representam, respectivamente, os subtipos menos e mais agressivo da doença. Desta forma, focamos nosso estudo nestas linhagens, e assim optamos por investigar a localização celular desta proteína nestes dois modelos, através dos ensaios de Western Blot de extratos de proteínas nucleares e citoplasmáticas, e imunofluorescência. A análise da expressão proteica de ASXL2 nos extratos nucleares e citoplasmáticos das células por WB permitiu observar que ASXL2 está expresso na fração

nuclear de ambos os modelos (Figura 5.1C), e permitiu também confirmar que MDA-MB-231 possui maior expressão de ASXL2, comparando-se com MCF-7.



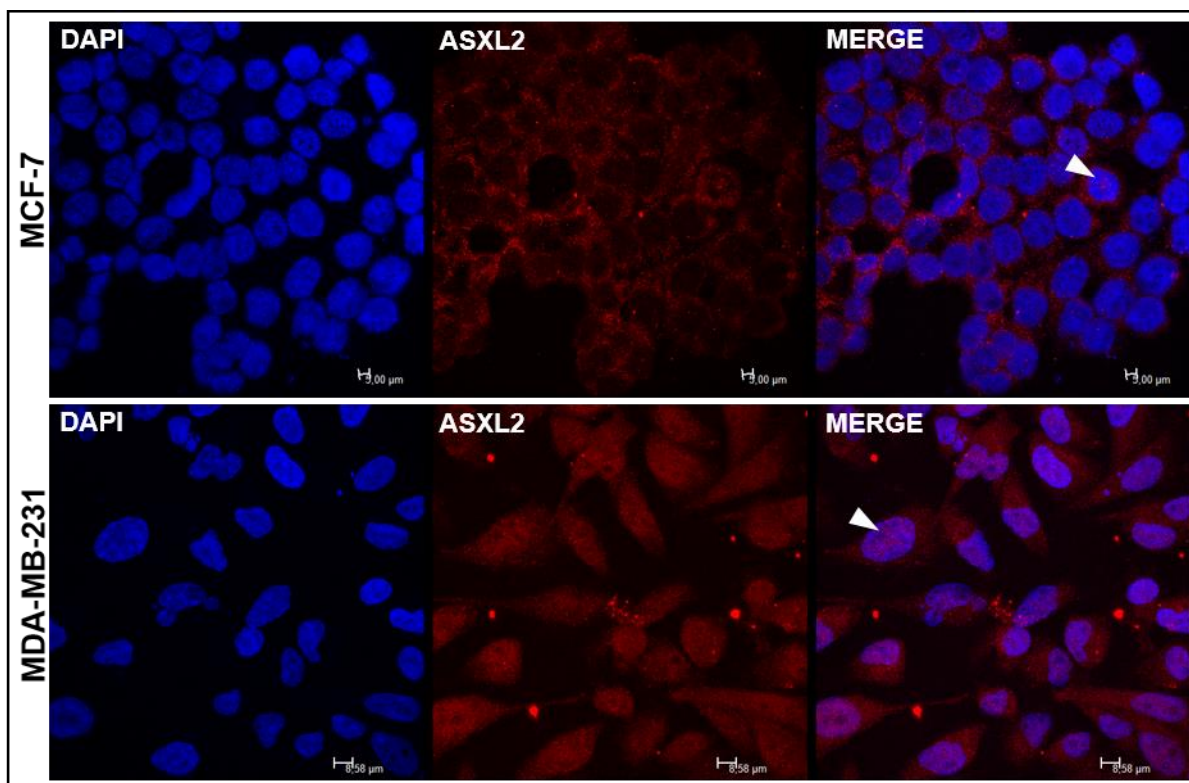
**Figura 5.1 – Expressão de ASXL2 em linhagens celulares de CM.** (A) Expressão gênica relativa do transcrito de *ASXL2* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH*, em linhagens celulares representativas dos subtipos de CM, comparadas pela expressão na linhagem normal HMEC (n=3). \*\*p<0,01. \*\*\*p<0,001. Barras de erro indicativas de desvio padrão. (B) Expressão proteica de ASXL2 em linhagens celulares representativas dos subtipos de CM. Foram resolvidos 30 µg de extratos proteicos totais das linhagens MCF-7, EVSA-T, HCC-1954 e MDA-MB-231 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-ASXL2 em diluição 1:1000. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau. (C) Expressão de ASXL2 em extratos de proteínas nucleares e citoplasmáticas de MCF-7 e MDA-MB-231. Foram resolvidos 30 µg de extratos proteicos de núcleo e citoplasma das linhagens MCF-7 e MDA-MB-231 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-ASXL2 em diluição 1:1000. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau.

A fim de refinar os dados de expressão gênica, realizamos o ensaio de imunofluorescência, no qual utiliza-se o marcador nuclear DAPI (4',6-diamidino-2-fenil-indol), que se liga ao DNA e emite luz azul quando excitado, permitindo discernir núcleo e citoplasma. Os resultados podem ser observados na Figura 5.2.

Na linhagem MCF-7, ASXL2 (em vermelho) encontra-se mais abundantemente no citoplasma, com focos de colocalização com DAPI no núcleo da célula (média de 38,66% de colocalização). Na linhagem MDA-MB-231, ASXL2 também está presente no citoplasma e



núcleo da célula, apresentando mais focos de colocalização com DAPI quando comparado com a linhagem MCF-7 (média de 53,78% de colocalização). Assim, nossos resultados sugerem que ASXL2 está diferencialmente expresso em CM.



**Figura 5.2 – Localização de ASXL2 em MCF-7 e MDA-MB-231.** Localização celular de ASXL2 por imunofluorescência em MCF-7 e MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpo específico para ASXL2, em diluição 1:20. Núcleos marcados em azul (DAPI) e ASXL2 em vermelho (ampliação 63x). Cabeça de seta: núcleo celular.

## 5.2 Modificadores de histona e remodeladores de cromatina encontram-se diferencialmente expressos e localizados em modelos *in vitro* de CM e podem estar colocalizados com ASXL2

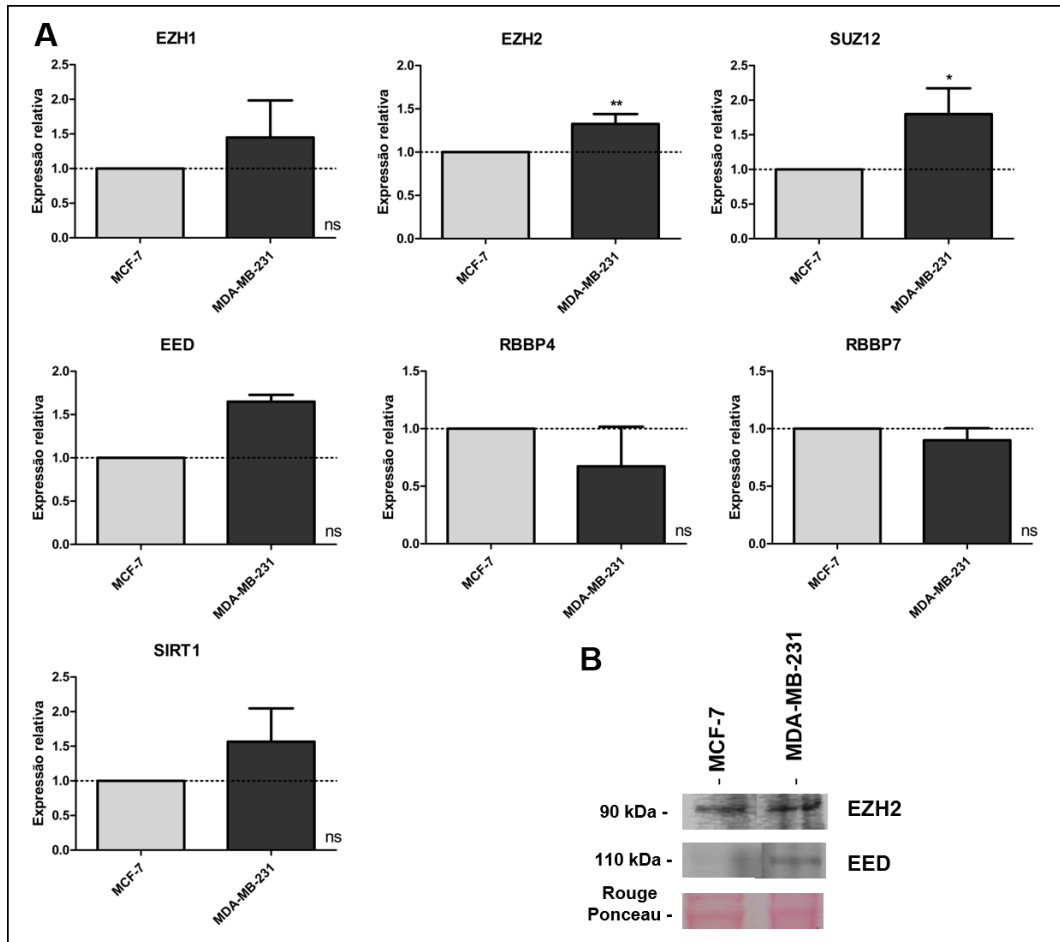
Uma vez que ASXL2 foi descrito como um potencial regulador e/ou parceiro de modificadores de histona e remodeladores de cromatina, a primeira pergunta a ser respondida seria sobre a expressão destas proteínas nos modelos de CM, além da sua localização celular e em conjunto com ASXL2.

A análise de membros das famílias Pc, Trx, SWI/SNF e NURF se deu por RT-qPCR em MCF-7 e MDA-MB-231, e, para alguns dos membros destas famílias, realizamos também os experimentos de WB e IF para avaliação da expressão proteica e localização celular, respectivamente.

Com relação à família Pc, avaliamos membros de seus dois complexos efetores, o PRC2 e o PRC1. Dentre os membros do complexo PRC2 (Figura 5.3), foi possível observar

um aumento significativo da expressão gênica de *EZH2*, um dos efetores do complexo, na linhagem MDA-MB-231. Esta expressão foi confirmada em nível de proteína em ambas as linhagens celulares, com maior intensidade em MDA-MB-231, confirmando os resultados de expressão gênica.

Outro membro com expressão gênica aumentada em MDA-MB-231 foi *SUZ12*, integrante *core* do complexo, que está cerca de 2 vezes mais expresso nesta linhagem quando comparado com MCF-7. A proteína EED, também um componente *core* do complexo PRC2, está mais expressa em MDA-MB-231. Os demais membros do complexo avaliados, dentre eles o efector *EZH1*, os mediadores de ligação ao nucleossomo *RBBP4* e *RBBP7* e a histona deacetilase *SIRT1*, tiveram pouca variação de expressão gênica entre os dois modelos estudados. Estes resultados sugerem que o complexo PRC2 poderia ser mais ativo na linhagem TN.

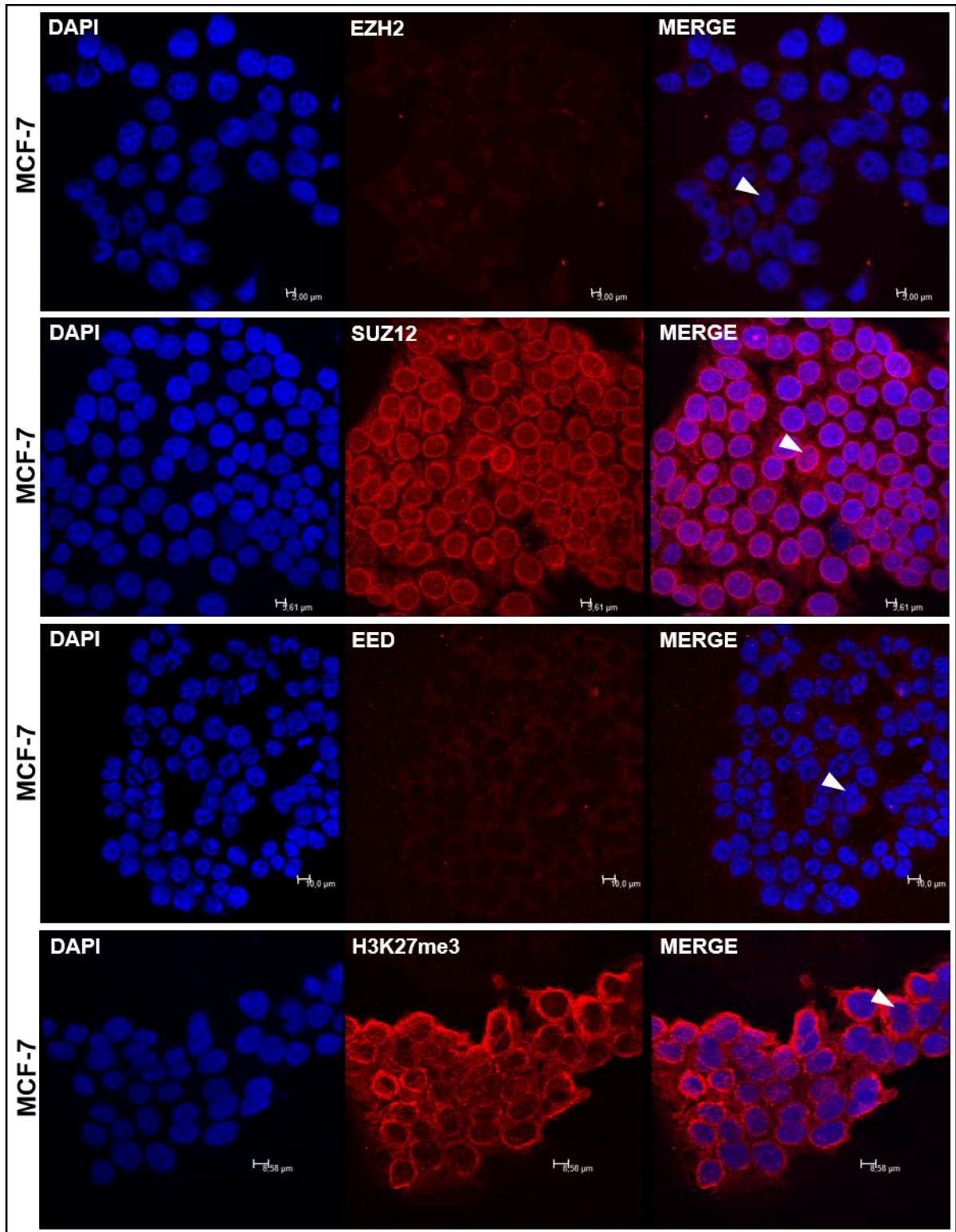


**Figura 5.3 – Expressão de membros de PRC2 em linhagens celulares de CM.** (A) Expressão gênica relativa dos transcritos de *EZH1*, *EZH2*, *SUZ12*, *EED*, *RBBP4*, *RBBP7* e *SIRT1* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). ns: diferença não significativa. \*p<0,05. \*\*p<0,01. Barras de erro indicativas de desvio padrão. (B) Expressão proteica de EZH2 e EED em MCF-7 e MDA-MB-231. Foram resolvidos 30 µg de extratos proteicos totais das linhagens MCF-7 e MDA-MB-231 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-EZH2 e anti-EED em diluição 1:500. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau.

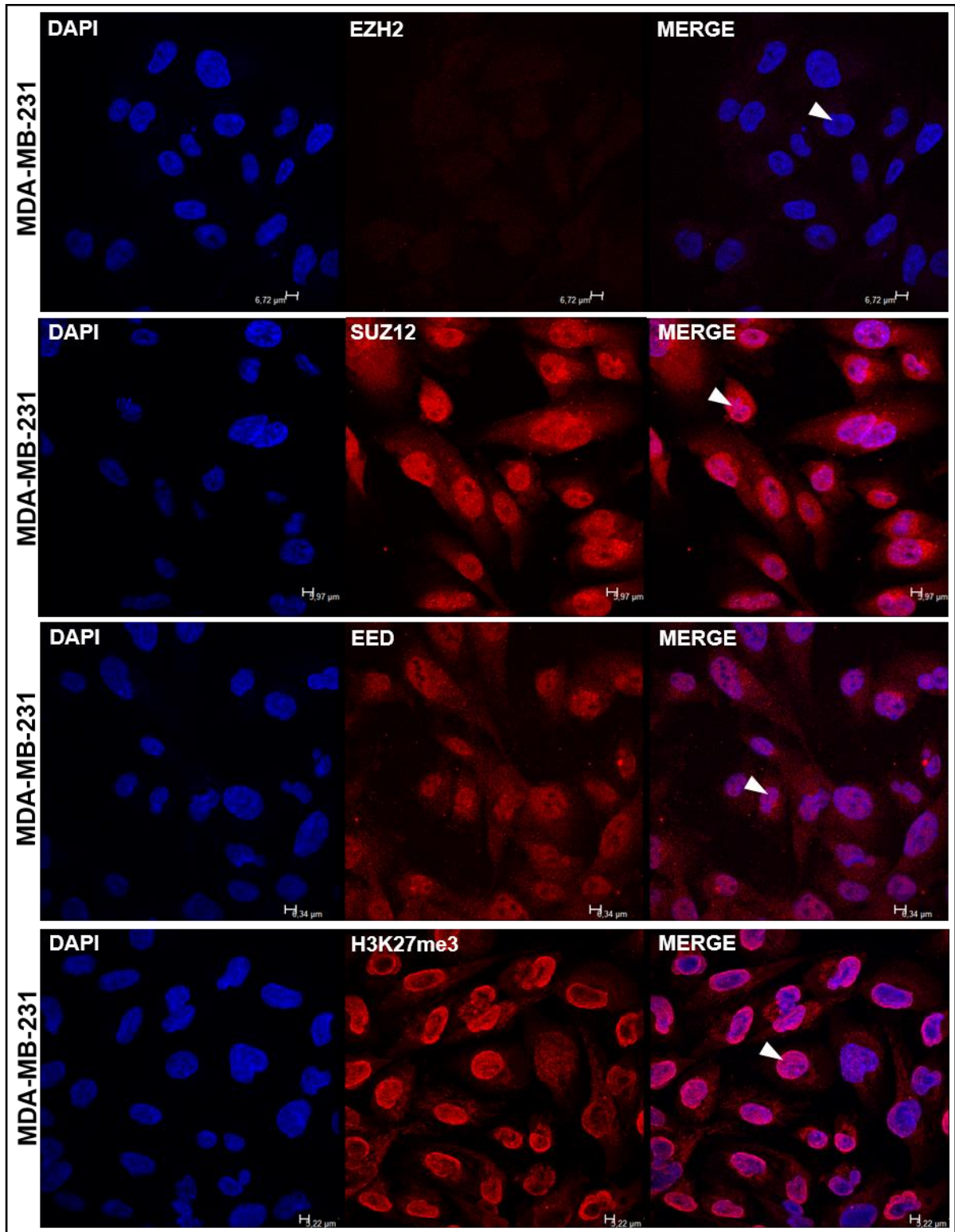
Alguns dos membros do complexo PRC2 foram avaliados em termos de localização celular em ambos os modelos (Figuras 5.4 e 5.5). Embora a marcação do do anticorpo anti-EZH2 tenha ficado fraca no experimento de imunofluorescência, podemos observar localização difusa, tanto citoplasmática quanto nuclear, desta proteína nas duas linhagens. SUZ12, por sua vez, está presente principalmente no núcleo e região perinuclear em ambos os modelos. Em MCF-7, observa-se um acúmulo perinuclear de SUZ12. A proteína EED foi encontrada em maior abundância no citoplasma de MCF-7, enquanto em MDA-MB-231 a mesma se localiza principalmente no núcleo.

Uma vez que os membros do complexo PRC2 apresentaram, em geral, maior expressão gênica e proteica em MDA-MB-231, bem como maior localização no núcleo desta linhagem, quando comparado com MCF-7, avaliamos a localização de sua marca de

silenciamento, H3K27me3, nos dois modelos, a fim de verificar se esta diferença em expressão e localização se reflete na função do complexo. Conforme esperado, a marca de metilação do complexo PRC2, apesar de estar presente no núcleo de ambas as linhagens celulares, se mostrou mais intensa em MDA-MB-231 (média de 71,28% de colocalização com o núcleo) em comparação com MCF-7 (média de 41,21% de colocalização com o núcleo), corroborando com os resultados de expressão e localização dos membros do complexo.

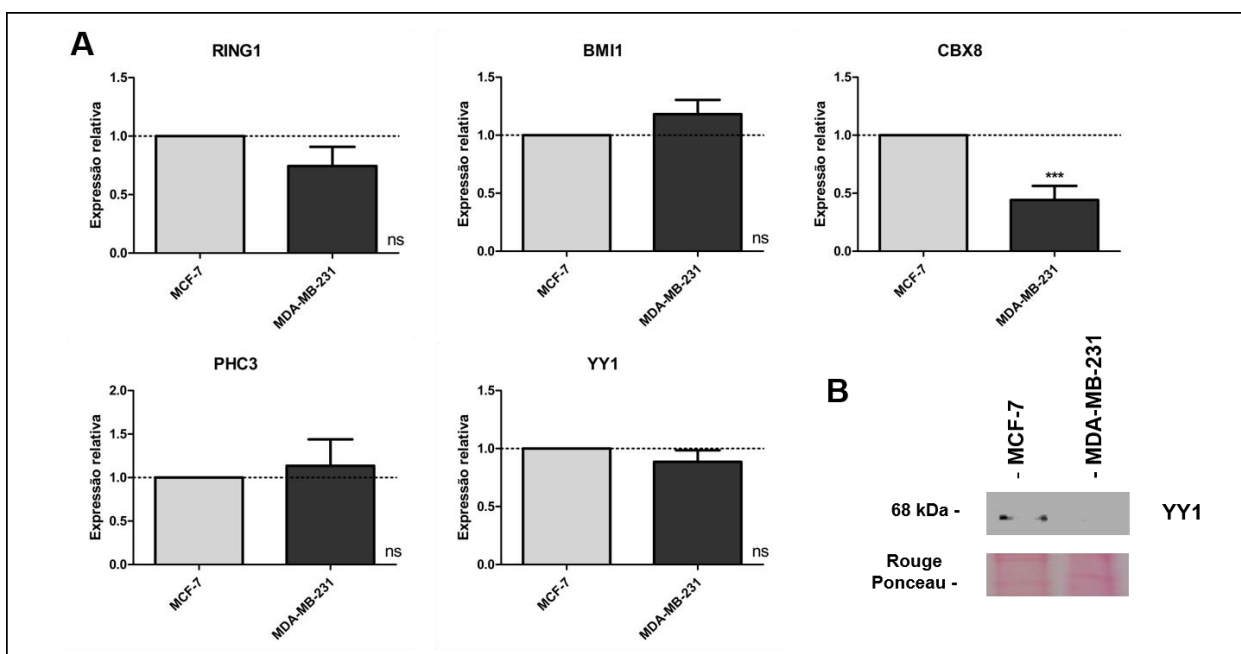


**Figura 5.4 - Localização de membros de PRC2 em MCF-7.** Localização celular de EZH2, SUZ12, EED e H3K27me3 por imunofluorescência em MCF-7. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpos específicos para EZH2 (1:10), SUZ12 (1:20), EED (1:50) e H3K27me3 (1:20). Núcleos marcados em azul (DAPI) e proteínas de interesse em vermelho (ampliação 63x). Cabeça de seta: núcleo celular.

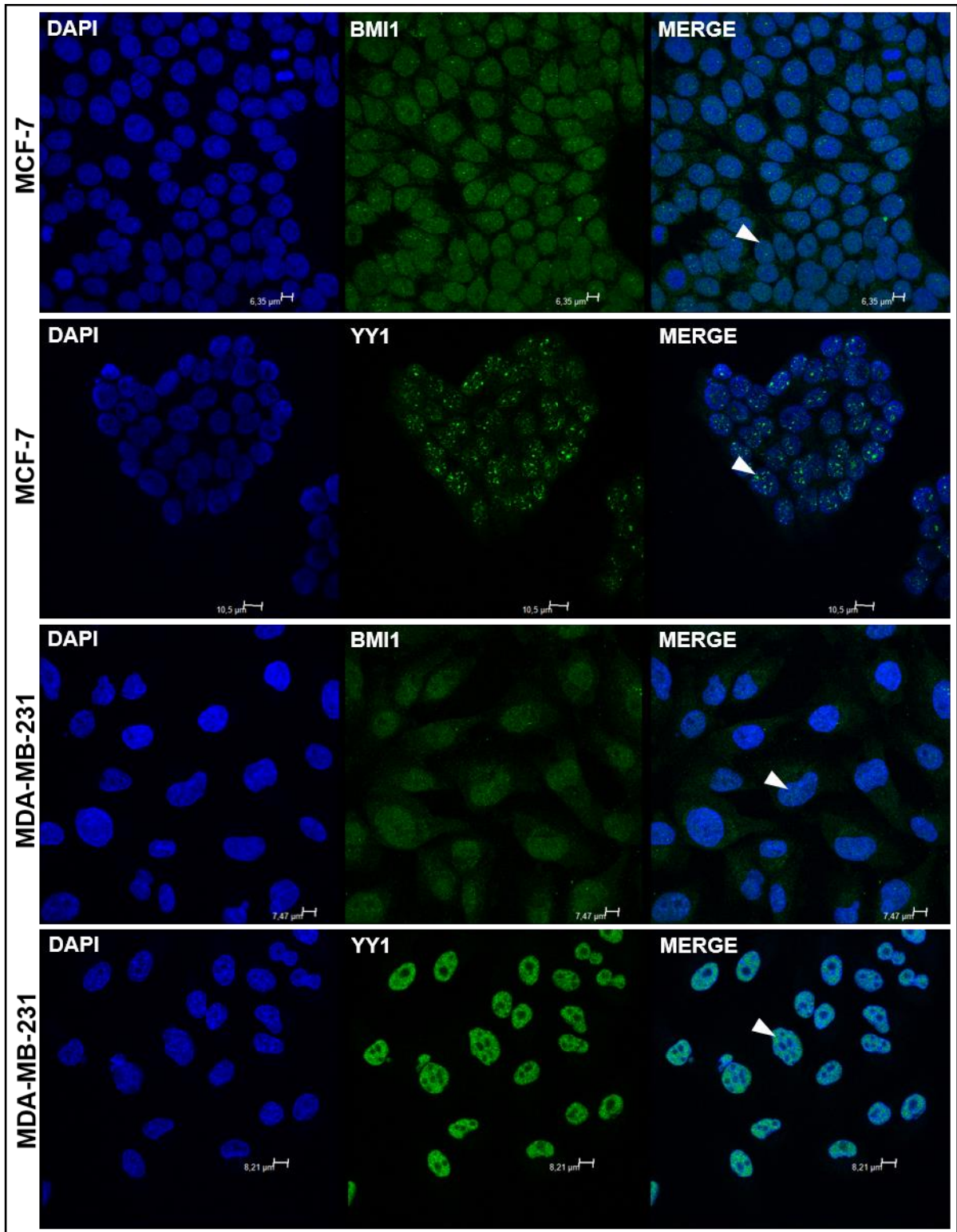


**Figura 5.5 - Localização de membros de PRC2 em MDA-MB-231.** Localização celular de EZH2, SUZ12, EED e H3K27me3 por imunofluorescência em MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpos específicos para EZH2 (1:10), SUZ12 (1:20), EED (1:50) e H3K27me3 (1:20). Núcleos marcados em azul (DAPI) e proteínas de interesse em vermelho (ampliação 63x). Cabeça de seta: núcleo celular.

A expressão gênica de membros do complexo PRC1 (Figura 5.6) tem pouca variação dentre as linhagens MCF-7 e MDA-MB-231, com a exceção de *CBX8*, com papel na associação do complexo com a cromatina, que encontra-se menos expresso na linhagem TN. Em experimento de Western Blot, a proteína YY1, capaz de se ligar ao DNA e de função acessória no complexo, foi encontrada com maior expressão em MCF-7. Em termos de localização (Figura 5.7), BMI1 e YY1 estão presentes no núcleo das linhagens MCF-7 e MDA-MB-231, com diferenças no padrão de localização. Enquanto BMI1 está majoritariamente localizado no núcleo de MCF-7, uma marcação no citoplasma pode ser observada na linhagem MDA-MB-231. Já YY1 apresenta-se com focos de marcação no núcleo de MCF-7, e se encontra mais difuso, ou seja, sem focos concentrados de marcação, no núcleo de MDA-MB-231. Estes dados sugerem que o complexo PRC1 poderia ser formado em ambas as linhagens.



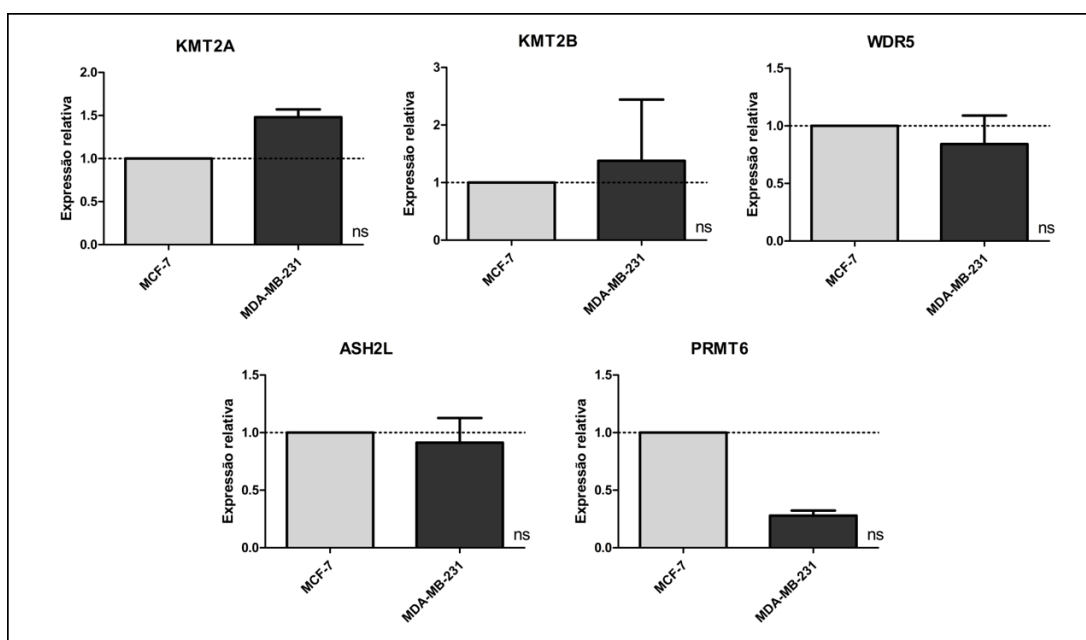
**Figura 5.6 - Expressão de membros de PRC1 em linhagens celulares de CM.** (A) Expressão gênica relativa dos transcritos de *RING1*, *BMI1*, *CBX8*, *PHC3* e *YY1* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). ns: diferença não significativa. \*\*\*p<0,001. Barras de erro indicativas de desvio padrão. (B) Expressão proteica de YY1 em MCF-7 e MDA-MB-231. Foram resolvidos 30 µg de extratos proteicos totais das linhagens MCF-7 e MDA-MB-231 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-YY1 em diluição 1:500. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau.



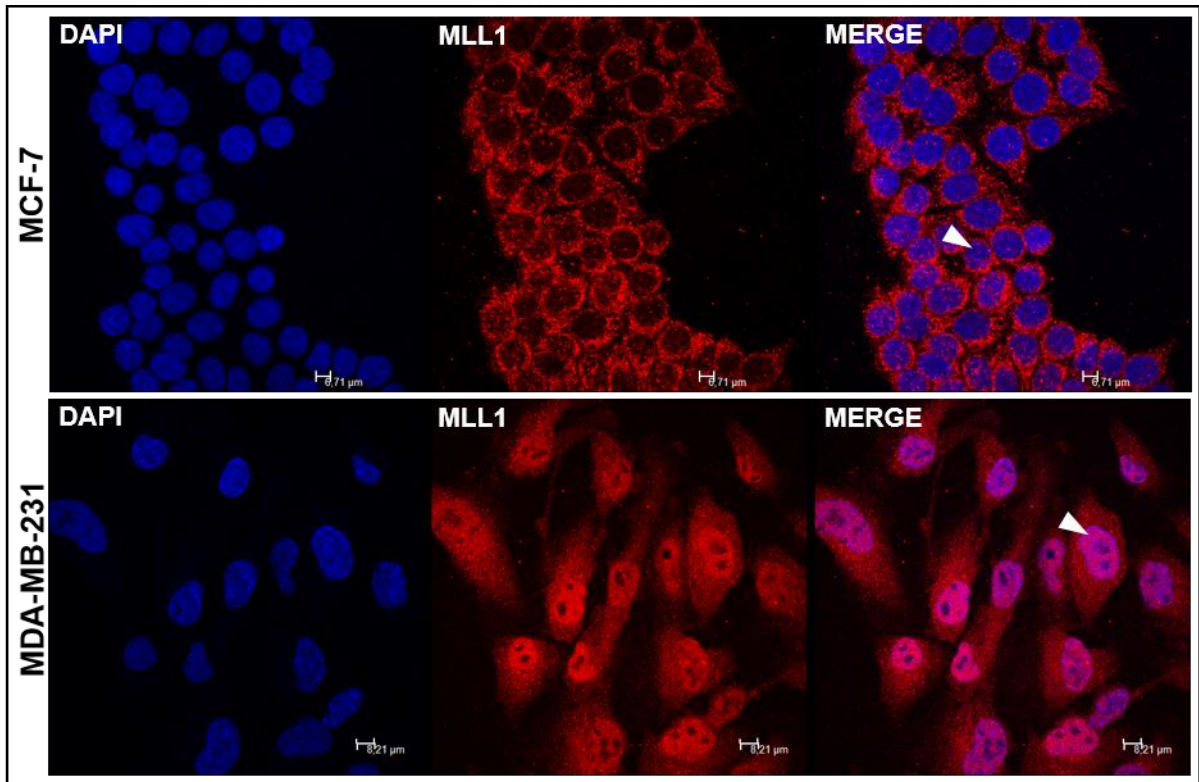
**Figura 5.7 - Localização de membros de PRC1 em MCF-7 e MDA-MB-231.** Localização celular de BMI1 e YY1 por imunofluorescência em MCF-7 e MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpos específicos para BMI1 (1:20) e YY1 (1:50). Núcleos marcados em azul (DAPI) e proteínas de interesse em verde (ampliação 63x). Cabeça de seta: núcleo celular.



Dentre os membros da família Trx, foram avaliados os componentes do complexo MLL (Figura 5.8), que não apresentaram diferenças significativas de expressão gênica entre as linhagens MCF-7 e MDA-MB-231. Avaliamos também a localização de MLL1 (KMT2A) (Figura 5.9), a proteína efetora deste complexo, observada principalmente no citoplasma de células MCF-7 e em abundância no núcleo de células MDA-MB-231, mostrando que o padrão de expressão gênica observado para esta proteína se relaciona com o aumento da localização nuclear da mesma em MDA-MB-231. Este achado sugere que MLL1 pode ser mais ativo na linhagem TN.

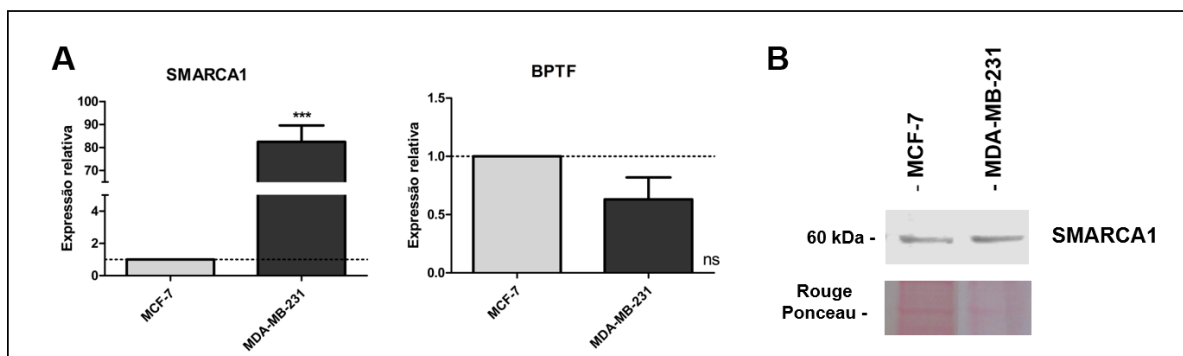


**Figura 5.8 - Expressão de membros de MLL em linhagens celulares de CM.** (A) Expressão gênica relativa dos transcritos de *KMT2A*, *KMT2B*, *WDR5*, *ASH2L* e *PRMT6* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). ns: diferença não significativa. Barras de erro indicativas de desvio padrão.

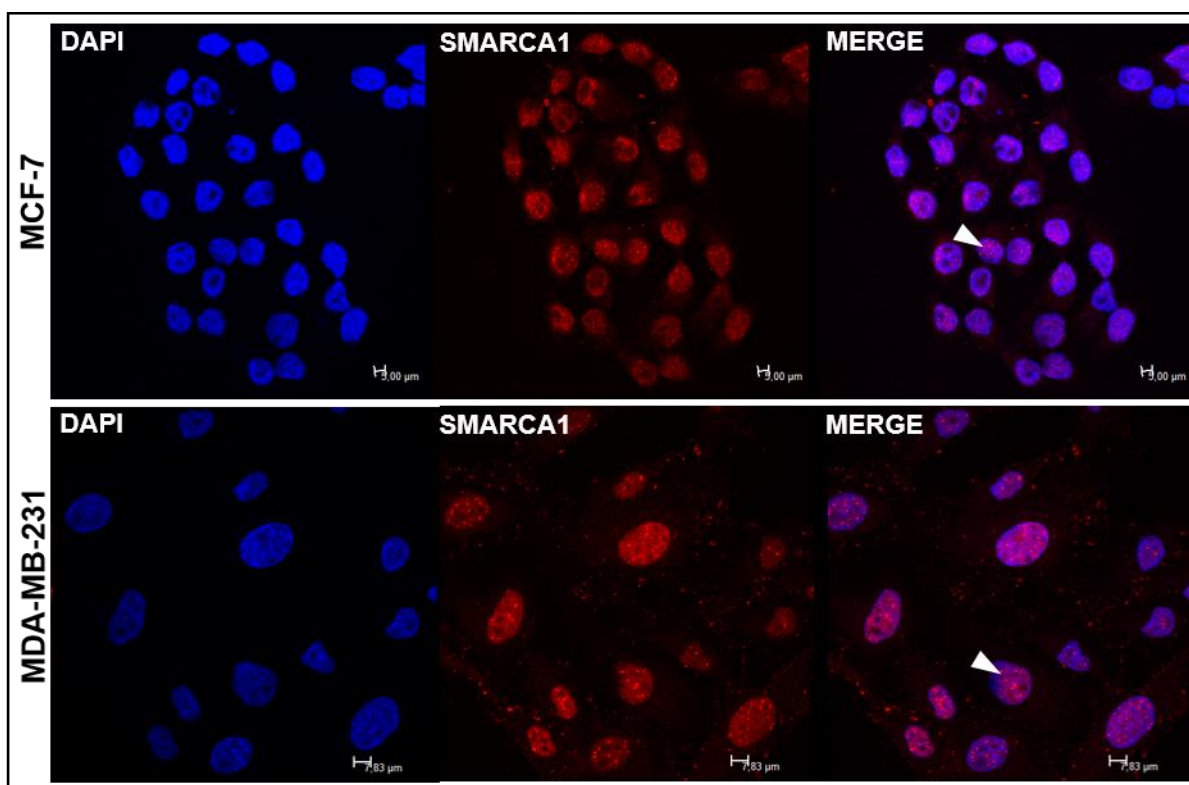


**Figura 5.9 - Localização de MLL1 em MCF-7 e MDA-MB-231.** Localização celular de MLL1 (KMT2A) por imunofluorescência em MCF-7 e MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpo específico para MLL1, em diluição 1:50. Núcleos marcados em azul (DAPI) e MLL1 em vermelho (ampliação 63x). Cabeça de seta: núcleo celular.

Na família NURF (Figura 5.10), foi possível observar um grande aumento na expressão gênica de *SMARCA1*, a ATPase deste complexo de remodelamento de nucleossomos, na linhagem MDA-MB-231 (*fold change* acima de 80). Porém, este aumento expressivo não é observado em níveis de proteína, nos quais *SMARCA1* foi encontrada com expressão semelhante em ambos os modelos. Nas duas linhagens, *SMARCA1* localiza-se principalmente no núcleo das células, onde seria capaz de promover o remodelamento da cromatina (Figura 5.11). Quanto ao membro *BPTF*, capaz de reconhecer a modificação de histona feita pelo complexo MLL, não foram observadas alterações significativas em sua expressão gênica. Pela localização nuclear de *SMARCA1* nas linhagens Luminal A e TN, não podemos afirmar que há diferença na função desta proteína entre os dois subtipos.



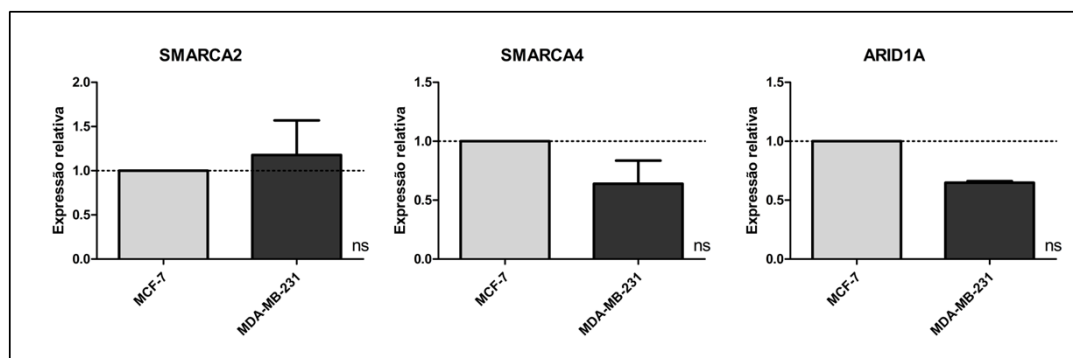
**Figura 5.10 - Expressão de membros de NURF em linhagens celulares de CM.** (A) Expressão gênica relativa dos transcritos de *SMARCA1* e *BPTF* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). ns: diferença não significativa. \*\*\*p<0,001. Barras de erro indicativas de desvio padrão. (B) Expressão proteica de *SMARCA1* em MCF-7 e MDA-MB-231. Foram resolvidos 30 µg de extratos proteicos totais das linhagens MCF-7 e MDA-MB-231 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-YY1 em diluição 1:1000. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau.



**Figura 5.11 - Localização de SMARCA1 em MCF-7 e MDA-MB-231.** Localização celular de *SMARCA1* por imunofluorescência em MCF-7 e MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpo específico para *SMARCA1*, em diluição 1:50. Núcleos marcados em azul (DAPI) e *SMARCA1* em vermelho (ampliação 63x). Cabeça de seta: núcleo celular.

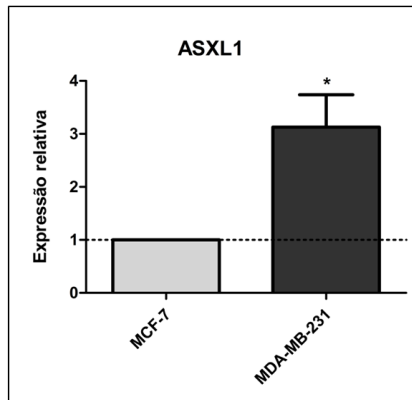
Dentre os membros de SWI/SNF, foi possível notar, na linhagem TN, um leve aumento na expressão de *SMARCA2*, uma das ATPases do complexo, e uma diminuição na expressão de

*SMARCA4*, também de função ATPase, e *ARID1A*, que pode estar envolvido no recrutamento do complexo para a cromatina, em comparação com o modelo Luminal A (Figura 5.12). As diferenças de expressão identificadas, porém, não foram estatisticamente significativas, ficando em aberto uma função deste complexo nas linhagens utilizadas como modelos de Luminal A e TN deste estudo.



**Figura 5.12 – Expressão de membros de SWI/SNF em linhagens celulares de CM.** Expressão gênica relativa dos transcritos de *SMARCA2*, *SMARCA4* e *ARID1A* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). ns: diferença não significativa. Barras de erro indicativas de desvio padrão.

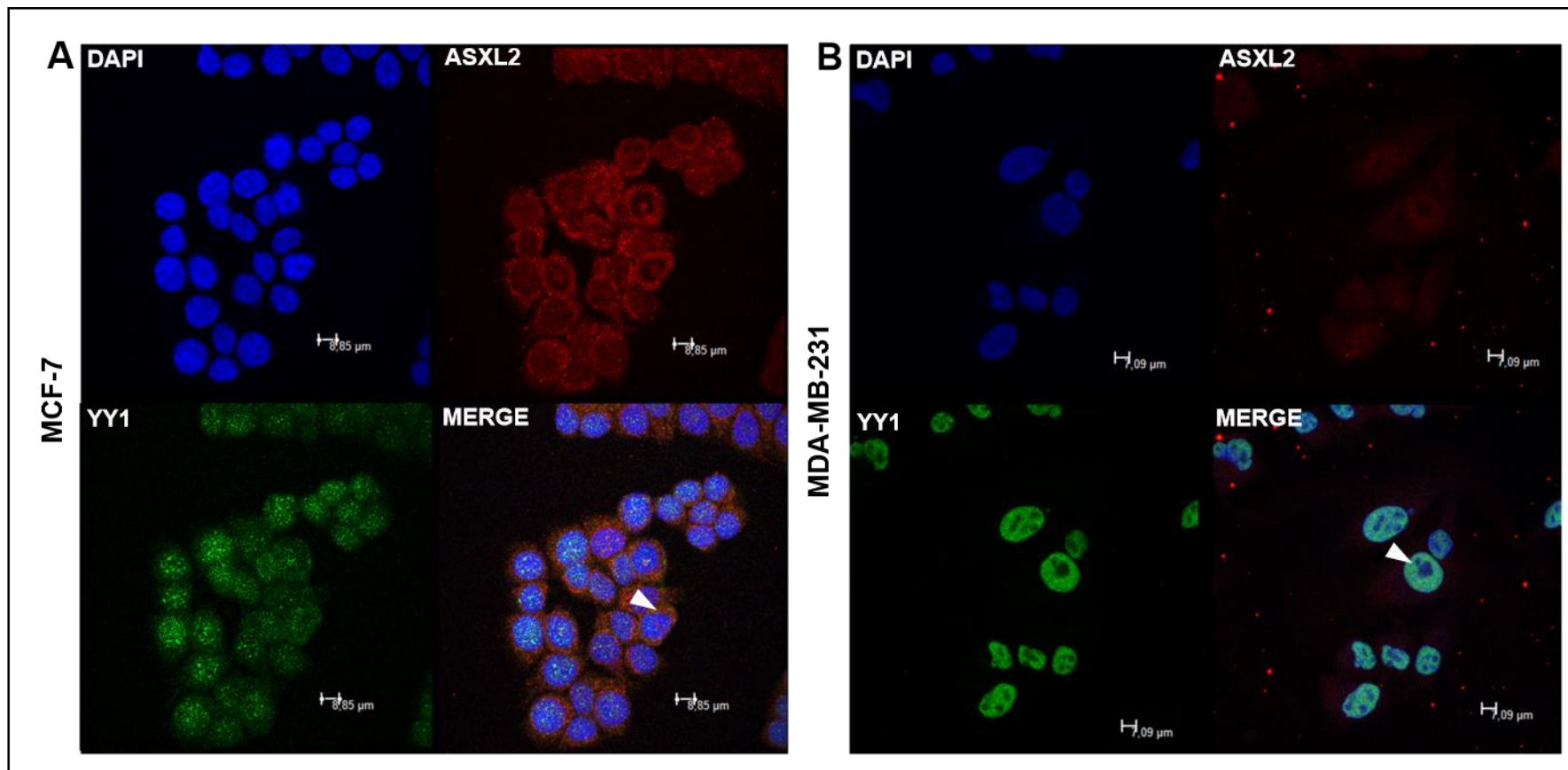
Além de membros das diversas famílias de modificadores de histona e remodeladores de cromatina, também avaliamos a expressão gênica de *ASXL1* nas duas linhagens celulares estudadas, uma vez que este gene também foi implicado em diversos tipos de tumores (Figura 5.13). Assim como *ASXL2*, *ASXL1* também está mais expresso na linhagem MDA-MB-231 em comparação com MCF-7. Esta diferença de expressão entre os dois modelos, porém, é menor para *ASXL1*, onde observou-se um *fold change* de aproximadamente 3, enquanto para *ASXL2* a diferença de expressão entre as linhagens chegou a aproximadamente 8.



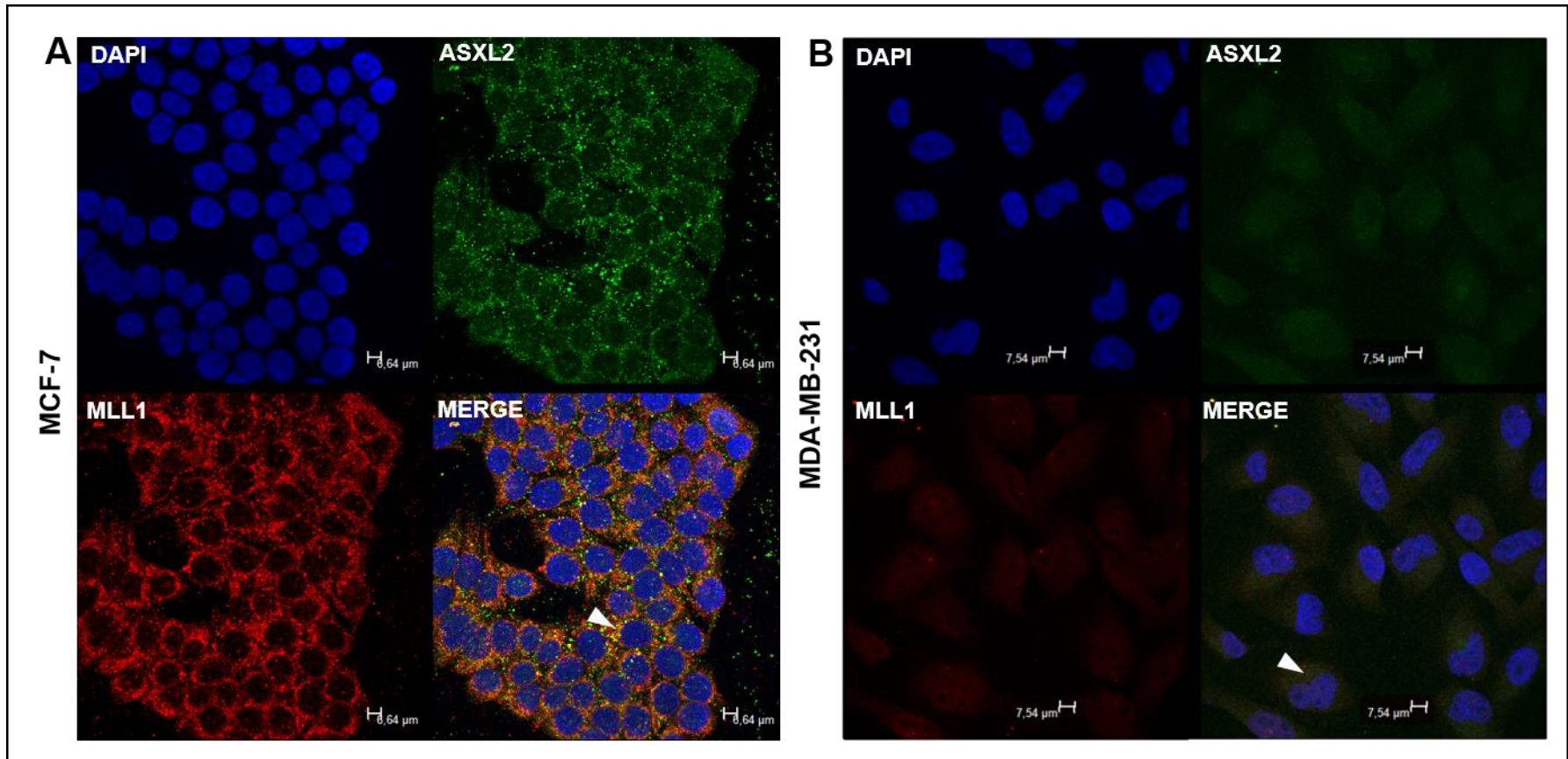
**Figura 5.13 - Expressão de *ASXL1* em linhagens celulares de CM.** Expressão gênica relativa do transcrito de *ASXL1* por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH* em linhagens celulares representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231) (n=3). \* $p < 0,05$ . Barras de erro indicativas de desvio padrão.

Uma vez que membros de complexos modificadores de cromatina estão diferentemente expressos e localizados entre as linhagens estudadas, procuramos avaliar se ASXL2 poderia estar colocalizada com os mesmos nos modelos de interesse. Para tal, realizamos a imunofluorescência de YY1, um membro Pc altamente localizado no núcleo de ambas as linhagens estudadas, e MLL1, um membro Trx cuja localização varia entre os dois modelos, em conjunto com ASXL2 (Figura 5.14). Em MCF-7, o resultado obtido mostra que ASXL2 está colocalizado com YY1 em alguns pontos do núcleo e citoplasma, havendo colocalização média de 8,08% entre ambas as proteínas. Já a colocalização de ASXL2 com MLL1 nesta linhagem é maior, numa média de 22,31%, porém esta é mais pronunciada no citoplasma.

Em MDA-MB-231, a colocalização de ASXL2 com YY1 é baixa, com média de 1,72%. A interação de ASXL2 com MLL neste modelo é de em média 19,84%, resultado próximo do obtido com MCF-7. Estes dados mostram, em conjunto, que ASXL2 não parece interagir com YY1, enquanto que uma possível interação poderia ocorrer para MLL1.



**Figura 5.14 – Colocalização de ASXL2 com YY1 em linhagens celulares de CM.** Localização celular de YY1 por imunofluorescência em (A) MCF-7 e (B) MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpos específicos para ASXL2 (1:20) e YY1 (1:50). Núcleos marcados em azul (DAPI) (ampliação 63x). Cabeça de seta: núcleo celular.



**Figura 5.15 – Colocalização de ASXL2 com MLL1 em linhagens celulares de CM.** Localização celular de MLL1 por imunofluorescência em (A) MCF-7 e (B) MDA-MB-231. Células foram cultivadas até atingir a confluência de 70%, fixadas e incubadas com anticorpos específicos para ASXL2 (1:20) e MLL1 (1:50). Núcleos marcados em azul (DAPI) (ampliação 63x). Cabeça de seta: núcleo celular.

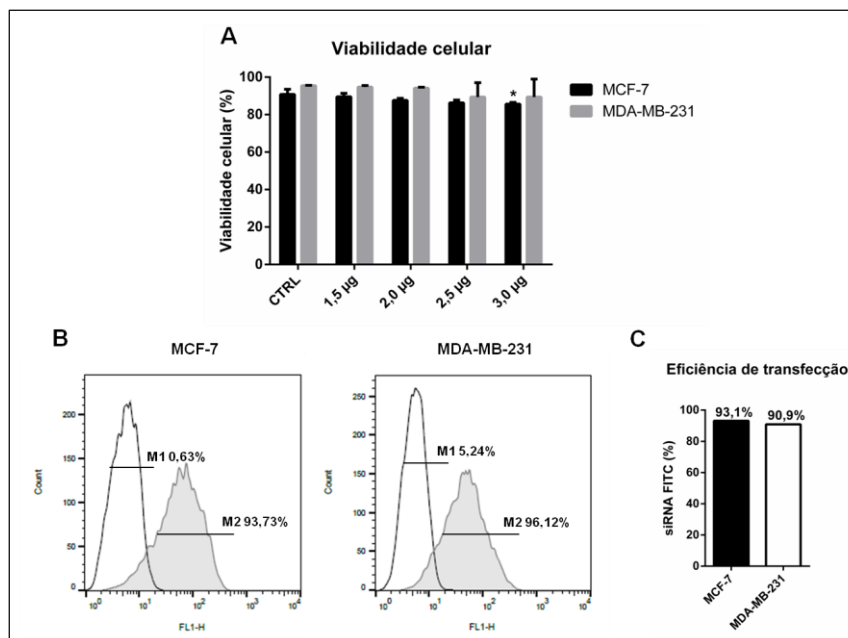
### 5.3 Silenciamento gênico de ASXL2

Com o objetivo de avaliar as vias e sinalizações celulares relacionadas a ASXL2 e diferenças em seu papel regulatório nos subtipos Luminal A e TN, realizamos o silenciamento transiente de ASXL2 nas linhagens MCF-7 e MDA-MB-231, representantes destes subtipos.

Inicialmente, a fim de atestar que os efeitos observados após transfecção com siRNA para ASXL2 seriam decorrentes apenas da ação do RNA de interferência, diferentes concentrações do reagente de transfecção Lipofectamina 3000 foram testadas, e seus efeitos na viabilidade celular avaliados por citometria de fluxo após marcação com PI. Decorridas 24 h do tratamento das células com a curva de dose (1,5; 2,0; 2,5; 3,0 µg), observou-se pouca redução na viabilidade celular nas doses utilizadas (Figuras 5.16A). Na linhagem MCF-7, porém, a dose de 3,0 µg de lipofectamina reduziu a viabilidade celular em 5,2% ( $p < 0,05$ ). Portanto, a fim de reduzir os efeitos de toxicidade nas células, optou-se por utilizar a dose de 2,5 µg de lipofectamina para ambas as linhagens.

Após definição da dose de lipofectamina, realizamos o teste de eficiência de transfecção com uma concentração de 50 nM de siRNA FITC, com o objetivo de assegurar que a concentração de siRNA utilizada seria capaz de exercer sua função biológica nas células. Na linhagem celular MCF-7, a eficiência de transfecção observada nas células tratadas com 50 nM de siRNA FITC foi de 93,1%, enquanto em MDA-MB-231 este valor foi de 90,9%, em comparação com o *mock* (Figuras 5.16B e 5.16C).





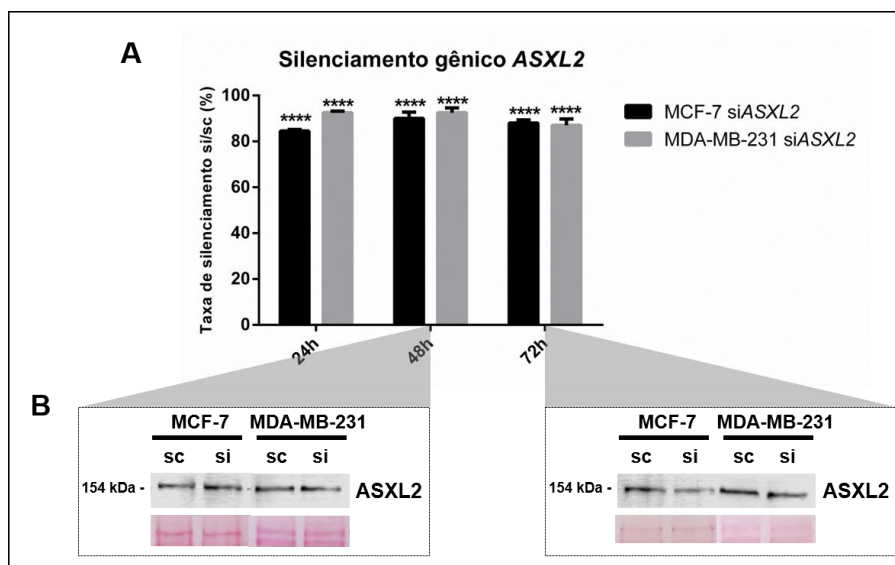
**Figura 5.16 – Padronização do experimento de silenciamento gênico de *ASXL2*.** (A) Padronização da concentração de lipofectamina 3000 para realização do experimento de transfecção nas linhagens MCF-7 e MDA-MB-231. A viabilidade celular das células incubadas por 24h com uma curva de doses de lipofectamina (1,5, 2,0, 2,5 e 3,0 µg) foi avaliada por citometria de fluxo após marcação com iodeto de propídio, e comparada com a viabilidade de células controle (CTRL). \*  $p < 0,05$ . Barras de erro indicativas de desvio padrão. (B) Histogramas representativos da análise por citometria de fluxo da eficiência de transfecção das linhagens MCF-7 e MDA-MB-231. Valores de mediana de intensidade de autofluorescência (M1) e da fluorescência emitida por células marcadas (M2) representados em porcentagem em ambos os histogramas. As condições avaliadas foram *mock* (área abaixo da curva em branco) e siRNA FITC 50 nM (área abaixo da curva em cinza). (C) Gráfico representativo da eficiência de transfecção de ambas as linhagens em comparação com o *mock*, utilizando 2,5 µg de lipofectamina e 50 nM de siRNA FITC. Valores obtidos através da média de intensidade de fluorescência observada em cada linhagem tratada com siRNA FITC, descontando-se os valores de autofluorescência obtidos no *mock*.

Após a padronização das concentrações de reagente de transfecção e siRNA a serem utilizadas, avaliamos a taxa de silenciamento gênico de *ASXL2*, através do tratamento das células com 50 nM de siRNA *ASXL2* nos tempos de 24, 48 e 72 h.

Conforme esperado, não foram observadas alterações significativas na expressão de *ASXL2* entre controle e *scramble*, após normalização com o gene endógeno *GAPDH*. Desta forma, as comparações da expressão gênica foram realizadas entre as células *scramble* e as células tratadas com siRNA *ASXL2*. Na linhagem celular MCF-7, as taxas médias de silenciamento obtidas em 24, 48 e 72 h foram de 84,5%, 90% e 88%, respectivamente. Em MDA-MB-231, por sua vez, as taxas médias de silenciamento foram de 92,5%, 92,5% e 87% nos tempos de 24, 48 e 72 h respectivamente (Figura 5.17A).

Para confirmação em nível protéico, realizamos extração de proteínas totais, após 48h e 72 h de silenciamento, e avaliação da expressão de *ASXL2* por WB. Em ambos os modelos, foi

possível observar a menor expressão de *ASXL2* nas amostras silenciadas após 72h, comparadas com as amostras *scramble* (Figura 5.17B). Assim, prosseguimos com as análises de microarranjo no tempo de 72h de silenciamento.



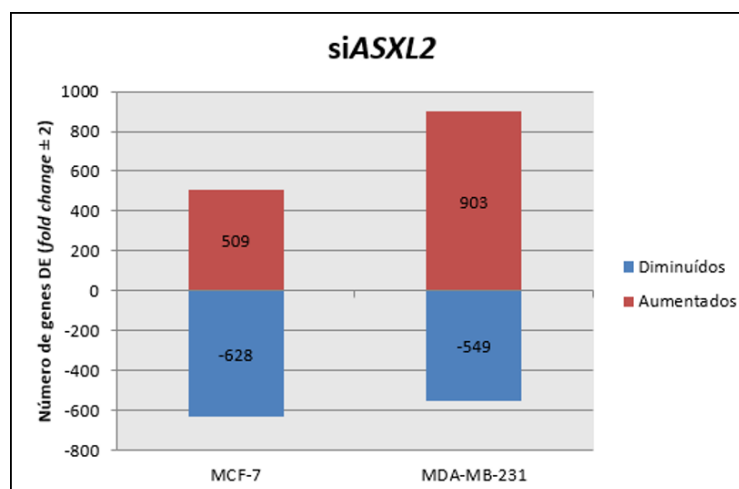
**Figura 5.17 – Silenciamento de *ASXL2* após a transfecção com siRNA.** (A) Taxa de silenciamento gênico. A quantificação relativa por RT-qPCR dos transcritos de *ASXL2* nas linhagens MCF-7 e MDA-MB-231 foi realizada nos tempos de 24, 48 e 72 h após transfecção com siRNA *ASXL2* (n=3). Os níveis transcricionais foram normalizados pelo gene endógeno *GAPDH*, e a taxa de silenciamento foi avaliada em comparação com os respectivos *scramble*. \*\*\*\* p<0,0001. Barras de erro indicativas de desvio padrão. (B) Silenciamento de *ASXL2* em nível proteico após 48h e 72h da transfecção com siRNA. Foram resolvidos 30 µg de extratos proteicos das linhagens MCF-7 e MDA-MB-231 *scramble* e siASXL2 em gel SDS-PAGE 10%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-*ASXL2* em diluição 1:1000.

#### 5.4 *ASXL2* tem impacto em vias de pluripotência e nas sinalizações de Wnt e TGF-beta em modelos *in vitro* de CM

As amostras de RNA total utilizadas no experimento de microarranjo foram inicialmente avaliadas em relação à sua integridade, e todas obtiveram valor de RIN igual a 10, equivalente à pontuação máxima de integridade de RNA, garantindo sua qualidade.

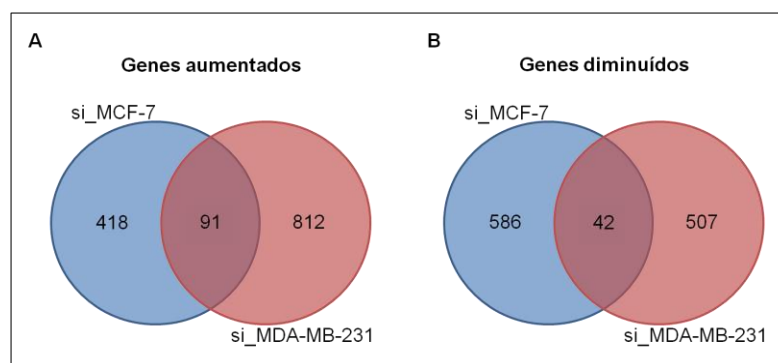
Após avaliação da integridade de RNA, prosseguimos com o experimento de microarranjo, a fim de identificar os genes diferencialmente expressos (DE) na condição siASXL2, em comparação com *scramble*, em ambas as linhagens celulares estudadas, e em seguida, relacionar tais diferenças na expressão com vias de sinalização e processos celulares possivelmente envolvidos com o papel de *ASXL2* no CM.

Como ponto de corte, selecionamos os genes cuja expressão diferencial (siASXL2 *versus* scramble) fosse de no mínimo 2 vezes (*fold change* =  $\pm 2$ ), a fim de realizar as análises subsequentes com maior representatividade de vias e processos celulares, e com o maior número de genes possível. Assim, em MCF-7, um total de 1137 genes foram encontrados DE após o silenciamento de ASXL2, sendo 509 com expressão aumentada e 628 com expressão diminuída, com um intervalo de *fold change* variando de -124 a 14. Na linhagem MDA-MB-231, 1452 genes apresentaram expressão diferencial após o silenciamento, sendo 903 com expressão aumentada e 549 com expressão diminuída, com um intervalo de *fold change* variando de -281 a 20. (Figura 5.18). As listas com os genes diferencialmente expressos em cada linhagem encontram-se nos Anexos A e B.



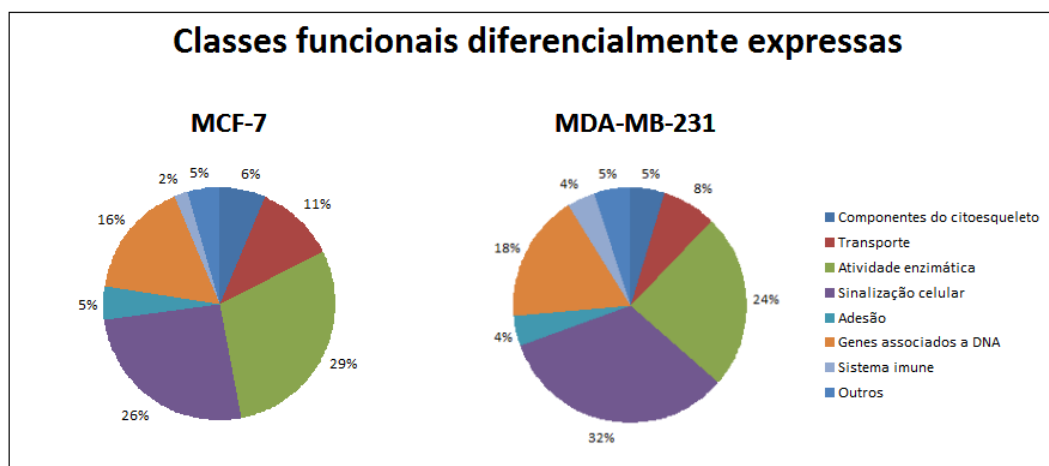
**Figura 5.18 – Perfil de expressão gênica global das linhagens MCF-7 e MDA-MB-231 silenciadas para ASXL2.** Foram identificados um total de 1137 genes diferencialmente expressos na linhagem MCF-7, sendo 509 aumentados e 628 diminuídos, e 1452 genes diferencialmente expressos na linhagem MDA-MB-231, sendo 903 aumentados e 549 diminuídos. *Fold change* (FC) = 2.

Após identificação dos genes DE, realizamos uma análise *in silico* de construção de Diagramas de Venn dos genes aumentados e diminuídos, a fim de verificar se os alvos de ASXL2 são comuns entre os modelos. Por meio desta análise, verificamos que um total de 133 genes DE são comuns entre as linhagens, sendo 91 destes aumentados e 42 diminuídos (Figura 5.19). Este resultado sugere que os genes alvo de ASXL2 são diferentes entre os modelos, uma vez que o número de genes DE em comum representam aproximadamente 10% da totalidade de genes DE identificados em cada uma das linhagens celulares.



**Figura 5.19 – Análise comparativa *in silico* por Diagrama de Venn das linhagens MCF-7 e MDA-MB-231 silenciadas para *ASXL2*.** (A) Genes com expressão aumentada em ambas as linhagens após comparação com seus respectivos *scramble*. (B) Genes com expressão diminuída em ambas as linhagens após comparação com seus respectivos *scramble*. Análise realizada com o *software online* disponível em [bioinformatics.psb.ugent.be/webtools/Venn/](http://bioinformatics.psb.ugent.be/webtools/Venn/). FC=2.

A fim de realizar uma categorização baseada na ontologia gênica, os genes DE foram submetidos à análise *in silico* no *software online* PANTHER. A partir desta análise, identificamos, em ambas as linhagens silenciadas, o enriquecimento de processos celulares tais como transporte, metabolismo e sinalização, associação a DNA, entre outros (Figura 5.20).



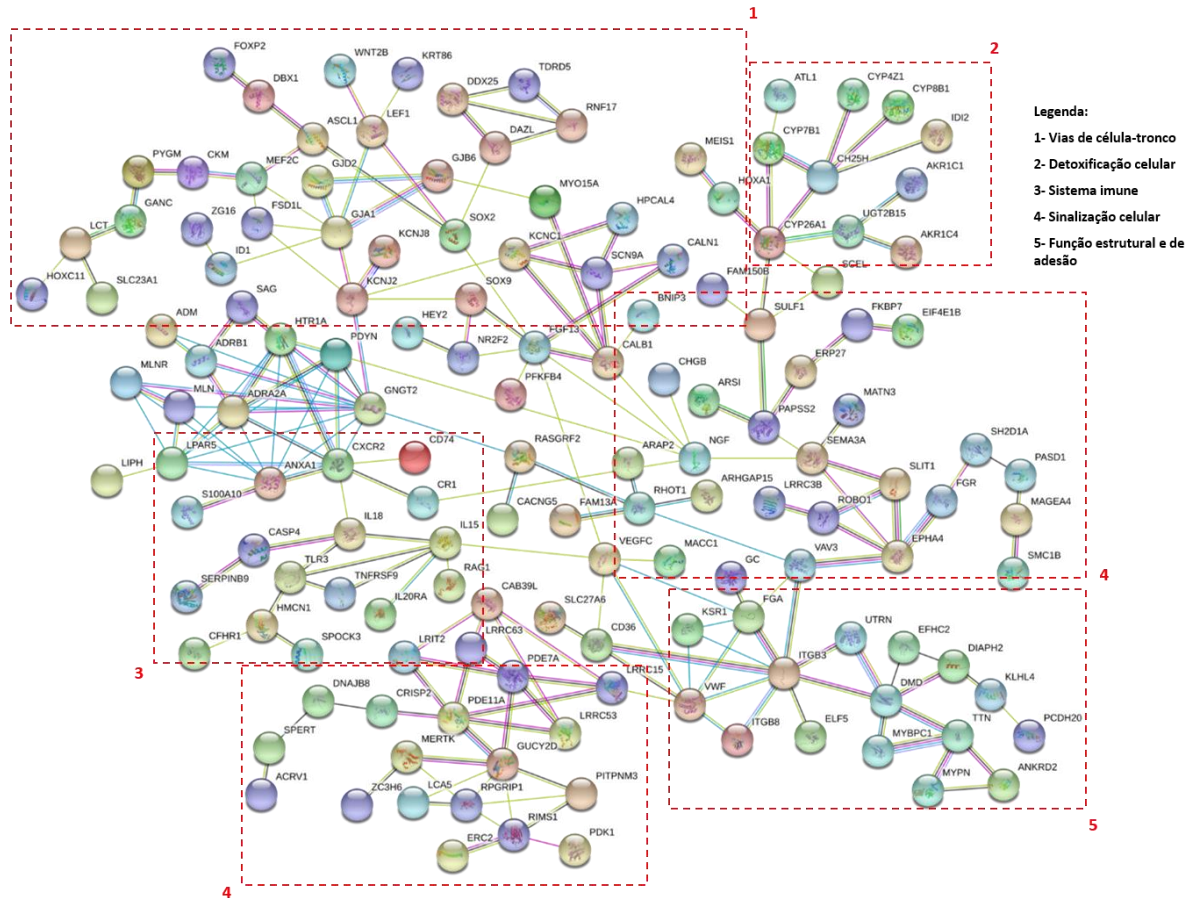
**Figura 5.20 - Classificação funcional de genes diferencialmente expressos após o silenciamento de *ASXL2* em MCF-7 e MDA-MB-231.** Gráficos de pizza indicativos da ontologia gênica após o silenciamento de *ASXL2* nas linhagens MCF-7 e MDA-MB-231. Valores expressos em porcentagem, em relação ao número total de genes DE avaliados pelo *software*.

Dentre os genes DE, figuram em sua maioria enzimas e componentes de vias de sinalização. Além disso, cerca de 16% dos genes DE em MCF-7 e 18% dos genes DE identificados em MDA-MB-231 estão de alguma forma relacionados com DNA, seja por meio de interação direta com o mesmo ou por ação como fator de transcrição. Outras classes funcionais

também foram identificadas, nas quais estão presentes genes associados com adesão celular, resposta imune, transporte de moléculas e tráfico de vesículas, dentre outros.

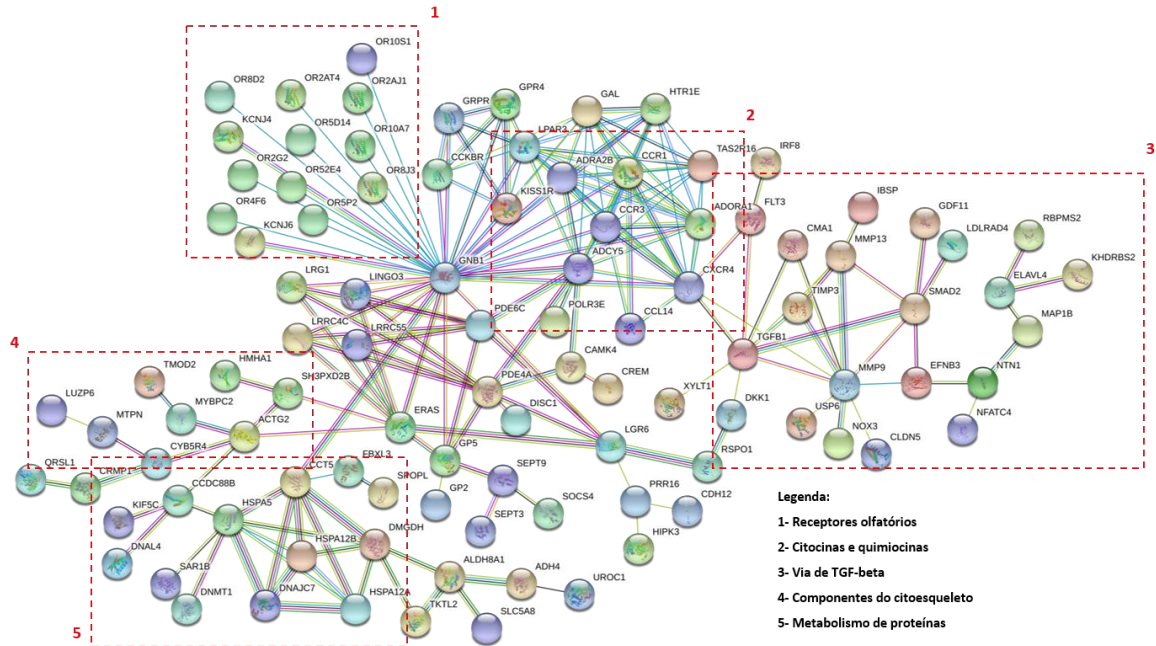
A fim de avaliar como os membros destas classes podem interagir entre si, em nível proteico, submetemos a lista de genes DE a uma análise no *software online* String. A partir desta análise, foram gerados mapas de interação entre os genes/proteínas aumentados e diminuídos em cada linhagem celular silenciada.

O mapa de interação relacionado com os genes diminuídos em MCF-7 após o silenciamento (Figura 5.21) mostra que componentes e reguladores de vias de células-tronco, como as vias de Wnt e HOX, podem interagir, direta ou indiretamente, com componentes envolvidos em detoxificação celular, resposta imune, adesão e regulação do citoesqueleto e sinalização celular. Dentre os genes associados às vias de células-tronco, destacamos fatores de transcrição como *LEF1* (membro da via canônica de Wnt), *SOX9*, *HOXA1* e *HOXC11*, com importantes funções no desenvolvimento embrionário e no câncer. Neste mapa, por exemplo, é possível observar que *HOXA1* interage com genes da família CYP, envolvidos no processo de detoxificação celular. Com relação aos genes associados à resposta imune, estão representados interleucinas e receptores de citocinas, capazes de interagir com genes envolvidos em transdução de sinal, que por sua vez podem interagir com genes relacionados à adesão e regulação do citoesqueleto, como as integrinas *ITGB3* e *ITGB8*, que promovem adesão à matriz extracelular.



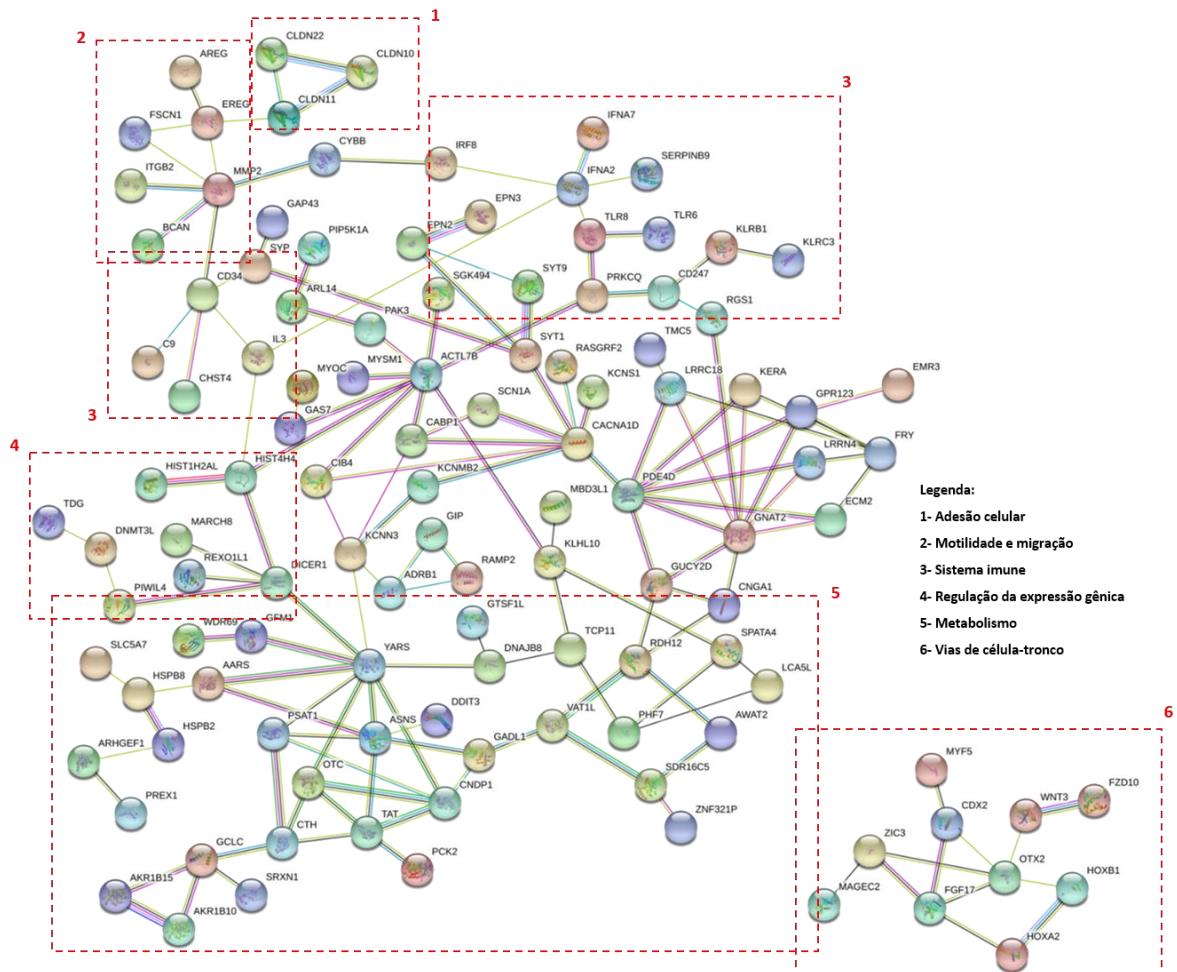
**Figura 5.21** – Mapa de interação dos genes encontrados diminuídos em MCF-7 após o silenciamento de *ASXL2*. Mapa de interação gerado com o auxílio do software online String versão 10.5 (string.db.org), a partir da lista de genes com expressão diminuída na linhagem analisada (*Fold change* = 2). Classes funcionais e vias de interesse foram sinalizadas por quadros vermelhos.

No mapa de interação dos genes aumentados em MCF-7 (Figura 5.22), podemos observar a presença de algumas proteínas, presentes na área central do mapa, capazes de integrar componentes de diferentes vias e processos. Dentre eles, destacamos citocinas e quimiocinas envolvidas com resposta imune, capazes de interagir com componentes da via de TGF beta, como *TGFB1*, um dos ligantes da via, e *SMAD2*, com função de modulação transcricional. Além disso, também observamos interações entre componentes envolvidos na modulação do citoesqueleto de actina e genes relacionados com metabolismo de proteínas, como chaperonas e membros das famílias *heat shock*.



**Figura 5.22 - Mapa de interação dos genes encontrados aumentados em MCF-7 após o silenciamento de *ASXL2*.** Mapa de interação gerado com o auxílio do software online String versão 10.5 (string.db.org), a partir da lista de genes com expressão aumentada na linhagem analisada (*Fold change* = 2). Classes funcionais e vias de interesse foram sinalizadas por quadros vermelhos.

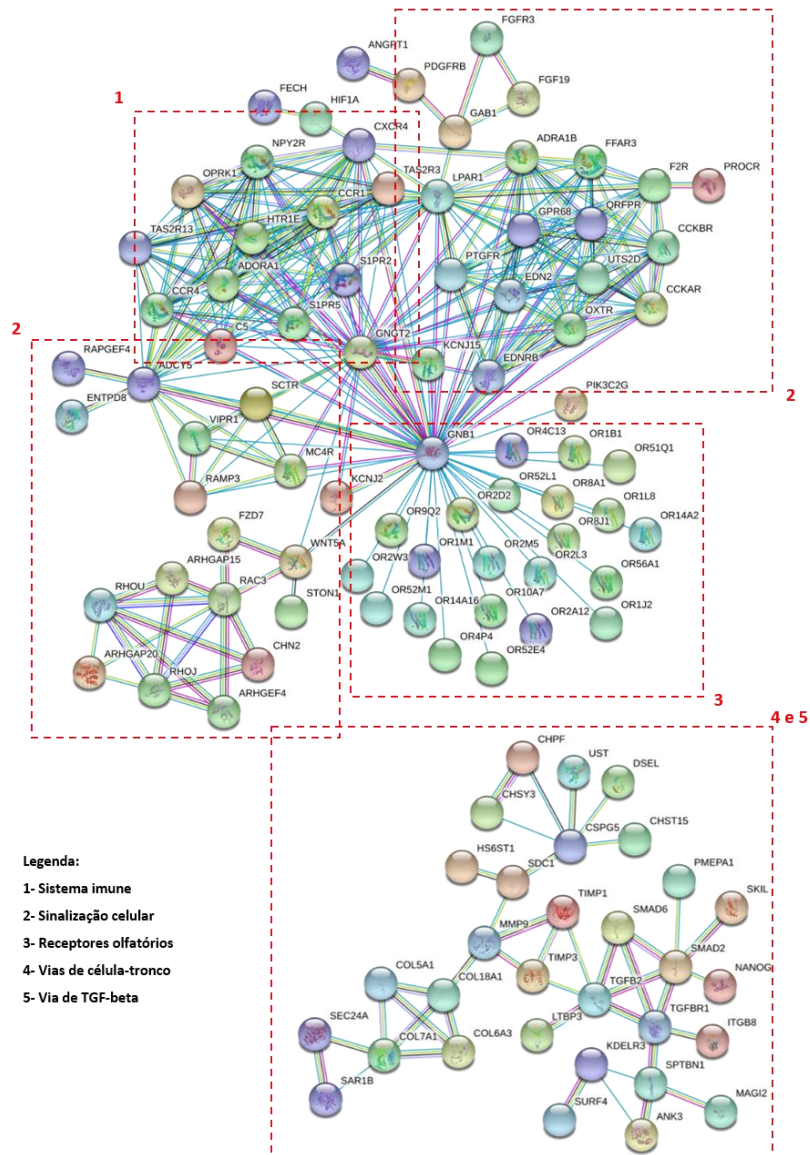
Na linhagem MDA-MB-231 silenciada (Figura 5.23), por sua vez, foram encontrados diminuídos genes associados à adesão celular, que interagem com genes de motilidade e migração, como *MMP2*. Também observamos a interação entre componentes do sistema imune, como interleucinas e receptores de citocinas. Componentes relacionados com regulação da expressão gênica, como *DNMT3L* e *TDG*, envolvidos respectivamente em metilação e demetilação de DNA, e *DICER1*, importante para a biogênese de miRNAs, são capazes de interagir com membros de vias metabólicas. Além disso, também foram identificadas interações entre componentes de vias de células-tronco, como *WNT3*, da via canônica de Wnt, capaz de interagir com genes da via de HOX, tais como *OTX2*, *HOXB1* e *HOXA2*.



**Figura 5.23 - Mapa de interação dos genes encontrados diminuídos em MDA-MB-231 após o silenciamento de *ASXL2*.** Mapa de interação gerado com o auxílio do software online String versão 10.5 ([string.db.org](http://string.db.org)), a partir da lista de genes com expressão diminuída na linhagem analisada (*Fold change* = 2). Classes funcionais e vias de interesse foram sinalizadas por quadros vermelhos.

No mapa de interação dos genes aumentados nesta linhagem (Figura 5.24), podemos observar uma maior densidade de componentes, quando comparado com o mapa de genes diminuídos, refletindo o que foi observado no perfil global de expressão gênica deste modelo silenciado. Dentre os genes aumentados nesta linhagem, também foram encontrados componentes da resposta imune, capazes de interagir com genes envolvidos em sinalização celular. Componentes destes dois grupos também são capazes de interagir com membros de vias de células-tronco, dentre eles *WNT5A*. Membros da via de TGF beta foram encontrados como capazes de interagir entre si, tais como *TGFB2* e *TGFBR1*, respectivamente ligante e receptor da via de TGF beta, e *SMAD6* e *SMAD2*, fatores de transcrição da mesma via, e também como capazes de interagir com *NANOG*, fator de transcrição envolvido em proliferação, renovação e pluripotência, já relacionado com células-tronco tumorigênicas.





**Figura 5.24 - Mapa de interação dos genes encontrados aumentados em MDA-MB-231 após o silenciamento de *ASXL2*.** Mapa de interação gerado com o auxílio do software online String versão 10.5 (string.db.org), a partir da lista de genes com expressão aumentada na linhagem analisada (*Fold change* = 2). Classes funcionais e vias de interesse foram sinalizadas por quadros vermelhos.

Após análise destes mapas com dados da literatura, foi possível associar algumas das classes funcionais identificadas no PANTHER com diversos grupos de genes que interagem entre si para promover tais funções na célula. Entretanto, fez-se necessária uma análise mais focada nas vias de sinalização e fatores de transcrição relacionados com os genes DE, uma vez que a análise pelo *software* String foi capaz de evidenciar interações de genes relacionados com vias relevantes

no câncer. Para isto, realizamos uma análise *in silico* das listas de identificação de genes DE em ambas as linhagens no *software online* Metacore®.

A partir da análise de potenciais fatores de transcrição capazes de regular os genes DE em MCF-7 e MDA-MB-231, foi possível observar fatores relacionados com pluripotência e manutenção de células-tronco, tais como SOX2, Oct3/4 e NANOG, dentre os 10 mais relevantes em ambos modelos. Dentre eles, dois fatores também apresentaram expressão gênica alterada após o silenciamento, SOX2 (*fold change* de -41 na linhagem MCF-7) e NANOG (*fold change* de 2 na linhagem MDA-MB-231). Também existe uma diferença de representatividade dos mesmos (*score*) e para ambos os modelos ainda foram identificados fatores "exclusivos", tais como HNF4-alpha e ESR1 (nuclear) para MCF-7 e TCF8 e GATA-2 para MDA-MB-231 (Tabela 5.1).

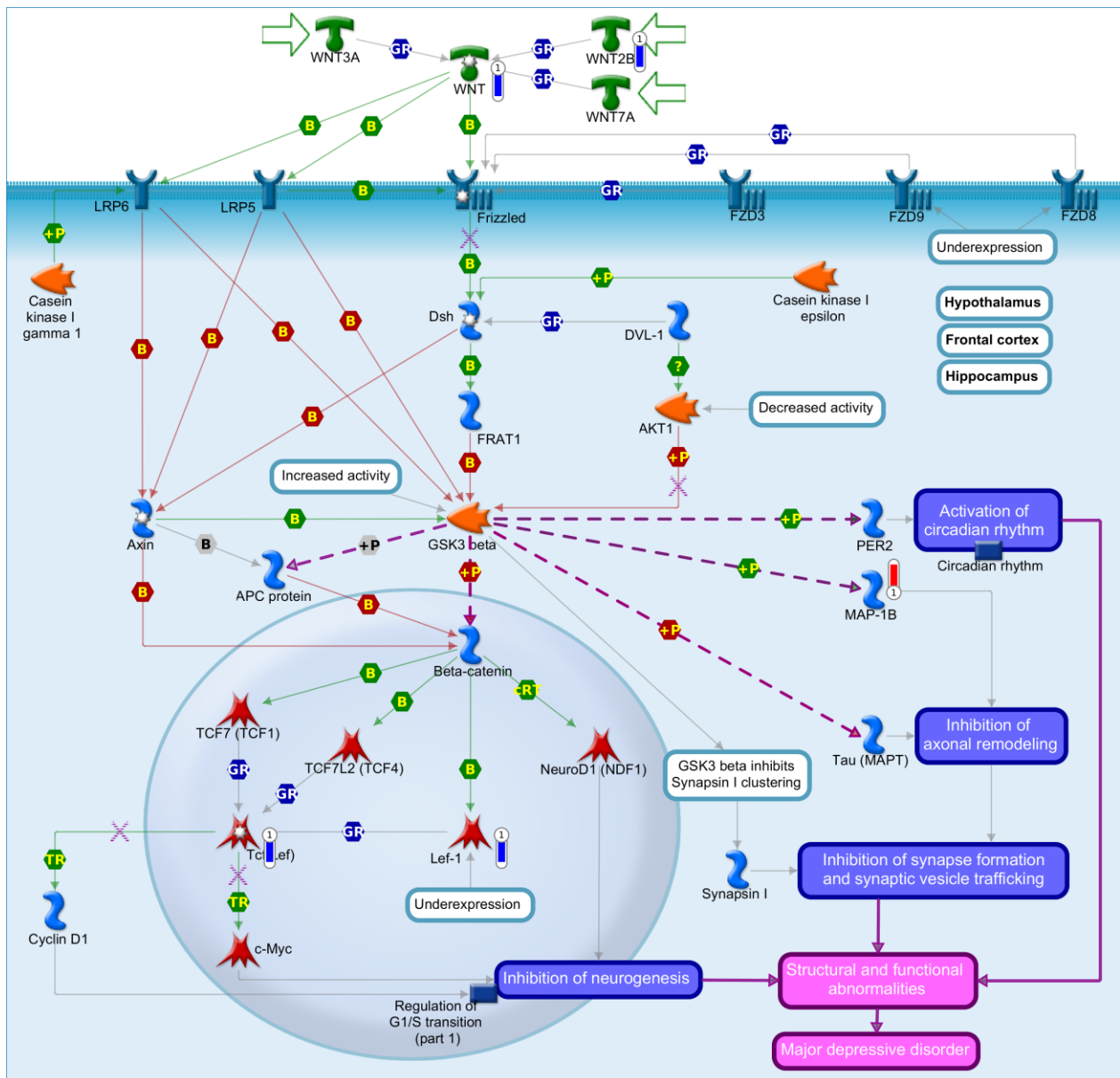
**Tabela 5.1 - Fatores de transcrição relacionados com os genes DE nos modelos celulares MCF-7 e MDA-MB-231 após silenciamento de ASXL2.**

Fatores de transcrição		
	MCF-7	MDA-MB-231
1	FOXP3	FOXP3
2	CREB1	Oct-3/4
3	SOX2	CREB1
4	Oct-3/4	ETS1
5	ETS1	SOX2
6	NANOG	NANOG
7	TCF7L1 (TCF3)	TCF7L1 (TCF3)
8	E2F1	E2F1
9	HNF4-alpha	TCF8
10	ESR1 (nuclear)	GATA-2

Os resultados de enriquecimento de vias de sinalização mostraram, em MCF-7, alta representatividade para Wnt e TGF-beta, além de vias relacionadas com remodelamento de citoesqueleto, adesão e sistema imune. Em MDA-MB-231, além de Wnt e TGF-beta, também foram identificadas vias de MAPK e vias de célula-tronco, tais como Notch e Hedgehog. Ainda, foram reportadas vias relacionadas a sistema imune, ciclo celular, apoptose, citoesqueleto, transição epitélio-mesenquimal e migração.

Com relação à sinalização de Wnt, observamos, em MCF-7 (Figura 5.25), uma diminuição na expressão de genes da via canônica, como por exemplo do ligante *WNT2B* e do

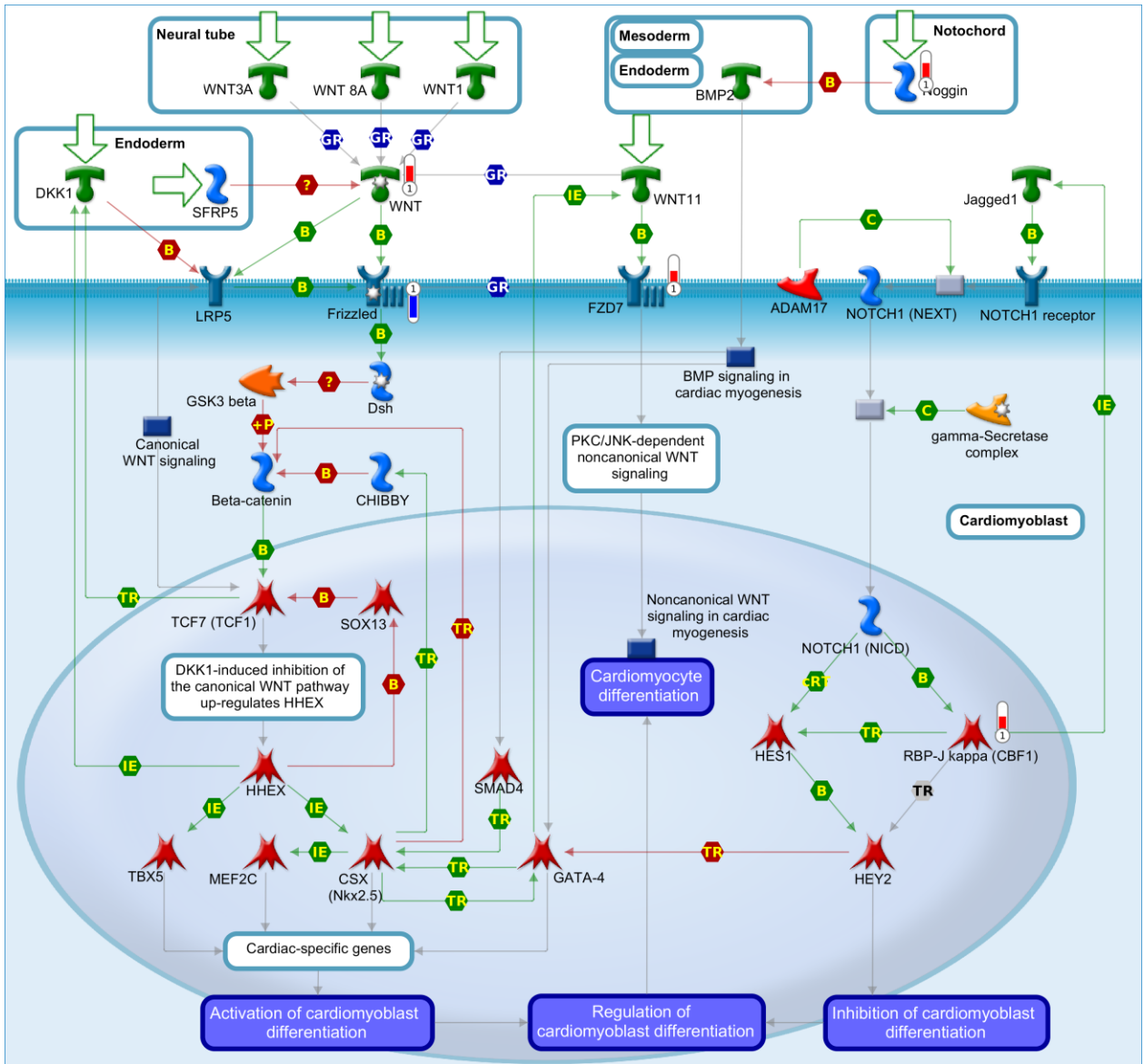
fator de transcrição *LEF1*. Além disso, a expressão aumentada de *DKK1*, um inibidor da via, e *MAP-1B* corrobora com a inibição na via de Wnt/ $\beta$ -catenina.



**Figura 5.25 – Via canônica de Wnt em MCF-7.** Mapa da via canônica de Wnt/ $\beta$ -catenina contendo os genes diferencialmente expressos em MCF-7 após o silenciamento gênico de *ASXL2* (*Fold change* = 2). Os termômetros indicam, em vermelho, a expressão aumentada e, em azul, a expressão diminuída. Legenda dos símbolos e interações disponível no Anexo C.

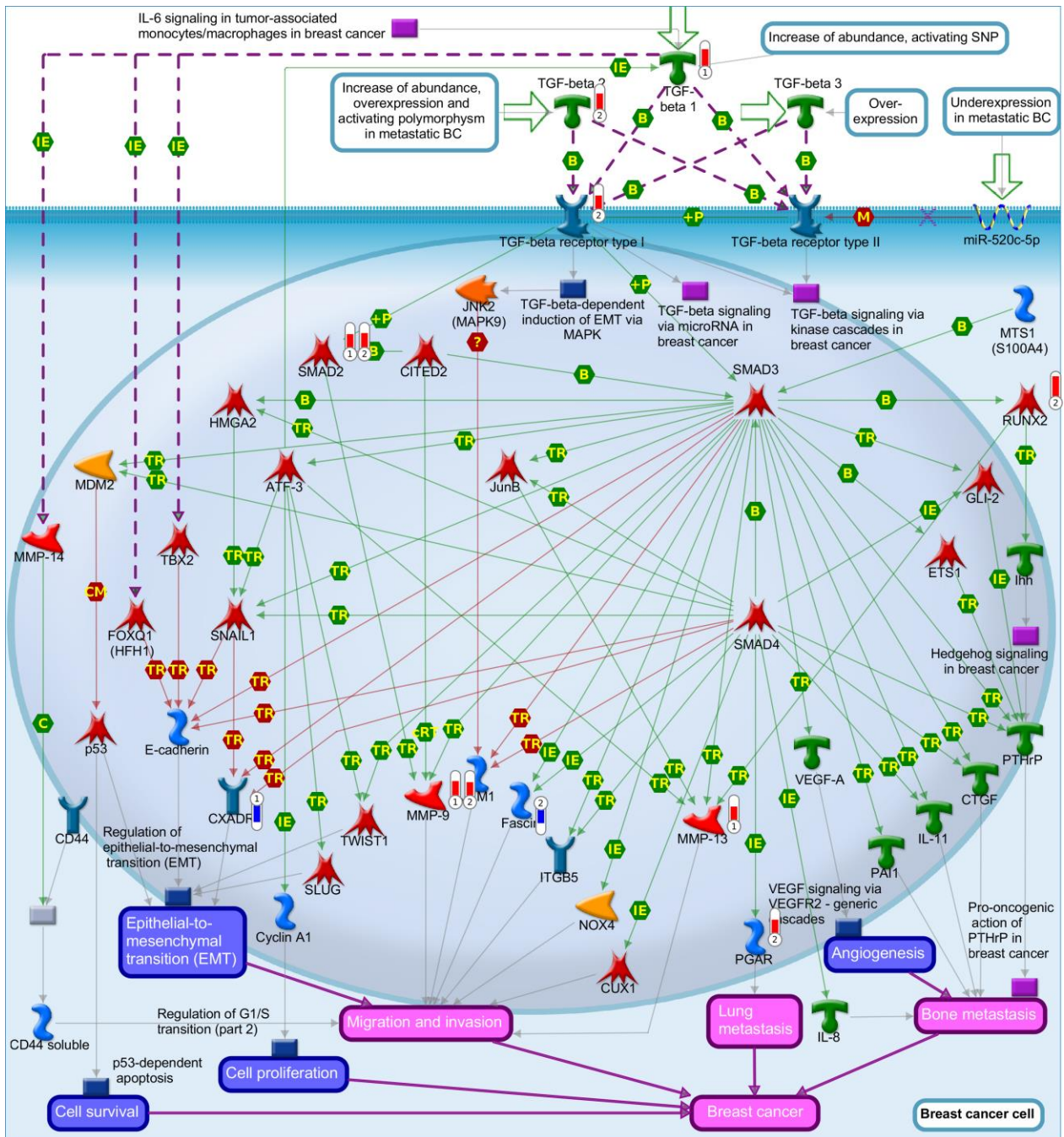
Em MDA-MB-231 (Figura 5.26), observamos tanto uma diminuição (*WNT3*, via canônica) quanto aumento (*WNT5A*, vias canônica e não-canônica) de expressão de ligantes. Ainda, o receptor *FZD10* (via canônica) está diminuído após silenciamento de *ASXL2*, ao passo

que *FZD7* (via não-canônica) está aumentado. O inibidor da via canônica *SOST* também está aumentado neste modelo, indicando uma possível ativação da via não-canônica sobre a canônica.



**Figura 5.26 – Via de sinalização de Wnt em MDA-MB-231.** Mapa da sinalização de Wnt com as vias Wnt/ $\beta$ -catenina (canônica) e não-canônica, contendo os genes diferencialmente expressos em MDA-MB-231 após o silenciamento gênico de *ASXL2* (*Fold change* = 2). Os termômetros indicam, em vermelho, a expressão aumentada e, em azul, a expressão diminuída. Legenda dos símbolos e interações disponível no Anexo C.

Com relação à sinalização de TGF-beta, foi possível, após o silenciamento de *ASXL2*, identificar em ambas as linhagens a expressão aumentada de seus membros, indicando uma maior ativação da via nos modelos *in vitro* de CM avaliados. Além dos genes expostos na Figura 5.27, outros membros da via se encontram DE, tais como *BMP3* e *SMAD6* em MDA-MB-231.



**Figura 5.27 – Via de sinalização de TGF-beta nos modelos celulares silenciados.** Mapa da via de TGF-beta contendo os genes diferencialmente expressos em (1) MCF-7 e (2) MDA-MB-231 após o silenciamento gênico de ASXL2 (*Fold change* = 2). Os termômetros indicam, em vermelho, a expressão aumentada e, em azul, a expressão diminuída. Legenda dos símbolos e interações disponível no Anexo C.

Componentes do sistema imune foram identificados tanto pela análise de mapas de interação (String) quanto pelo enriquecimento de vias (Metacore®). Dentre eles, diversas interleucinas e seus receptores estão DE entre as linhagens, em geral, diminuídos em MCF-7 e

aumentados em MDA-MB-231, após o silenciamento de *ASXL2* (Tabela 5.2). Interleucinas relacionadas com as vias de PI3K-Akt e Ras foram identificadas como diminuídas em MCF-7, enquanto interleucinas relacionadas com as vias de NFκB, TGF beta e perfil IL17 estão aumentadas em MDA-MB-231.

**Tabela 5.2 – Interleucinas identificadas DE nas linhagens celulares MCF-7 e MDA-MB-231 após o silenciamento gênico de *ASXL2*.**

Gene	Diferença de expressão (si <i>ASXL2</i> vs <i>scramble</i> )	
	MCF-7	MDA-MB-231
<i>IL31</i>	-41.254425	-
<i>IL18</i>	-3.778779	-
<i>IL20RA</i>	-3.0827854	-
<i>IL15</i>	-2.4421966	-
<i>IL27RA</i>	2.04133	-
<i>IL24</i>	2.6553617	-
<i>IL21</i>	-	-3.8507588
<i>IL3</i>	-	-2.9630442
<i>IL36RN</i>	-	-2.219776
<i>IL41I</i>	-	2.016042
<i>IL17C</i>	-	2.0225925
<i>IL17B</i>	-	2.0438595
<i>IL21R</i>	-	2.218276
<i>IL36B</i>	-	2.3435435
<i>IL12RB2</i>	-	2.348554
<i>IL10RB</i>	-	2.491403
<i>IL1B</i>	-	2.8455312

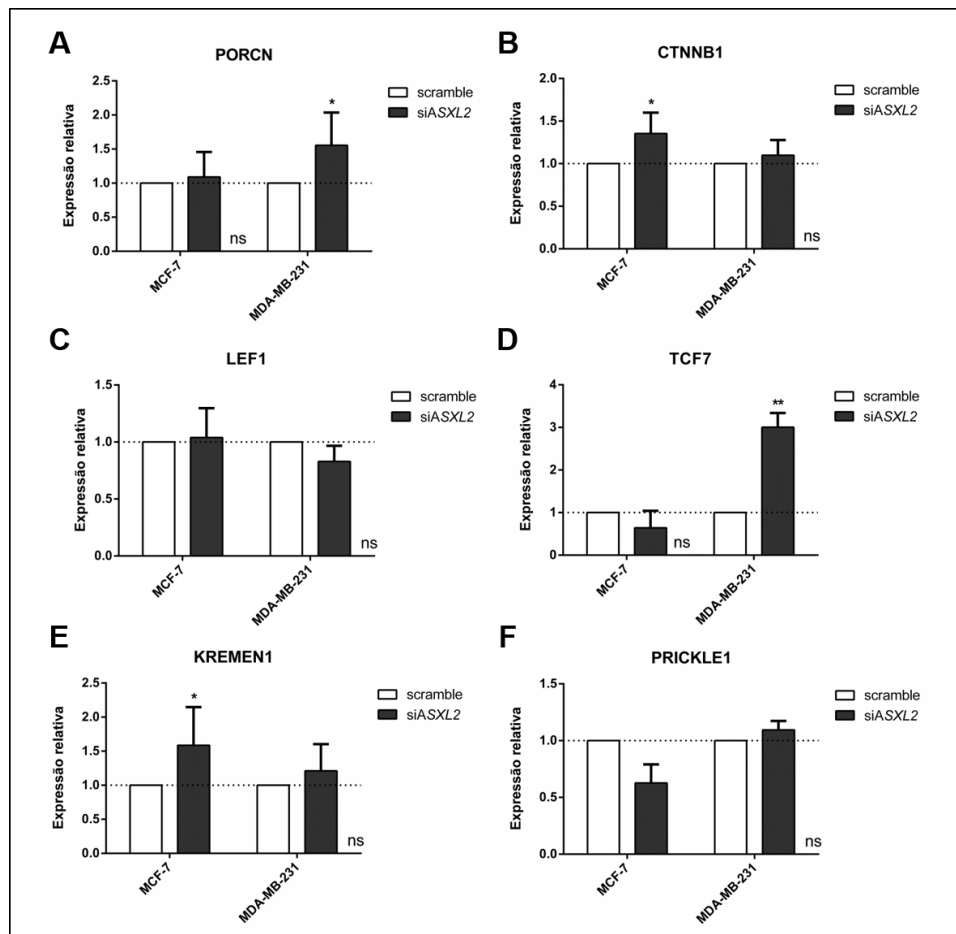
Diferenças de expressão representadas em *fold change*.

“-“-: Genes não DE (FC=2) nas linhagens.

Alguns dos genes encontrados como DE após o silenciamento de *ASXL2*, bem como genes associados às vias e processos biológicos de interesse, foram selecionados para validação e avaliação por RT-qPCR nas linhagens *scramble* e silenciadas. Os genes avaliados foram *PORCN*, *CTNNB1*, *LEF1*, *TCF7*, *KREMEN1*, *PRICKLE1*, *TGFB1*, *SMAD3*, *SMAD4* e *RUNX2*.

Avaliamos a expressão de membros da sinalização de Wnt, que foi encontrada alterada em ambas as linhagens no ensaio de microarranjo. O gene *PORCN*, envolvido em processamento das proteínas WNT na via canônica, está significativamente aumentado em MDA-MB-231 silenciada ( $p < 0,05$ ) (Figura 5.28A). O gene *CTNNB1*, que codifica a proteína β-catenina, co-ativadora dos fatores de transcrição da via canônica de Wnt, foi encontrado aumentado em ambas as linhagens após o silenciamento, principalmente em MCF-7 ( $p < 0,05$ ) (Figura 5.28B). Um dos fatores de transcrição da via canônica, o gene *LEF1*, que foi encontrado como diminuído em

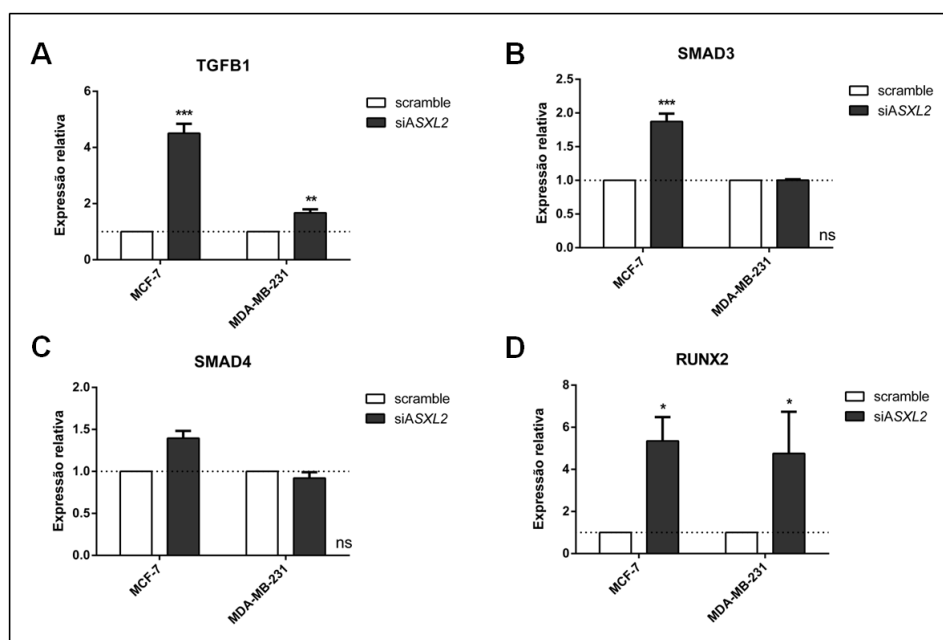
MCF-7 após o silenciamento (*fold change* de -2,23), permaneceu pouco alterado em ambas as linhagens (Figura 5.28C). Já um outro fator de transcrição, o gene *TCF7*, teve sua expressão pouco diminuída em MCF-7 e aumentada em cerca de 3 vezes ( $p < 0,01$ ) em MDA-MB-231 silenciada (Figura 5.28D). Estes resultados, combinados com aqueles observados no microarranjo, podem sugerir uma possível ativação da via canônica de Wnt na linhagem TN após o silenciamento de *ASXL2*. Dentre os inibidores da sinalização de Wnt, *KREMEN1*, que inibe a via canônica, está aumentado principalmente em MCF-7 ( $p < 0,05$ ) (Figura 5.28E). *PRICKLE1*, regulador negativo da via de polaridade planar, foi encontrado diminuído em MCF-7 e pouco aumentado em MDA-MB-231 (Figura 5.28F).



**Figura 5.28 - Expressão gênica relativa de genes envolvidos na sinalização de Wnt após o silenciamento gênico de *ASXL2* em MCF-7 e MDA-MB-231.** Quantificação relativa por RT-qPCR dos transcritos de (A) *PORCN*, (B) *CTNNB1*, (C) *LEF1*, (D) *TCF7*, (E) *KREMEN1* e (F) *PRICKLE1*, componentes da via de Wnt, com resultados normalizados pelo gene endógeno *GAPDH*, em linhagens celulares controle (*scramble*) e silenciadas para *ASXL2* (*siASXL2*) representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231). ns: diferença não significativa. \* $p < 0,05$ . \*\* $p < 0,01$ . Barras de erro indicativas de desvio padrão.

Dentre os componentes da via de sinalização de TGF beta analisados, o gene *TGFB1*, que codifica um dos ligantes que ativam a via, foi encontrado com a expressão aumentada em MCF-7 (*fold change* médio de 4,5,  $p < 0,001$ ) após o silenciamento, corroborando com os resultados do microarranjo. Este gene também está aumentado em cerca de 1,67 vezes ( $p < 0,01$ ) na linhagem Triplo Negativa (Figura 5.29A). *SMAD3* e *SMAD4*, capazes de interagir entre si para ativar a transcrição de diversos genes alvo da via, apresentaram expressão aumentada em MCF-7 e não alterada em MDA-MB-231 após o silenciamento (Figura 5.29B e 5.29C). Em MCF-7 silenciada para *ASXL2*, o aumento na expressão de *SMAD3* foi estatisticamente significativo ( $p < 0,001$ ).

O fator de transcrição *RUNX2*, por sua vez, apresentou níveis elevados de expressão em ambos os modelos celulares após o silenciamento de *ASXL2* (Figura 5.29D), com valores de *fold change* próximos a 5 ( $p < 0,05$ ). Este resultado corrobora com o aumento da expressão observado em MDA-MB-231 no experimento de microarranjo (*fold change* de 3).



**Figura 5.29 – Expressão gênica relativa de genes envolvidos em vias de sinalização após o silenciamento gênico de *ASXL2* em MCF-7 e MDA-MB-231.** Quantificação relativa por RT-qPCR dos transcritos de (A) *TGFB1*, (B) *SMAD3*, (C) *SMAD4*, componentes da via de TGF beta; e (D) *RUNX2*, fator de transcrição; com resultados normalizados pelo gene endógeno *GAPDH*, em linhagens celulares controle (*scramble*) e silenciadas para *ASXL2* (*siASXL2*) representativas dos subtipos Luminal A (MCF-7) e Triplo Negativo (MDA-MB-231). ns: diferença não significativa. \* $p < 0,05$ . \*\* $p < 0,01$ . \*\*\* $p < 0,001$ . Barras de erro indicativas de desvio padrão.

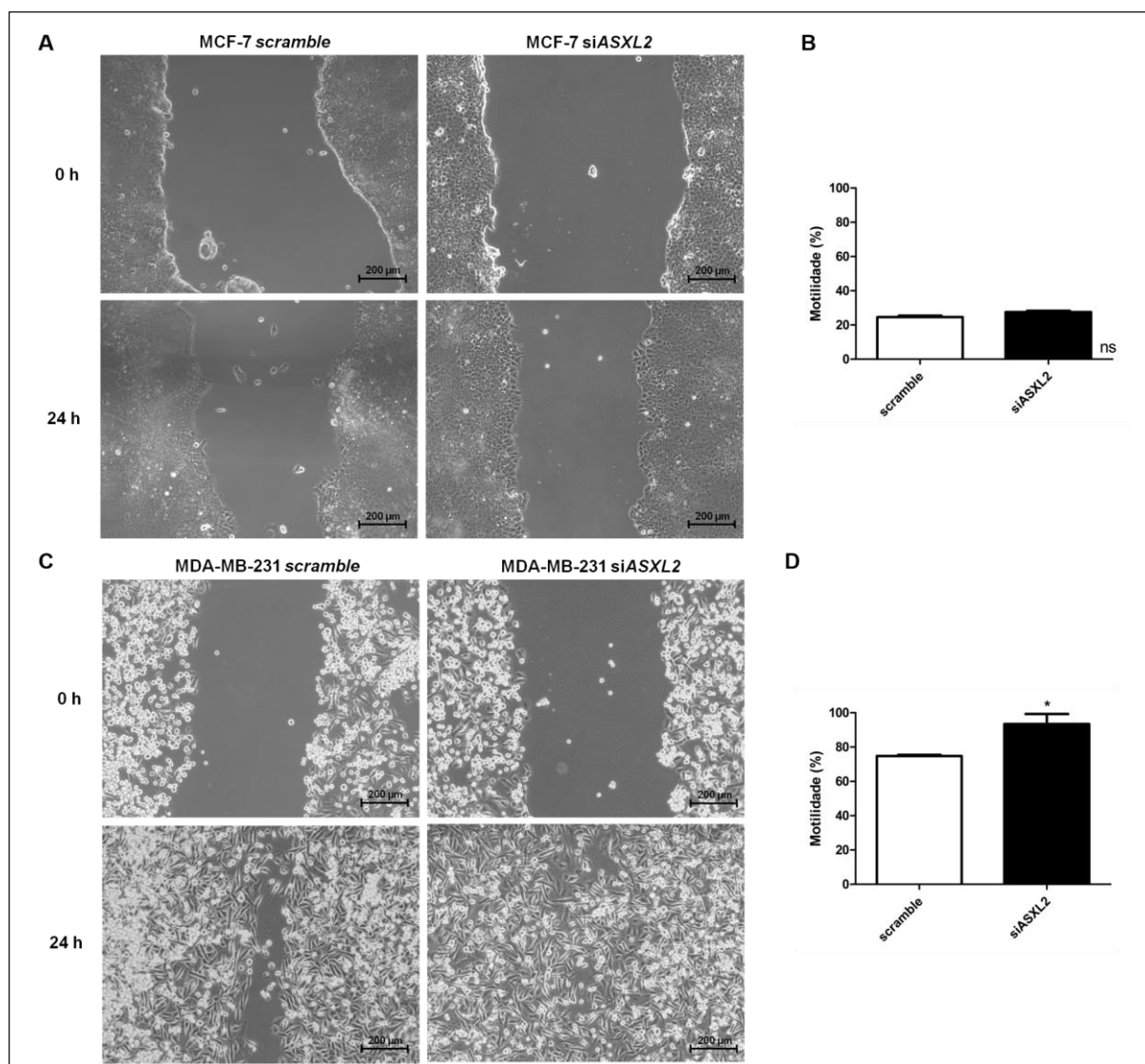


De forma geral, os dados obtidos com a abordagem utilizada foram capazes de mostrar que *ASXL2* pode contribuir direta ou indiretamente para a regulação de vias de sinalização importantes para o processo de tumorigênese em ambos os modelos de CM avaliados.

Uma vez que o silenciamento gênico de *ASXL2* foi capaz de identificar vias de sinalização envolvidas com adesão, regulação do citoesqueleto, migração e motilidade celular, realizamos o ensaio funcional de *wound healing* em células MCF-7 e MDA-MB-231 silenciadas (*siASXL2*) ou não (*scramble*), a fim de avaliar o efeito do mesmo no potencial migratório de células de CM.

Em MCF-7, não foram observadas alterações estatisticamente significativas ( $p=0,0672$ ) no estreitamento da fenda entre as condições *scramble* e *siASXL2* (Figura 5.30A e 5.30B). Em 24 h, o percentual de fechamento na linhagem silenciada comparada com a linhagem controle aumentou em cerca de 3%, indicando que, neste modelo, *ASXL2* pode causar baixo impacto na motilidade celular.

Dentre os subtipos moleculares de CM, o Triplo Negativo, por ser o mais agressivo e com maior potencial metastático, é também o que possui maior potencial migratório. Na linhagem MDA-MB-231, após 24 h, foi possível observar um fechamento mais eficiente da fenda no modelo silenciado quando comparado com o controle (Figura 5.30C). O silenciamento de *ASXL2* aumentou em cerca de 20% ( $p<0,05$ ) o potencial de migração na linhagem MDA-MB-231 (Figura 5.30D), o que pode estar associado à regulação direta ou indireta de vias e genes envolvidos neste processo.



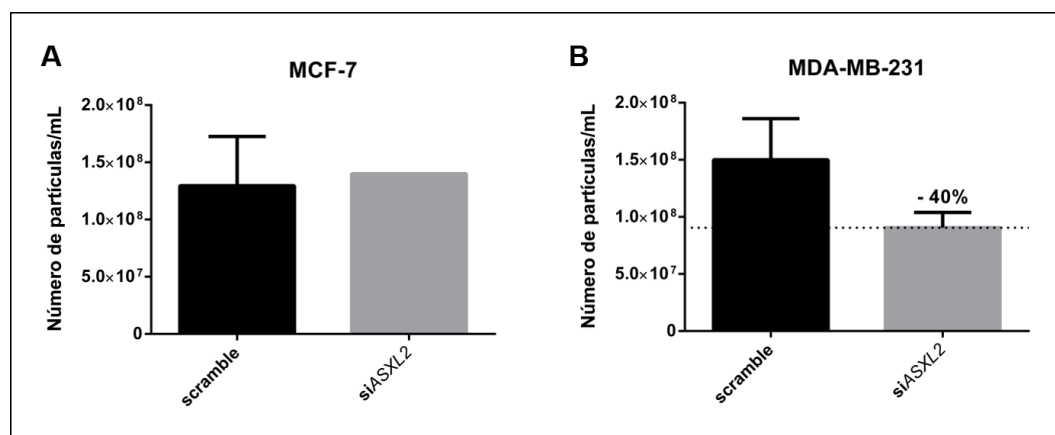
**Figura 5.30 – Ensaio de *wound healing* após silenciamento de *ASXL2* em MCF-7 e MDA-MB-231.** Imagens representativas da avaliação da migração celular após 24 h de silenciamento do gene *ASXL2* nas linhagens celulares (A) MCF-7 e (C) MDA-MB-231. Aumento de 100x. (B) e (D) Gráfico representativo do potencial de motilidade celular, indicado pelo percentual de fechamento do espaço aberto no tempo inicial (0 h) em MCF-7 e MDA-MB-231, respectivamente (n=3). ns = não significativo. \*p<0,05. Barras de erro indicativas de desvio padrão.

A análise com o software PANTHER também permitiu identificar a expressão diferencial de genes relacionados com tráfego de membrana e formação de vesículas extracelulares após o silenciamento de *ASXL2*, em especial os membros da família das sinaptotagminas (SYT), envolvidas com exocitose de neurotransmissores e hormônios. Em MCF-7, o gene *SYT8* foi encontrado diminuído em 2 vezes após o silenciamento. Em MDA-MB-231, por sua vez, quatro genes desta família apresentaram expressão diminuída com o silenciamento de *ASXL2*: *SYT1* (fold change de -3,9), *SYT6* (fold change de -2,5), *SYT9* (fold change de -2,1) e *SYT15* (fold change de -2,3). Além destes, os genes *SYP* e *SYTL3*, que também possuem papel no tráfego de

vesículas, foram encontrados com expressão diminuída neste modelo (*fold change* de -2,2 e -3,8, respectivamente).

Uma vez que genes envolvidos com tráfego de membrana e formação de vesículas extracelulares foram encontrados DE nas linhagens celulares silenciadas, principalmente no modelo TN, optamos por avaliar a contagem de microvesículas no sobrenadante das células silenciadas ou não, a fim de investigar se há diferença na liberação de vesículas entre os modelos após o silenciamento.

Não houve diferença na concentração de microvesículas encontradas no sobrenadante de MCF-7 (silenciada versus *scramble*), como podemos observar na Figura 5.31A. Entretanto, em MDA-MB-231, observou-se uma diminuição da concentração de microvesículas em aproximadamente 1,7 vezes, correspondendo a 40% (Figura 5.31B). Estes resultados evidenciam uma ação de ASXL2 na formação de vesículas extracelulares.

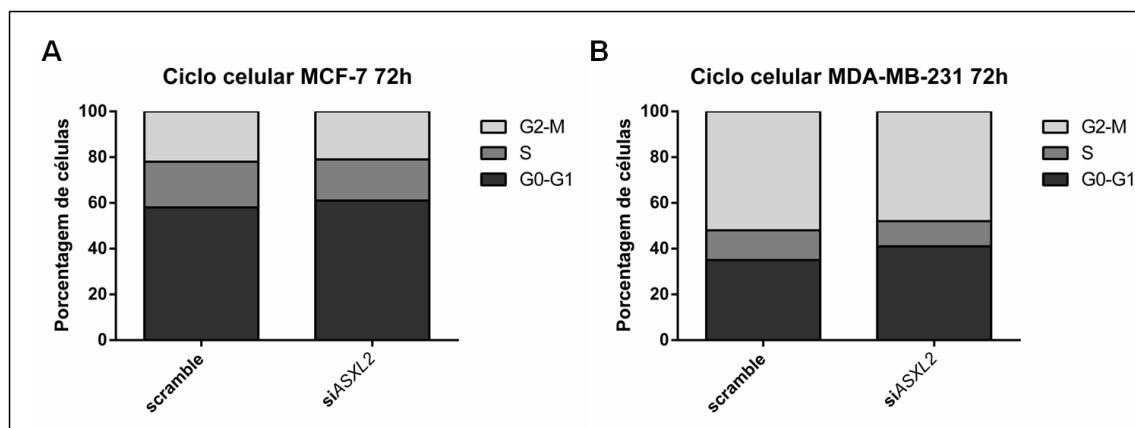


**Figura 5.31 – Avaliação da contagem de microvesículas no sobrenadante das culturas celulares de MCF-7 e MDA-MB-231 após silenciamento de ASXL2.** Gráficos representativos do número de microvesículas identificadas a cada 1 mL de sobrenadante nas linhagens (A) MCF-7 e (B) MDA-MB-231, silenciadas ou não para ASXL2 (n=2). Barras de erro indicativas de desvio padrão.

Além disso, uma vez que vias e genes relacionados a ciclo celular foram encontrados DE nas linhagens silenciadas para ASXL2, avaliamos, por citometria de fluxo, a porcentagem de células em cada fase do ciclo celular nas linhagens MCF-7 e MDA-MB-231 (*scramble* e *siASXL2*), comparando-as, a fim de elucidar se o silenciamento de ASXL2 pode causar efeitos diretos no ciclo.

Na Figura 5.32, é possível observar que houve uma pequena diminuição na quantidade de células nas fases S + G2-M após o silenciamento, levando a um acréscimo de cerca de 17% de células em G0-G1 na linhagem MDA-MB-231, enquanto que em MCF-7 o aumento foi menos

pronunciado (5%). Este experimento é sugestivo do envolvimento de ASXL2 na regulação do ciclo celular (n=1).

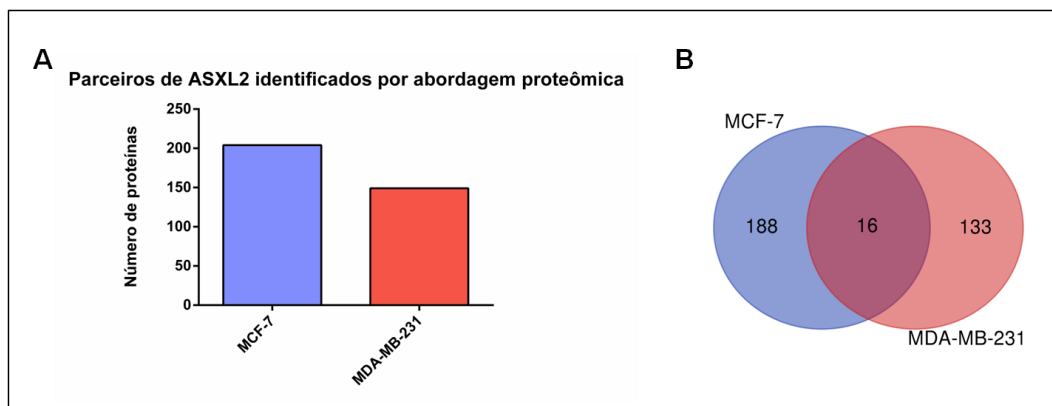


**Figura 5.32 – Avaliação do ciclo celular após silenciamento de ASXL2 em MCF-7 e MDA-MB-231.** Gráficos representativos do percentual de células das linhagens (A) MCF-7 e (B) MDA-MB-231, silenciadas ou não para ASXL2, em cada fase do ciclo celular (n=1). As células foram marcadas com solução contendo iodeto de propídio e avaliadas em citômetro de fluxo (canalFL3 670LP). Total de 10.000 eventos avaliados em cada condição.

### 5.5 ASXL2 associa-se a proteínas envolvidas com modificações de histona e remodelamento de cromatina em modelos *in vitro* de CM

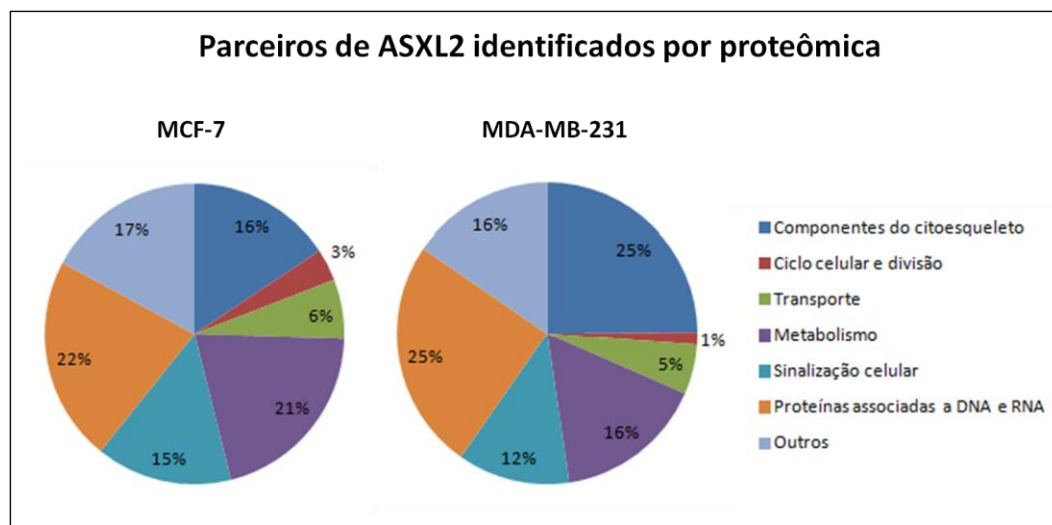
A fim de investigar a função ETP de ASXL2 no CM, buscamos identificar os potenciais parceiros desta proteína nos subtipos moleculares Luminal A e Triplo Negativo. Para tal, realizamos a imunoprecipitação de ASXL2 nas linhagens MCF-7 e MDA-MB-231, seguida de análise proteômica dos imunoprecipitados.

Após a exclusão de redundâncias e similaridades com os imunoprecipitados de IgG normal, foi possível identificar um total de 204 proteínas imunoprecipitadas com ASXL2 na linhagem MCF-7, e um total de 149 proteínas para a linhagem MDA-MB-231. Dentre as proteínas identificadas, 16 foram comuns a ambas as linhagens (Figura 5.33).



**Figura 5.33 – Parceiros de ASXL2 identificados por abordagem proteômica.** Foram identificados um total de 204 proteínas imunoprecipitadas com ASXL2 na linhagem MCF-7, e um total de 149 na linhagem MDA-MB-231, após normalização com IgG em ambas as linhagens. **(A)** Gráfico de barras e **(B)** Diagrama de Venn representativos das proteínas identificadas com ASXL2 em MCF-7 e MDA-MB-231.

As proteínas identificadas foram categorizadas baseadas na ontologia gênica em um total de 7 classes funcionais. As categorias funcionais são: (1) Componentes do citoesqueleto; (2) Ciclo celular e divisão; (3) Transporte; (4) Metabolismo; (5) Sinalização celular; (6) Proteínas associadas a DNA e RNA; e (7) Outros, esta última contendo proteínas envolvidas em outros processos ou de função pouco conhecida. A porcentagem de proteínas identificadas em cada classe pode ser observada na Figura 5.34, e as listagens das proteínas por classe encontram-se nos Anexos D e E. Em ambas as linhagens foi possível identificar pelo menos uma proteína pertencente a cada grupo.



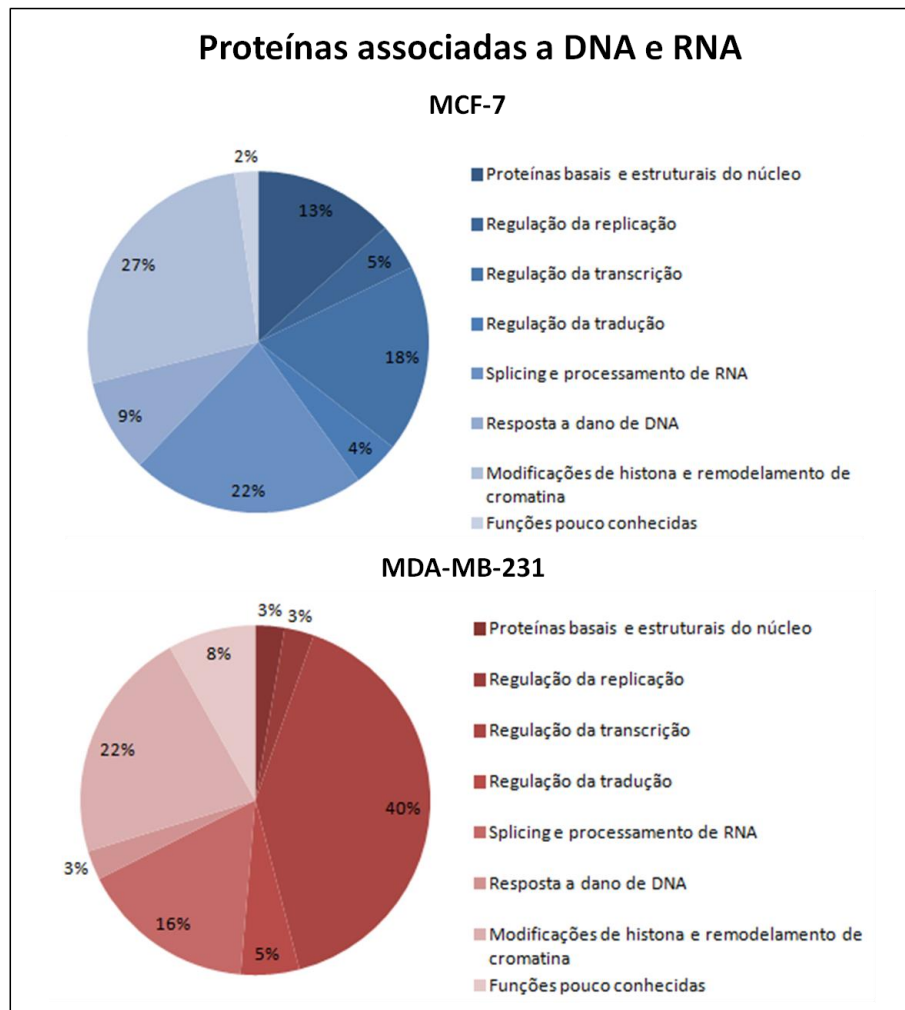
**Figura 5.34 – Classificação funcional de parceiros de ASXL2 identificados por abordagem proteômica.** Gráficos de pizza indicativos da função de proteínas imunoprecipitadas com ASXL2 nas linhagens MCF-7 e MDA-MB-231. Valores expressos em porcentagem, em relação ao número total de proteínas identificadas em cada linhagem.

Em MCF-7, a maioria dos parceiros de ASXL2 identificados são proteínas associadas a DNA e RNA (22%). Logo em seguida figura a classe Metabolismo (21%), cujos membros são, principalmente, enzimas envolvidas em diversos processos metabólicos da célula. Outros grupos com grande representação nesta linhagem são Componentes do citoesqueleto (16%), tais como queratinas e tubulinas, e Sinalização celular (15%), classe na qual figuram diversos receptores e transdutores de sinal.

Em MDA-MB-231, as classes com maior representatividade são Componentes do citoesqueleto e Proteínas associadas a DNA e RNA, que juntas correspondem a 50% das proteínas identificadas em associação com ASXL2 nesta linhagem (25% de identificação em cada classe). Proteínas envolvidas em metabolismo e sinalização celular somam 28% do total.

Das 16 proteínas comuns a ambas as linhagens, 8 são associadas a DNA e RNA, enquanto as demais estão relacionadas a metabolismo, sinalização, transporte ou funções desconhecidas.

As proteínas associadas a DNA e RNA identificadas após a imunoprecipitação (45 em MCF-7 e 37 em MDA-MB-231) foram subdivididas em outras 8 classes, de acordo com o processo com o qual estão envolvidas: (1) Proteínas basais e estruturais do núcleo, (2) Regulação da replicação, (3) Regulação da transcrição, (4) Regulação da tradução, (5) *Splicing* e processamento de RNA, (6) Resposta a dano de DNA, (7) Modificações de histona e remodelamento de cromatina, e (8) Funções pouco conhecidas. A porcentagem de proteínas de cada subclasse pode ser observada na Figura 5.35, e as listagens por subclasse estão disponíveis nos Anexos F e G.



**Figura 5.35 – Classificação funcional de parceiros de ASXL2 associados a DNA e RNA identificados por abordagem proteômica.** Gráficos de pizza indicativos da função de proteínas associadas a DNA e RNA imunoprecipitadas com ASXL2 nas linhagens MCF-7 e MDA-MB-231. Valores expressos em porcentagem, em relação ao número total de proteínas associadas a DNA e RNA identificadas em cada linhagem.

Em MCF-7, 12 proteínas estão envolvidas em modificações de histona e remodelamento de cromatina, o que corresponde a 27% das proteínas associadas a DNA e RNA identificadas nesta linhagem. Na linhagem MDA-MB-231, por sua vez, 8 proteínas foram identificadas nesta subclasse, totalizando 22% do total de proteínas associadas a DNA e RNA identificadas.

Outras subclasses com grande representatividade em ambas as linhagens foram Regulação da transcrição (18% em MCF-7 e 40% em MDA-MB-231) e *Splicing* e processamento de RNA (22% em MCF-7 e 16% em MDA-MB-231).

As proteínas presentes na subclasse “Função pouco conhecida” são aquelas que, apesar de sabidamente serem capazes de se ligar a moléculas de DNA e/ou RNA, ainda não tiveram seu

papel completamente elucidado, e representam 3% das proteínas associadas a DNA e RNA identificadas em MCF-7 e 8% em MDA-MB-231.

Dentre os modificadores de histona e remodeladores de cromatina, estão presentes membros de complexos de acetilação, proteínas que recrutam histona deacetilases, demetilases, remodeladores de cromatina dependentes de ATP, dentre outros. Dois deles, as proteínas ANKRD11 e JADE1 (PHF17), foram identificados associados a ASXL2 em ambas as linhagens celulares. As proteínas envolvidas em modificação de histonas e remodelamento de cromatina encontradas em cada linhagem celular, bem como sua função de acordo com a literatura, estão discriminadas na Tabela 5.3.

Dentre as proteínas identificadas em associação com ASXL2 na linhagem MCF-7, destacamos CBX8, cujo transcrito correspondente foi encontrado com expressão aumentada nesta linhagem, conforme mostrado anteriormente; e MLL1, corroborando com os resultados de colocalização com ASXL2 previamente observados. Em MDA-MB-231, destaca-se BPTF, encontrado com menor expressão gênica nesta linhagem.

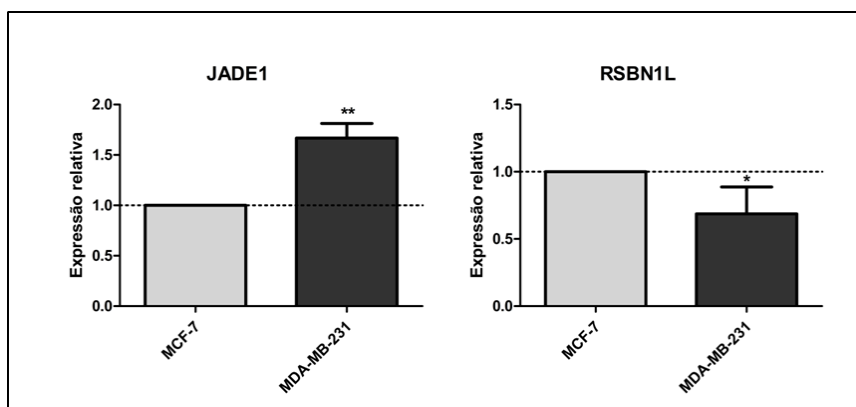


**Tabela 5.3 – Parceiros de ASXL2 envolvidos em modificação de histonas e remodelamento de cromatina identificados por abordagem proteômica.**

<b>Símbolo</b>	<b>Nome da proteína</b>	<b>Função</b>
<b>MCF-7</b>		
ANKRD11	Ankyrin repeat domain-containing protein 11	Inibe ativação da transcrição, pode recrutar HDACs.
BAZ2B	Bromodomain adjacent to zinc finger domain protein 2B	Potencial papel em ativação da transcrição, com preferência pela marca H3K14ac.
BCOR	BCL-6 corepressor	Corepressor transcricional, interage com HDACs.
CBX8	Chromobox protein homolog 8	Componente do complexo PRC1, que promove a marca de silenciamento H2AK119ub.
JADE 1 (PHF17)	Protein Jade-1	Componente do complexo HBO1, que promove acetilação em H4.
JADE3 (PHF16)	Protein Jade-3	Componente do complexo HBO1, que promove acetilação em H4.
KAT6A	Histone acetyltransferase KAT6A	Componente do complexo MOZ/MORF, promove acetilação em H3K9.
KAT6B	Histone acetyltransferase KAT6B	Componente do complexo de acetilação de H3 MOZ/MORF.
KDM5A (RBBP2)	Lysine-specific demethylase 5A	Promove demetilação em H3K4.
KMT2A (MLL1)	Histone-lysine N-methyltransferase 2A	Subunidade catalítica do complexo MLL, promove a marca de ativação H3K4me.
SMARCAD1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1	Membro da família SWI/SNF, medeia desacetilação em H3 e H4.
ZMYND11	Zinc finger MYND domain-containing protein 11	Reconhece H3.3K36me3, atua como corepressor da transcrição por regular a RNAPol II.
<b>MDA-MB-231</b>		
ANKRD11	Ankyrin repeat domain-containing protein 11	Inibe ativação da transcrição, pode recrutar HDACs.
ANKRD12	Ankyrin repeat domain-containing protein 12	Inibe ativação da transcrição, pode recrutar HDACs.
BPTF	Nucleosome-remodeling factor subunit BPTF	Membro da família NURF, reconhece H3K4me3 e facilita a transcrição.
FAM208A	Protein TASOR	Componente do complexo HUSH, que medeia a repressão em H3K9me3.
GTF3C4	General transcription factor 3C polypeptide 4	Possui atividade de HAT específica para H3.
JADE 1 (PHF17)	Protein Jade-1	Componente do complexo HBO1, que promove acetilação em H4.
RSBN1L	Round spermatid basic protein 1-like protein	Parálogo de RSBN1, que demetila H4K20me2.
SMARCD3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3	Membro da família SWI/SNF, envolvida em remodelamento de cromatina.

Por meio de PCR quantitativo em tempo real, avaliou-se a expressão gênica de *JADE1*, cuja proteína foi identificada associada a *ASXL2* em ambas as linhagens, e *RSBN1L*, cuja proteína foi identificada em MDA-MB-231, a fim de verificar se há diferenças de expressão nestes genes entre as linhagens celulares estudadas. A partir desta análise, foi possível observar

que *JADE1* possui maior expressão relativa em MDA-MB-231, enquanto *RSBN1L* encontra-se com a expressão diminuída nesta linhagem quando comparada com MCF-7 (Figura 5.36).

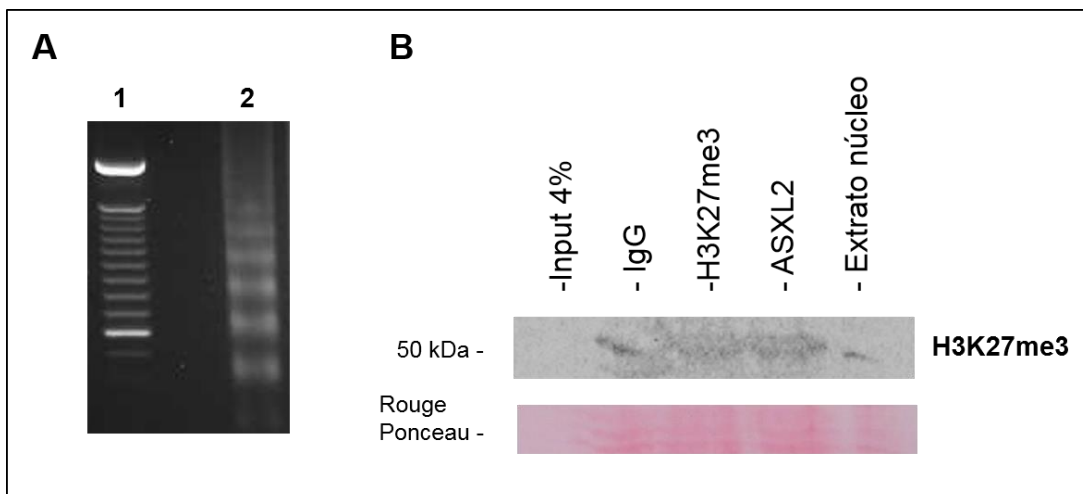


**Figura 5.36 – Expressão gênica relativa de *JADE1* e *RSBN1L* nas linhagens celulares MCF-7 e MDA-MB-231.** Quantificação relativa por RT-qPCR, com resultados normalizados pelo gene endógeno *GAPDH*, dos genes *JADE1* e *RSBN1L* em MCF-7 e MDA-MB-231 (n=3). \* p<0,05, \*\* p<0,01. Barras de erro indicativas de desvio padrão.

## 5.6 ASXL2 colocaliza-se com uma marca de silenciamento da expressão gênica na linhagem MCF-7

Uma vez que os dados obtidos por análise proteômica sugerem que ASXL2 está relacionado com modificadores de histona e remodeladores de cromatina de diversas famílias, envolvidos tanto em ativação quanto repressão da expressão gênica, decidiu-se avaliar sua colocalização na cromatina com um marcador de silenciamento, H3K27me3.

Para tal, foi realizado o ensaio de imunoprecipitação de cromatina (ChIP) de ASXL2 e H3K27me3 na linhagem MCF-7. Os imunoprecipitados foram submetidos ao experimento de WB com anticorpo específico para a marca de metilação. Através deste ensaio, foi possível confirmar a imunoprecipitação de H3K27me3 e verificar que esta marca de silenciamento está associada a ASXL2 na cromatina (Figura 5.37).



**Figura 5.37 – Colocalização entre H3K27me3 e ASXL2 em cromatina da linhagem MCF-7.** (A) Resolução em gel de agarose 1% da cromatina da linhagem MCF-7 pós-digestão. 1- Padrão de peso molecular de 100 pb; 2- Cromatina digerida. (B) Análise da expressão de H3K27me3 por WB em produtos do ensaio de ChIP de IgG (controle), H3K27me3 e ASXL2, e em extrato nuclear de MCF-7. A totalidade dos produtos de ChIP e 30  $\mu$ g de extrato de proteínas nucleares da linhagem MCF-7 foram resolvidos em gel SDS-PAGE 15%, transferidos para membrana de nitrocelulose e incubados com anticorpo anti-H3K27me3 em diluição 1:350. Controle de carregamento da membrana de nitrocelulose representativo do WB por coloração com Rouge Ponceau.

## 6. DISCUSSÃO

As alterações genéticas e epigenéticas tem papel importante na carcinogênese, uma vez que são capazes de promover alterações na expressão gênica. Atualmente, sabe-se que ambas agem em conjunto no câncer, e que modificações epigenéticas podem de fato contribuir para iniciar a tumorigênese, sugerindo que alterações no epigenoma possam ser a chave para a iniciação e progressão de diversos tumores (YOU e JONES, 2012). A metilação do DNA foi o primeiro mecanismo epigenético descrito, e é mais amplamente estudado, enquanto a ação de ncRNAs tem sido mais explorada recentemente (TOLLERVEY e LUNIAK, 2012).

O mecanismo epigenético menos compreendido é o remodelamento da cromatina. Este mecanismo é complexo, pois envolve diversos tipos de modificações que causam efeitos diferentes; e dinâmico, uma vez que as modificações são reversíveis e podem ser adicionadas e removidas por diversas proteínas, agindo individualmente e/ou em complexos. Em câncer, alguns modificadores de histona e remodeladores de cromatina foram descritos como mutados ou desregulados, destacando a importância de seu estudo na contribuição para esta doença.

Membros das famílias de modificadores de histona Pc e Trx foram descritos inicialmente em *Drosophila melanogaster*, com função de manter estados silenciados ou ativos da cromatina, respectivamente, em genes específicos durante o desenvolvimento embrionário do animal, através de uma fina regulação espaço-temporal. As proteínas Pc e Trx são conservadas em organismos multicelulares, inclusive em humanos (STEFFEN e RINGROSE, 2014).

Em um nível acima dos efetores das modificações de histonas, encontram-se seus possíveis reguladores, os ETPs, também descritos em *Drosophila*, que apresentam homólogos em humanos. Apesar da primeira descrição dos ETPs ter ocorrido em 2000, pelos autores Gildea, Lopez e Shearn, pouco ainda se sabe sobre sua função no balanço entre silenciamento e ativação de genes em processos celulares importantes, como no desenvolvimento embrionário, e nas doenças, incluindo o câncer.

Em humanos, membros da família de ETPs ASXL foram descritos como mutados em câncer, incluindo o CM. Porém, a maior parte dos dados disponíveis sobre esta família está associada com mutações em tumores hematológicos. Pouco se sabe sobre as proteínas ASXL em termos funcionais e de expressão, e seu papel em tumores sólidos, como o CM, também é pouco explorado. Além disso, o fato das proteínas ASXL terem um domínio de ligação ao DNA sugere mecanismos de atuação adicionais, quando comparados com seus homólogos em *Drosophila*.

Dados em CM mostraram que linhagens celulares positivas para RE apresentam expressão de ASXL2 (PARK *et al.*, 2016). Assim, a realização de um *screening* de sua expressão em outros modelos *in vitro* de CM, representantes dos diferentes subtipos moleculares da doença, evidenciou que ASXL2 se encontra diferencialmente expresso entre as mesmas. Nossos resultados corroboram com dados de Park *et al.* (2016), uma vez que identificamos em nosso trabalho uma alta expressão de ASXL2 em MCF-7; entretanto nossos dados também apontaram como modelo de interesse a linhagem MDA-MB-231, do subtipo TN, que apresentou alta expressão de ASXL2 e presença desta proteína no núcleo.

As linhagens celulares MCF-7 e MDA-MB-231 representam dois subtipos moleculares que diferem em relação à agressividade, grau histológico e prognóstico em CM, e uma vez que dados integrados sobre a expressão e localização de modificadores de histona em CM são limitados, a avaliação de membros Pc, Trx e remodeladores de cromatina nestes modelos nos permitiu elucidar que membros destes complexos estão distribuídos de forma diferente entre os modelos *in vitro* de CM.

Os complexos da família Pc, PRC1 e PRC2, foram descritos por ocupar os mesmos alvos na cromatina. Porém, o mecanismo exato de recrutamento destes complexos e sua função em conjunto ainda são pouco compreendidos. De acordo com a hipótese clássica, proposta em *Drosophila*, PRC2 é recrutado para alvos não-modificados na cromatina e promove sua marca característica de metilação em H3K27. Esta marca, por sua vez, é reconhecida por uma proteína CBX, componente do complexo PRC1, que em seguida insere a modificação H2AK119ub. Porém, um novo mecanismo foi descoberto recentemente, no qual PRC1 é o primeiro complexo a promover modificação no alvo, e a partir desta, PRC2 é recrutado para inserir sua marca característica de silenciamento. Além de atuarem em conjunto, PRC2 e PRC1 podem ser recrutados individualmente para a cromatina, e alguns de seus membros também são capazes de atuar isoladamente, fora do complexo (BLACKLEDGE, ROSE e KLOSE, 2015).

Em geral, os membros de PRC2 avaliados foram encontrados com expressão mais aumentada em MDA-MB-231 em comparação com MCF-7, dentre eles um dos efetores do complexo, a proteína EZH2. O papel de EZH2 no câncer é amplamente estudado, e sabe-se que a expressão alterada desta proteína está associada com diversos tipos de neoplasias, inclusive o CM. Holm *et al.* (2012) identificaram expressão gênica e proteica de EZH2 aumentada nas linhagens MCF-7 e MDA-MB-231, com um aumento mais significativo na linhagem TN, corroborando com nossos resultados.

Em uma análise utilizando amostras de tumores extraídas de pacientes com CM, Pietersen *et al.* (2008) identificaram, por PCR em tempo real, baixa expressão de *EZH2* em tumores classificados como Luminal A e superexpressão deste gene no subtipo TN, também corroborando com nossos resultados. Além disso, os autores relacionaram altos níveis de mRNA de *EZH2* com diminuição na sobrevivência e aumento de proliferação em CM. Também, altos níveis de *EZH2* foram reportados como marcadores de agressividade, pior prognóstico, grau histológico mais avançado, maior índice de proliferação (Ki-67), e diferença na expressão de receptores, estando relacionados com o fenótipo TN (YOO e HENNIGHAUSEN, 2012; JIANG *et al.*, 2015; WANG *et al.*, 2015; NEUSQUEN *et al.*, 2016; WASSEF e MARGUERON, 2016).

SUZ12 e EED, também membros de PRC2, possuem aumento de expressão na linhagem MDA-MB-231. Estes já foram descritos como aumentados em outros tipos de câncer, como leucemias, câncer de pulmão, câncer colorretal e câncer gástrico (PIZZATTI *et al.*, 2010; XIA *et al.*, 2015; LIU *et al.*, 2016; SAN *et al.*, 2017).

Os membros de PRC2 podem regular diversos genes, dentre eles supressores tumorais, tais como *RUNX3* e *CDKN1C*, que inibem proliferação; *FOXC1* e *CDH1* (E-caderina), envolvidos em inibição de migração e invasão; e *RAD51*, com função em reparo de danos ao DNA (YOO e HENNIGHAUSEN, 2012; YAMAGUCHI e HUNG, 2014; XIA *et al.*, 2015; KIM e ROBERTS, 2016; HU *et al.*, 2018). Em linhagens celulares e tumores de CM do tipo TN, o gene *WWC1*, que também possui função de supressão tumoral e inibição da transição epitélio-mesenquimal, encontra-se silenciado por *EZH2*, através de H3K27me3, em combinação com a metilação de seu promotor por DNMT1 (LIU *et al.*, 2018). Em tumores TN, *EZH2* também é capaz de silenciar genes *TIMP*, conhecidos por inibir a ação de metaloproteinases (MMPs) na promoção de metástase. Este silenciamento resulta em aumento na expressão de MMPs e na atividade invasiva destes tumores (CHIEN *et al.*, 2018).

Assim, o silenciamento de genes alvo por PRC2, ao favorecer processos como proliferação e migração, pode contribuir para a tumorigênese e fenótipos agressivos no CM, o que corrobora com nossos resultados, uma vez que observamos alta expressão de membros de PRC2 e H3K27me3 nos dois modelos avaliados, mas principalmente no modelo TN, o mais agressivo.

Quanto ao complexo PRC1, BMI1, o membro mais estudado em câncer, está altamente expresso em tumores de mama RE positivos, geralmente associado a um melhor prognóstico (PIETERSEN *et al.*, 2008; CHOI *et al.*, 2009; WANG *et al.*, 2014). Nossos resultados

corroboram com estes dados da literatura, uma vez que observamos marcação mais intensa de BMI1 no núcleo de células MCF-7, positivas para RE, em comparação com as células MDA-MB-231. Em outros tipos de câncer, como leucemias e hepatocarcinomas, a alta expressão de BMI1 está relacionada com pior prognóstico (PIETERSEN *et al.*, 2008).

Em células MCF-7 superestimuladas com o ligante do receptor de estrogênio 17 $\beta$ -estradiol, Wang *et al.* (2014) observaram um aumento na expressão de BMI1, com subsequente diminuição na expressão de p16-INK4a, um de seus alvos. A expressão aumentada de BMI1 pode contribuir para o tumor através do silenciamento deste alvo, que está envolvido em controle do ciclo celular, supressão tumoral e ativação de TP53, levando assim a um aumento na proliferação (PIETERSEN *et al.*, 2008; RAMAIAH e VAISHNAVE, 2018).

Dados sobre a ação da proteína RING1, um dos membros efetores do complexo PRC1, na regulação epigenética do câncer são escassos. Nosso estudo mostrou pouca diferença de expressão de *RING1* dentre as linhagens avaliadas. Shen *et al.* (2018) propuseram uma função de RING1 independente do complexo PRC1 em hepatocarcinoma e câncer colorretal, onde RING1 pode regular p53 diretamente, através de ubiquitinação desta proteína supressora de tumor, resultando em sua degradação e contribuindo para a proliferação e sobrevivência de células tumorais.

Na literatura, há poucas informações disponíveis sobre o papel de CBX8 em CM. Em outros tipos de câncer, porém, como carcinoma hepatocelular, câncer de bexiga e câncer colorretal, CBX8 está associado com um potencial oncogênico e pior prognóstico, por exemplo através de sua ação em complexo com BMI1, que contribui para o silenciamento do supressor de tumor p16-INK4a (MAERTENS *et al.*, 2009; YUAN *et al.*, 2017; SONG *et al.*, 2018; ZHANG *et al.*, 2018). Entretanto, encontramos *CBX8* significativamente diminuído na linhagem TN, de pior prognóstico, quando comparado com MCF-7, sugerindo que este gene possa ter um papel diferente em CM.

A proteína YY1 é um fator de transcrição capaz de reprimir e ativar diversos promotores, e pode agir como membro Pc através de ligação a alvos no DNA e recrutamento de complexos Pc para estabelecimento e manutenção de estados silenciados (WAN *et al.*, 2012). YY1 foi descrita como superexpressa em diversos tipos de neoplasias, incluindo câncer gástrico, câncer de próstata e CM (KAUFHOLD, GARBÁN e BONAVIDA, 2016). Neste último, Wan *et al.* (2012) observaram alta expressão de YY1 e localização predominantemente nuclear em MCF-7, corroborando com nossos resultados. Segundo os autores, YY1 também está altamente expressa

em tumores, quando comparado com células de mama normal, e localiza-se principalmente na fração nuclear das células, conforme observamos em nossos modelos, nos quais esta proteína foi encontrada em abundância no núcleo. YY1 pode regular a transcrição de diversos alvos, incluindo *HER2*, sendo capaz de ativar este oncogene, *p53* e HDACs. Assim como RING1, YY1 pode promover a ubiquitinação e consequente degradação de proteínas alvo, incluindo os supressores de tumor *p53* e *p27*. Recentemente, YY1 foi implicado na fosforilação e ativação de AKT, que está associada com um pior prognóstico (WAN *et al.*, 2012; BASU *et al.*, 2014; ZHANG *et al.*, 2016).

O estado ativado da cromatina depende de vários fatores, dentre eles a ação do complexo MLL, de função antagônica a Pc. Assim como observamos dentre os membros de PRC1, houve pouca alteração na expressão gênica dos componentes deste complexo nas linhagens MCF-7 e MDA-MB-231. Entretanto, observamos que o membro efetor do grupo, MLL1, apresenta localização diferencial entre as linhagens, sendo mais abundante no núcleo do modelo TN.

Em 2013, Rabello *et al.* identificaram um perfil de expressão de genes *KMT2 (MLL)* similar entre linhagens celulares de CM, ou seja, com pouca alteração entre os modelos, corroborando com nossos resultados. Em comparação com amostras de mama normal, os autores observaram expressão diminuída dos genes *MLL* em linhagens celulares e amostras de CM, sugerindo uma desregulação destes genes na doença.

De acordo com Jeong *et al.* (2011), no núcleo de células MCF-7 superestimuladas com o ligante 17 $\beta$ -estradiol, RE é capaz de recrutar para seus alvos complexos SWI/SNF, que promovem o remodelamento e abertura da cromatina e em seguida recrutam MLL1 para promoção da metilação em H3K4 e consequente manutenção de um estado ativado nos genes alvo de RE. Em nosso estudo, entretanto, em termos de localização nesta linhagem, MLL1 foi encontrada em maior abundância no citoplasma. Especulamos que esta localização pode estar associada ao fato de que não superestimulamos nossas células com o ligante de RE. Em condições de ativação, MLL1 poderia ser translocada para o núcleo para exercer sua função na ativação de genes alvo.

Em geral, dados da literatura relacionam a atividade do complexo MLL com CM positivo para RE. Nossos resultados apontam que este complexo também pode ter um papel, ainda a ser investigado, no modelo TN da doença, uma vez que a expressão de seus membros neste modelo se assemelha à observada no modelo RE positivo, havendo também alta localização do efetor MLL1 no núcleo.



Enquanto os membros do complexo MLL avaliados apresentaram, em geral, pouca variação de expressão entre os modelos, *PRMT6*, que promove uma marca que antagoniza a metilação de H3K4 promovida pelo complexo MLL (H3R2me) (SCHUETTENGRUBER *et al.*, 2011), foi encontrado aumentado em MCF-7. Segundo Sun *et al.* (2014), *PRMT6* está superexpresso em diversos tipos de câncer, incluindo o CM, onde age como um *enhancer* da atividade de RE, atuando como um oncogene, sugerindo um papel específico em modelo RE+. Nossos dados corroboram com essa premissa, e sugerem que *PRMT6* possa atuar como um antagonista da atividade de MLL.

Ainda no contexto de ativação da expressão gênica, os resultados obtidos de *SMARCA1*, a ATPase do complexo de remodelamento de nucleossomos NURF, corroboram com os de Ye *et al.* (2009), que observaram uma maior expressão deste gene na linhagem MDA-MB-231, em comparação com MCF-7. Estes autores também observaram que, apesar de *SMARCA1* estar expresso tanto em tecidos normais quanto em tecidos tumorais, seu silenciamento levou à inibição do crescimento, dano no DNA e apoptose apenas nas células tumorais, indicando que estas células podem ser mais dependentes da atividade deste gene quando comparadas com células saudáveis.

Em outros modelos tumorais, *SMARCA1* parece ter função anti-tumoral, como por exemplo em câncer cervical, no qual Eckey *et al.* (2012) observaram que o silenciamento de *SMARCA1* em células HeLa levou a um aumento na proliferação e migração celulares, associado com uma ativação da via de Wnt e seus genes alvo envolvidos em migração e proliferação. Entretanto, o achado da localização de *SMARCA1* no núcleo de nossos modelos indica que esta proteína pode ser ativa em CM, necessitando de mais estudos acerca de sua função.

O resultado de expressão gênica sugere que *RBBP4* e *RBBP7* possam atuar como membros de NURF, uma vez que as quantificações de seus transcritos estão mais próximas de BPTF, membro de NURF, do que de membros de PRC2. *RBBP4* e *RBBP7* foram descritos como capazes de interagir com BCL11A, um fator de transcrição superexpresso em tumores TN, resultando no recrutamento de complexos de modificação epigenética, tais como PRC2 e NuRD, para seus alvos na cromatina, podendo contribuir para o CM (MOODY *et al.*, 2017). Em células MCF-7, *RBBP4* e *RBBP7* foram encontradas no núcleo, interagindo diretamente com RE e contribuindo para a regulação positiva ou negativa de diversos genes alvo deste receptor (CREEKMORE *et al.*, 2008). Uma possível atuação de *RBBP4/7* como NURF em CM ainda precisa ser elucidada.

Dentre nossos modelos, observamos pouca diferença de expressão gênica dos membros SWI/SNF avaliados. Corroborando com nossa análise, DeCristofaro *et al.* (2001), observaram que, em nível de proteína, a expressão de SMARCA2 e SMARCA4 não difere dentre as linhagens MCF-7 e MDA-MB-231.

Diversos estudos apontam potenciais papéis da família SWI/SNF no câncer. Aproximadamente 1 em cada 5 tumores pode apresentar mutações nos genes que codificam os membros de SWI/SNF e/ou proteínas acessórias do complexo (KADOCH *et al.*, 2013; SHAIN e POLLACK, 2013). Em tumores de mama, por exemplo, uma alta expressão de SMARCA2 e SMARCA4 é observada em comparação com tecidos adjacentes normais, e está relacionada com baixa sobrevivência (WU *et al.*, 2015). SMARCA4 também é capaz de promover proliferação de células TN, inclusive MDA-MB-231, através de diversos mecanismos (WU *et al.*, 2015), e pode ser recrutada por RE em MCF-7 para a coativação de seus genes alvo (DIRENZO *et al.*, 2000). Estes dados indicam que SWI/SNF pode estar envolvido na regulação de ambos os subtipos de CM, e estudos futuros podem permitir elucidar ainda mais seu papel nesta doença.

O gene *ASXL1*, assim como *ASXL2*, apresentou maior expressão na linhagem TN. Em níveis comparativos, porém, a diferença de expressão foi mais significativa para *ASXL2*. Entretanto, não se deve descartar um potencial papel de *ASXL1* no CM, uma vez que mutações neste gene já foram associadas à doença em tumores RE positivos (KATOH, 2015).

O estudo de colocalização de *ASXL2* com membros de Pc (YY1) e Trx (MLL1) permitiu verificar que YY1 possui baixa colocalização com *ASXL2*. Entretanto, a colocalização de *ASXL2* com MLL1 nos chamou atenção, pois Park *et al.* (2016) já haviam identificado a associação de *ASXL2* com MLL2, proteína que exerce a mesma função de MLL1, na linhagem MCF-7 superestimulada com o ligante de RE. A interação entre *ASXL2* e MLL1 não havia sido descrita anteriormente, e uma vez que a linhagem MCF-7 não foi estimulada em nosso estudo, pode-se sugerir que a associação entre *ASXL2* e MLL1 observada neste modelo possa ser ainda maior em um estado de ativação da célula.

Os resultados de expressão e localização nos mostraram que há diferença na distribuição de *ASXL2* e membros Pc e Trx entre as linhagens celulares avaliadas, bem como de membros associados e remodeladores de cromatina. Esta diferença, porém, não aponta para a predominância de uma marca de modificação específica em um modelo específico, indicando que ambos os grupos possam ter papel nas linhagens celulares, podendo atuar em conjunto com *ASXL2* diferentemente nas mesmas. A ação dos modificadores de histona e remodeladores

avaliados em conjunto com ASXL2 pode, assim, depender da localização de cada grupo na cromatina.

Para melhor compreender e avaliar funções e vias de sinalização relacionadas direta e indiretamente com ASXL2 em CM Luminal A e TN, realizamos seu silenciamento gênico nos modelos representativos destes subtipos, MCF-7 e MDA-MB-231, seguido de análise da expressão gênica global por microarranjo. Para a escolha do ponto de corte, optamos por mantê-lo igual ou acima de 2, a fim de obtermos o maior número possível de genes DE.

Em uma análise inicial, observamos que os efeitos do silenciamento gênico de *ASXL2* foram mais pronunciados na linhagem TN, uma vez que uma maior quantidade de genes DE foram encontrados nesta linhagem, em comparação com o modelo Luminal A. Além disso, notamos que os genes regulados por ASXL2 diferiam entre os modelos, havendo pouca sobreposição entre eles. Apesar disso, em análises posteriores, identificamos que estes genes estavam, em sua maioria, relacionados com as mesmas vias e processos celulares, sugerindo que ASXL2 é capaz de regular as mesmas vias em ambos os modelos, agindo porém com efeitos diferenciados nas mesmas.

Uma análise dos fatores de transcrição relevantes nas linhagens silenciadas para *ASXL2* foi realizada com o Metacore®. Tal análise leva em conta o panorama das alterações de expressão gênica observadas após o silenciamento de *ASXL2*, e gera uma lista com os fatores de transcrição mais relevantes que possam estar regulando os genes DE identificados nos modelos avaliados. Conforme esperado, o fator de transcrição ESR1 (receptor de estrógeno nuclear) foi apontado como relevante na linhagem MCF-7 silenciada, modelo celular que representa o subtipo Luminal A, sugerindo que ASXL2 pode ter um papel específico neste subtipo, uma vez que seu silenciamento é capaz de alterar a expressão de genes regulados por RE. Em MDA-MB-231, por sua vez, os fatores de transcrição TCF8 e GATA-2, envolvidos respectivamente em indução de transição epitélio-mesenquimal em tumores e manutenção da proliferação em células-tronco (RODRIGUEZ-BRAVO *et al.*, 2017; ZHANG *et al.*, 2019), foram identificados dentre os 10 mais relevantes, indicando uma possível contribuição de ASXL2 na regulação do fenótipo mais agressivo, característico deste modelo.

Fatores de transcrição relacionados com pluripotência e vias de células-tronco, por sua vez, foram encontrados como relevantes em ambas as linhagens silenciadas, com representatividades diferentes. Oct3/4, SOX2 e NANOG, marcadores clássicos de células-tronco embrionárias, estão associados, dentre outros, com proliferação, transição epitélio-mesenquimal e

agressividade (GWAK *et al.*, 2017). Os dois últimos foram encontrados como DE após o silenciamento de *ASXL2*, indicando que podem agir como (1) possíveis reguladores dos processos e vias associados a *ASXL2*, mas também como (2) potenciais alvos, diretos ou indiretos, de *ASXL2*.

Oct3/4 é um importante marcador de prognóstico em CM, e está associado com os subtipos mais agressivos desta doença, corroborando com sua maior relevância em nosso modelo TN silenciado (GWAK *et al.*, 2017). O gene *SOX2*, apontado como mais relevante no modelo Luminal A silenciado, foi encontrado diminuído nesta linhagem. Como outros membros da família SOX, *SOX2* pode atuar como supressor de tumor através da regulação do ciclo celular e apoptose, inibindo o crescimento celular, ou como oncogene, promovendo proliferação, progressão tumoral e migração (XU e YANG, 2017). Em CM, assim como Oct3/4, *SOX2* também está associado com os subtipos mais agressivos (GWAK *et al.*, 2017). *NANOG*, apesar de apresentar a mesma representatividade em ambos os modelos, parece ter mais relevância em MDA-MB-231, uma vez que foi encontrado aumentado nesta linhagem após o silenciamento de *ASXL2*. Este gene está superexpresso em diversos tipos de câncer, incluindo o CM, onde pode contribuir para a proliferação e a manutenção de um estado não-diferenciado de células-tronco tumorais (SEYMOUR, TWIGGER e KAKULAS, 2015), corroborando com o modelo TN.

Além de *SOX2*, outros genes da família SOX foram encontrados DE após o silenciamento de *ASXL2*. De forma geral, os membros da família SOX possuem importantes funções no desenvolvimento embrionário e diferenciação celular, podendo ativar e inibir a transcrição de genes (DONG, WILHELM e KOOPMAN, 2004; CASTILLO e SANCHEZ-CESPEDES, 2012). Assim como *SOX2*, o gene *SOX9*, diminuído em MCF-7 silenciada, possui papel contexto-dependente (XU e YANG, 2017). Em CM, *SOX9* está associado com um pior prognóstico (RICHTIG *et al.*, 2017). O gene *SOX21*, diminuído em MDA-MB-231 após o silenciamento de *ASXL2*, já foi descrito como superexpresso em glioblastoma e câncer de pulmão, entretanto seu papel nestas doenças não é conhecido (DONG, WILHELM e KOOPMAN, 2004).

Também envolvida em vias de células-tronco, a família de fatores de transcrição HOX é essencial para o desenvolvimento embrionário, mas seus componentes podem ter papéis importantes no câncer, onde são capazes de agir como ativadores ou repressores de transcrição (BHATLEKAR, FIELDS e BOMAN, 2014). Em alguns tecidos, genes *HOX* com função de supressores de tumor podem estar silenciados, enquanto outros genes da mesma família podem

estar mais expressos, com efeitos oncogênicos. Dentre os mecanismos envolvidos na regulação do câncer por genes *HOX*, estão a transição epitélio-mesenquimal e a invasão, a ativação de vias anti-apoptóticas, a indução de proliferação, e a diminuição de genes *HOX* que promovem diferenciação. Em CM, o gene *HOXA5* foi descrito como um regulador direto de p53, e *HOXB13* pode estar relacionado com migração, invasão e resistência à terapia com tamoxifeno em CM positivo para RE (SHAH e SUKUMAR, 2010).

Em MCF-7 silenciada, encontramos expressão diminuída de *HOXA1*, *HOXC11* e *HOXD8*, já descritos como diminuídos em CM quando comparado com tecido normal (BHATLEKAR, FIELDS e BOMAN, 2014). Em MDA-MB-231 silenciada, os genes *HOX* diminuídos foram *HOXA2*, que também foi descrito como diminuído em tumores de mama (BHATLEKAR, FIELDS e BOMAN, 2014), e *HOXB1*, capaz de prevenir a diferenciação de células-tronco neuronais (SHAH e SUKUMAR, 2010). Uma vez que estes genes estão normalmente diminuídos em CM, o silenciamento de *ASXL2* poderia levar a uma diminuição ainda maior em sua expressão, possivelmente favorecendo o fenótipo agressivo e sugerindo um papel de *ASXL2* na regulação de genes *HOX*. Nossos dados, porém, não permitem elucidar se o mecanismo desta regulação envolve ação direta de *ASXL2*, ou indireta através de potencial papel ETP, uma vez que genes *HOX* são os alvos clássicos de modificações epigenéticas por membros Pc e Trx.

Além dos fatores de transcrição discutidos, análises de enriquecimento permitiram identificar que vias de sinalização de células-tronco, tais como Notch, Hedgehog e Wnt, também estão alteradas após o silenciamento, demonstrando o envolvimento de *ASXL2* nas mesmas.

A via de Notch, por exemplo, pode operar em diversos contextos, e assim como outras vias de sinalização relacionadas com a regulação de células-tronco, pode contribuir para o câncer ou ter função de supressão tumoral (BRAY, 2016). Nos modelos silenciados, o gene *HEY2*, que atua como repressor da transcrição, foi identificado como diminuído em MCF-7, enquanto os genes *DLL1*, ligante da via, e *RBPJ*, que contribui para a transcrição de genes alvo quando a via se encontra ativada, estão aumentados em MDA-MB-231. Em CM, a expressão alterada de membros da via de Notch é reportada, e contribui para a iniciação, progressão e manutenção do tumor (KONTOMANOLIS *et al.*, 2018).

Uma das principais sinalizações em desenvolvimento embrionário e manutenção de células-tronco é a sinalização de Wnt, que pode ser encontrada como aberrante em diversos tipos

de câncer, onde contribui para manutenção de células-tronco tumorais e processos como proliferação e metástase. A sinalização de Wnt é realizada através da via canônica, dependente de  $\beta$ -catenina, e das vias não-canônicas de polaridade celular planar e mediada por  $\text{Ca}^{2+}$  (POHL *et al.*, 2017; ZHAN, RINDTORFF e BOUTROS, 2017).

Em ambos os modelos silenciados para *ASXL2*, encontramos um enriquecimento da sinalização de Wnt, porém com um papel aparentemente diferente entre as linhagens. Em MCF-7, observamos uma desregulação na via canônica, com o aumento na expressão de dois de seus inibidores (*DKK1* e *KREMEN1*) e pouca alteração ou diminuição na expressão dos demais membros da via, dentre eles o ligante *WNT2B* e o fator de transcrição *LEF1*. Em MDA-MB-231, diversos membros da sinalização de Wnt foram encontrados como aumentados, com destaque para componentes de vias não-canônicas.

A sinalização de Wnt tem um papel conhecido no desenvolvimento e progressão de CM do subtipo TN, havendo desregulação das vias canônica e não-canônicas. Seu enriquecimento, que pode ocorrer em até 50% das pacientes de CM, está associado com menor sobrevivência e pior prognóstico. (POHL *et al.*, 2017; ZAHN, RINDTORFF e BOUTROS, 2017).

Em um comparativo entre as células *scramble* de ambas as linhagens, notamos que MDA-MB-231 apresenta, em geral, expressão aumentada de ligantes e receptores da sinalização de Wnt, corroborando com dados da literatura que indicam uma ativação das vias de Wnt no subtipo TN (POHL *et al.*, 2017). Enquanto o ligante *WNT3* e o receptor *FZD10*, da via canônica, foram encontrados diminuídos após o silenciamento, outros membros desta via tiveram expressão ainda mais aumentada, como *TCF7*, componente efetor da via, e *PORCN*, que atua no processamento de ligantes, indicando que estes poderiam estar mais disponíveis para ativação da via. Além disso, observamos um aumento na expressão de *WNT5A*, o ligante mais proeminente das vias não-canônicas (ZAHN, RINDTORFF e BOUTROS, 2017) e *FZD7*, receptor que pode atuar em ambas as vias, sugerindo que a via não-canônica possa estar ainda mais aumentada em decorrência do silenciamento.

Assim, nossos resultados sugerem *ASXL2* como um potencial regulador da via de Wnt em ambos os modelos avaliados. Em MCF-7 silenciada, a via canônica de Wnt pode estar negativamente regulada. Em MDA-MB-231, por outro lado, as vias canônica e não-canônica de Wnt, normalmente ativadas no subtipo TN, parecem estar mais acentuadas após o silenciamento de *ASXL2*, com destaque principalmente para as vias não-canônicas, podendo contribuir para a exacerbação do fenótipo agressivo desta linhagem.

Outra via de sinalização alterada após o silenciamento é a de TGF-beta, enriquecida em ambas as linhagens. Esta via apresenta funções opostas durante o desenvolvimento e a progressão do câncer. Em estágios iniciais da tumorigênese, a regulação de TGF-beta tem um papel de supressão tumoral, através da inibição da proliferação e da indução de apoptose. Entretanto, em estágios mais avançados da doença, a via contribui para processos como inibição da resposta imune, transição epitélio-mesenquimal, migração e metástase (KATZ *et al.*, 2013; XIE *et al.*, 2018).

Ratificamos que a linhagem MDA-MB-231 não-silenciada apresenta, normalmente, altos níveis de expressão dos componentes da via avaliados, de acordo com dados da literatura que relacionam aumento na via de TGF-beta com fenótipos mais agressivos (BARCELLOS-HOFF e ARKHURST, 2009). Após o silenciamento de *ASXL2*, a expressão da via é exacerbada em ambos os modelos. Em MCF-7, este aumento na expressão dos membros da via reduz, ainda que timidamente, a diferença de expressão que normalmente é observada em relação à linhagem MDA-MB-231.

Dentre os membros aumentados, destacamos os ligantes *TGFB1*, *TGFB2* e *BMP3*, o receptor *TGFBRI*, e os genes *SMAD*, mediadores da sinalização da via. Quando ativadas pelos receptores da via, as proteínas SMAD2 e SMAD3 formam um complexo com SMAD4 e são translocadas para o núcleo da célula, onde são capazes de regular genes alvo, dentre eles *SNAIL*, *SLUG* e *TWIST*, envolvidos na transição epitélio-mesenquimal, e metaloproteinases de matriz (*MMPs*) (AHMADI *et al.*, 2018).

Além desta via, nossas análises, em conjunto, apontaram para alterações em outros genes relacionados com os processos de adesão, migração e transição epitélio-mesenquimal, como as integrinas, que promovem adesão à matriz extracelular e podem estar com expressão reduzida em tumores, incluindo o CM (KOISTINEN e HEINO, 2013); as claudinas, que também possuem um papel na adesão celular, sendo componentes importantes em junções oclusivas, e cuja diminuição na expressão pode levar à motilidade e invasão em tumores metastáticos (TABARIES e SIEGEL, 2017); e as metaloproteinases de matriz (*MMPs*), que através de seu papel na quebra da matriz extracelular podem contribuir para a migração, invasão e metástase, sendo encontradas superexpressas em diversos tipos de tumores (NABESHIMA *et al.*, 2002). Em nossos modelos silenciados, as *MMPs* foram encontradas em sua maioria aumentadas, enquanto as integrinas e claudinas estão em sua maioria diminuídas.

O fator de transcrição *RUNX2* também foi encontrado alterado em nossos modelos. Este gene é encontrado frequentemente expresso em células tumorais metastáticas, incluindo células de CM, e pode contribuir para a metástase de diversas maneiras, dentre elas na promoção de autofagia como uma maneira de reciclar nutrientes, contribuindo para que células metastáticas sobrevivam ao stress nutricional associado com a metástase. *RUNX2* também atua regulando genes envolvidos com migração e angiogênese, que podem favorecer o fenótipo metastático, e em CM, sua alta expressão pode estimular a migração em células MDA-MB-231 e MCF-7 (LEONG *et al.*, 2010; FERRARI *et al.*, 2013; TANDON *et al.*, 2017; VISHAL *et al.*, 2017).

Nos nossos modelos, em condições normais das células (*scramble*), observamos que MDA-MB-231 expressa níveis elevados do gene *RUNX2*, corroborando com dados de Leong et al. (2010). Após o silenciamento de *ASXL2*, os níveis de expressão do gene *RUNX2* aumentaram ainda mais neste modelo, em cerca de 5 vezes, conforme validamos por RT-qPCR. Em MCF-7, Leong et al. (2010) reportam que não há expressão proteica de *RUNX2*. Em nossos resultados, porém, observamos que, após o silenciamento, os níveis de transcrito de *RUNX2* aumentam em mais de 5 vezes neste modelo. Assim, podemos inferir que ambos os modelos são sensíveis ao silenciamento de *ASXL2* no que diz respeito à expressão do gene *RUNX2*. Em MDA-MB-231, o aumento da expressão de *RUNX2* após o silenciamento poderia estar associado com uma maior agressividade, uma vez que este gene tem importantes funções em migração e metástase.

É sabido que a linhagem MDA-MB-231 possui alto potencial migratório e metastático (ANTALIS *et al.*, 2011; BAILEY *et al.*, 2012) quando comparada com a linhagem MCF-7, de baixa motilidade (PROIETTI *et al.*, 2018). Em MCF-7, houve um pequeno aumento na migração após o silenciamento de *ASXL2*, de cerca de 3% ( $p=0,0672$ ), que pode estar relacionado com o enriquecimento da sinalização de TGF-beta, com o aumento na expressão de *MMPs* (*MMP9/13*) e *RUNX2*, e com a diminuição na expressão das integrinas *ITGB3* e *ITGB8*. Este aumento na migração, por ser baixo e não estatisticamente significativo, evidencia que o enriquecimento de uma via de sinalização não é capaz de aproximar, por si só, a capacidade migratória de MCF-7, um modelo epitelial, àquela observada em MDA-MB-231, um modelo mesenquimal.

Em MDA-MB-231, por sua vez, o silenciamento de *ASXL2* permitiu aumentar ainda mais o potencial migratório, em cerca de 20% ( $p<0,05$ ), também condizente com o enriquecimento de TGF-beta, o aumento na expressão gênica de *MMPs* (*MMP8/9/12*) e *RUNX2*, bem como a diminuição da integrina *ITGB2* e das claudinas *CLDN10*, *11* e *12*. O gene *MMP21*, associado com progressão de tumor e pior prognóstico em câncer gástrico (WU *et al.*, 2013; CHANG *et al.*,



2014; ZHANG *et al.*, 2018), foi encontrado com a expressão extremamente diminuída (*fold change* de -104) em MDA-MB-231 após o silenciamento. Uma diminuição na expressão de *MMP2* (*fold change* de -2) também foi observada após o silenciamento. Estas diminuições, porém, parecem não interferir na capacidade migratória das células, indicando que, neste modelo, o aumento na expressão dos demais genes identificados relacionados com migração, incluindo outras MMPs, podem, em conjunto, sobrepujar a diminuição observada em *MMP21* e *MMP2*.

Ainda, a avaliação de ciclo celular não mostrou enriquecimento das células nas fases G2-M, indicando que estas não estão com maior capacidade proliferativa após o silenciamento. Isto ratifica que os resultados obtidos no experimento de *wound healing* são, de fato, reflexo da capacidade migratória das células. Além disso, no modelo TN, o aumento de expressão de interleucinas relacionadas com vias de TGF-beta, NFκB e MAPK, dentre outras, envolvidas em inflamação e agressividade (WANG *et al.*, 2015; ALINEJAD *et al.*, 2017; BIE *et al.*, 2017), pode também estar relacionado com o fenótipo mais agressivo observado nestas células após o silenciamento.

No contexto de processos biológicos relacionados com fenótipo mais agressivo, está o transporte de moléculas por microvesículas. Este consiste em um importante meio de troca de informações entre células, e no câncer, pode promover a transferência de DNA, RNA, receptores, MMPs, dentre outros, podendo induzir a ativação de vias de sinalização pró-tumorais e a inibição de vias anti-tumorais (WU *et al.*, 2017). A contagem de microvesículas, diminuída no modelo TN silenciado, corrobora com a diminuição na expressão de genes envolvidos na formação das mesmas, por exemplo os genes *SYT*; porém não se deve descartar a hipótese de que o conteúdo das vesículas não esteja interferindo no fenótipo mais agressivo observado nessas células em comparação com o controle. Estudos futuros poderão permitir um melhor entendimento da relação de *ASXL2* com a formação de microvesículas e seu impacto na agressividade de células de CM.

Com o silenciamento de *ASXL2*, notamos que há um impacto na expressão de genes relacionados com funções importantes para a tumorigênese. Uma vez que *ASXL2* é uma proteína pouco estudada em termos funcionais e de regulação da expressão gênica, foi aplicada uma abordagem proteômica para a identificação de seus potenciais parceiros nos modelos celulares avaliados, e assim contribuir para aprofundar nosso conhecimento em relação a seu papel na regulação de vias importantes para o CM.

Em nossos modelos celulares, notamos que ASXL2 é capaz de se associar, direta ou indiretamente, a diversas proteínas com as mais variadas funções, dentre elas componentes do citoesqueleto, enzimas, transportadores, mas também a proteínas de ligação a DNA e RNA capazes de regular a expressão gênica em diversos níveis, levantando novamente o envolvimento de ASXL2 neste controle. Este é um achado importante, pois, assim como observamos no experimento de microarranjo, ASXL2 pode ser capaz de regular direta ou indiretamente genes envolvidos com o controle da expressão gênica alterados no câncer, tais como *DNMT1*, *DNMT3L* e *TDG*, que possuem papel na metilação de DNA; e *DICER1*, que atua no processamento de miRNAs (KHOSHNAW *et al.*, 2012; HE *et al.*, 2014; YANG *et al.*, 2015; ZHANG e XU, 2017).

Uma vez que buscamos compreender o papel de ASXL2 na regulação da expressão gênica, mais especificamente no mecanismo epigenético de remodelamento de cromatina, focamos nosso estudo em modificadores de histona e remodeladores de cromatina que foram encontrados associados à nossa proteína de interesse. Na análise proteômica, identificamos membros envolvidos tanto em ativação quanto repressão da expressão gênica, não havendo sobreposição entre regulação positiva e negativa nos nossos modelos. Além disso, nossos resultados apontam que, nos modelos de CM avaliados, ASXL2 pode interagir com modificadores de histona e remodeladores de cromatina não-clássicos. Ainda assim, alguns dos parceiros identificados são componentes clássicos das famílias Pc e Trx, como é o caso de CBX8, MLL1 e BPTF.

No modelo Luminal A, um dos parceiros de ASXL2 é CBX8, membro do complexo PRC1. Observamos que, em termos de expressão gênica, *CBX8* está expresso de forma mais abundante neste modelo quando comparado com MDA-MB-231. Este resultado sugere um potencial papel de CBX8 em MCF-7, porém, fica em aberto o efeito desta associação na contribuição para o fenótipo Luminal A. Neste mesmo modelo, ASXL2 também pode se associar com MLL1. Anteriormente, observamos esta mesma associação através do experimento de imunofluorescência, dando robustez ao nosso resultado e indicando que ASXL2 pode formar complexo com MLL1 em MCF-7, possivelmente regulando sua atividade. Em MCF-7, ASXL2 também é capaz de se associar a *KDM5A*, uma demetilase de função antagônica ao complexo MLL. Em CM, o gene *KDM5A* pode estar amplificado ou superexpresso, e está associado com maior proliferação (HOU *et al.*, 2012). Este resultado indica que, na linhagem Luminal A, ASXL2 pode estar implicado na regulação do *status* de metilação em H3K4.

A proteína KAT6B (MORF), encontrada como parceira de ASXL2 no modelo Luminal A, está envolvida em ativação da expressão gênica, pois é capaz de promover acetilação em H3 (ULLAH *et al.*, 2008). KAT6B pode atuar estimulando a transcrição de genes alvo de RUNX2, através de interação com este fator de transcrição, enquanto RUNX2 é capaz de regular negativamente o potencial de ativação transcricional de KAT6B (PELLETIER *et al.*, 2002). A proteína KAT6A (MOZ), também encontrada como parceira de ASXL2 em MCF-7, possui a mesma função de KAT6B, e a mesma capacidade de interação com RUNX2 para ativação de genes alvo (PELLETIER *et al.*, 2002). Porém, o potencial de inibição da atividade HAT por RUNX2 não foi descrito para esta proteína, apenas para KAT6B. Com base nestes dados, a partir da observação do aumento na expressão gênica de *RUNX2* após o silenciamento, e sabendo que ASXL2 é capaz de interagir com a marca de silenciamento H3K27me3 na linhagem MCF-7, conforme observamos no experimento de CHIP, propomos que em células MCF-7 não-silenciadas, ASXL2 interage com H3K27me3 na cromatina, direta ou indiretamente, e com KAT6B, dentre outras proteínas. KAT6B, por sua vez, poderia promover acetilação nas histonas e, por consequência, auxiliar na abertura da cromatina, culminando na ativação de genes alvo. O silenciamento de ASXL2 neste modelo aumenta a expressão gênica de *RUNX2*. RUNX2 poderia então inibir KAT6B e conseqüentemente seu papel na abertura da cromatina, e assim, a estrutura da cromatina permaneceria fechada, e os genes regulados por ASXL2, silenciados.

Em MDA-MB-231, fica em aberto esta regulação via ASXL2 e RUNX2, uma vez que 1) não foi possível identificar interação entre ASXL2 e KAT6B por abordagem proteômica; 2) esta linhagem já possui altos níveis de RUNX2 e ASXL2, podendo estes competir pela possível ligação com KAT6B; 3) o experimento de CHIP, que mostraria um possível envolvimento de ASXL2 com a marca de silenciamento e potencial recrutamento de complexos ativadores, não foi realizado nesta linhagem.

Todavia, em MDA-MB-231, a proteína BPTF (complexo NURF) foi identificada como potencial parceiro de ASXL2, e apesar de, em termos de expressão gênica, estar menos presente nesta linhagem, pode estar associada a ASXL2. Ainda, RSBN1L, também identificada como parceira de ASXL2 no modelo TN, é paróloga de RSBN1, uma HDM que remove a marca de dimetilação em H4K20. Ambas as proteínas ainda são pouco estudadas, mas sabe-se que a alta expressão de RSBN1 em CM do tipo Luminal está associada com melhor prognóstico (ENCINAS *et al.*, 2018). Em nossos modelos, RSBN1L também possui expressão gênica aumentada na linhagem Luminal A em comparação com MCF-7. Sua interação com ASXL2 no

modelo TN ainda não havia sido descrita, sendo portanto necessários maiores estudos para sua compreensão.

Componentes do complexo HBO1 (JADE1/3), que promove acetilação em H4, uma marca de ativação da expressão gênica, também foram identificados como parceiros de ASXL2 em nossos modelos. A proteína JADE1, a mais estudada, tem o papel de direcionar o complexo HBO1 para seus alvos na cromatina através de seu domínio PHD. Alguns estudos sugerem ainda um papel de JADE1 em regular negativamente a via de Wnt, através da ubiquitinação de  $\beta$ -catenina. Em câncer renal, JADE1 apresenta baixos níveis de expressão, e sugere-se um papel de supressão tumoral para esta proteína (CHITALIA *et al.*, 2008; SAKSOUK *et al.*, 2009; PANCHENKO, 2016; XIAO-FEN *et al.*, 2016; HAN *et al.*, 2018). Em nossos modelos, JADE1 é parceira de ASXL2 em ambos, mas está mais expressa em MDA-MB-231, podendo indicar que ela tenha potencial função oncogênica e não como reguladora negativa de Wnt neste modelo, uma vez que esta via é descrita como enriquecida no subtipo TN.

O papel e funções de JADE3, associada a ASXL2 na linhagem, MCF-7, ainda são pouco conhecidos, mas sabe-se que esta proteína também possui o domínio PHD. No câncer, JADE3 foi encontrada superexpressa em câncer colorretal, e pode ter sua expressão induzida pela sinalização de Wnt (JIAN *et al.*, 2018), o que corrobora com nossos dados de expressão.

Como proposto em 2001 por Brock e Van Lohuizen, as proteínas ETP podem atuar de diversas maneiras, dentre elas como subunidades de complexos de remodelamento da cromatina, ou como co-ativadores ou co-repressores de modificadores de histona. Nossos resultados permitem observar associações ainda não descritas entre ASXL2 e proteínas modificadoras de histona e remodeladoras da cromatina, podendo sugerir um papel ETP. Entretanto, nossos dados não são capazes de elucidar o tipo de relação entre as proteínas e, conseqüentemente, seu efeito em genes alvo. Ainda assim, nossos achados são um passo importante para a compreensão da função de ASXL2 nos modelos Luminal A e TN de CM.

## 7. CONCLUSÕES

- ❖ As linhagens celulares MCF-7 (Luminal A) e MDA-MB-231 (TN) apresentam expressão e localização diferencial de ASXL2, inclusive no núcleo, e este pode estar implicado em CM;
- ❖ Membros das famílias de modificadores de histona e remodeladores de cromatina Polycomb, Trithorax, SWI/SNF e NURF estão diferencialmente expressos entre as linhagens, e o complexo PRC2 parece estar mais ativo na linhagem TN, enquanto ASXL2 parece estar mais associado a MLL na linhagem Luminal A;
- ❖ ASXL2 está envolvido, em ambas as linhagens celulares, na regulação de vias de sinalização tais como Wnt e TGF-beta, que estão relacionadas com marcas de agressividade no câncer, como pluripotência, proliferação e migração;
- ❖ ASXL2 é capaz de se interagir, direta ou indiretamente, com modificadores de histona e remodeladores de cromatina, dentre eles membros Pc, Trx, NURF e outras proteínas não-clássicas envolvidas neste mecanismo epigenético;
- ❖ ASXL2 associa-se, na cromatina, à marca de silenciamento característica de PRC2 (H3K27me3) no modelo Luminal A, e esta associação possivelmente influencia na regulação da expressão gênica.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

ABDEL-WAHAB, O.; DEY, A. The ASXL-BAP axis: new factors in myelopoiesis, cancer and epigenetics. **Leukemia**, v. 27, p. 10-15, 2013.

AHMADI, A. *et al.* Transforming growth factor- $\beta$  signaling: Tumorigenesis and targeting for cancer therapy. **Journal of Cellular Physiology**, 2018.

ALKHATIB, S. G.; LANDRY, J. W. The Nucleosome Remodeling Factor. **FEBS Letters**, v. 585, n. 20, p. 3197-3207, 2011.

ALINEJAD, V. *et al.* The role of IL17B-IL17RB signaling pathway in breast cancer. **Biomedicine and Pharmacotherapy**, v. 88, p. 795-803, 2017.

ALLIS, C. D. *et al.* Epigenetics. Nova Iorque: Cold Spring Harbor Laboratory Press, 2007.

ALORAIIFI, F. *et al.* Gene analysis techniques and susceptibility gene discovery in non-BRCA1/BRCA2 familial breast cancer. **Surgical Oncology**, v. 24, n. 2, p. 100-109, 2015.

ANTALIS, C. J. *et al.* Migration of MDA-MB-231 breast cancer cells depends on the availability of exogenous lipids and cholesterol esterification. **Clinical and Experimental Metastasis**, v. 28, n. 8, p. 733-741, 2011.

AUDIA, J. E.; CAMPBELL, R. M. Histone Modifications and Cancer. **Cold Spring Harbor Perspectives in Biology**, v. 8, n. 4, 2016.

BAILEY, C. K. High Motility of Triple-negative Breast Cancer Cells Is Due to Repression of Plakoglobin Gene by Metastasis Modulator Protein SLUG. **The Journal of Biological Chemistry**, v. 287, n. 23, p. 19472-19486, 2012.

BANNISTER, A. J.; KOUZARIDES, T. Regulation of chromatin by histone modifications. **Cell Research**, n. 21, p. 381-395, 2011.

BARCELLOS-HOFF, M. H.; ARKHURST, R. J. Transforming growth factor- $\beta$  in breast cancer: too much, too late. **Breast Cancer Research**, v. 11, n. 1, 2009.

BASKIND, H. A. *et al.* Functional Conservation of *Asxl2*, a Murine Homolog for the *Drosophila* Enhancer of Trithorax and Polycomb Group Gene *Asx*. **PLoS One**, v. 4, n. 3, 2009.

BASSE, C.; AROCK, M. The increasing roles of epigenetics in breast cancer: Implications for pathogenicity, biomarkers, prevention and treatment. **International Journal of Cancer**, v. 137, n. 12, p. 2785-2794, 2014.

BASU, A. *et al.* YY1 DNA binding and interaction with YAF2 is essential for Polycomb recruitment. **Nucleic Acids Research**, v. 42, n. 4, p. 2208-2223, 2014.

BERSE, B.; LYNCH, J. A. Molecular Diagnostic Testing in Breast Cancer. **Seminars in Oncology Nursing**, v. 31, n. 2, p. 108-121, 2015.

BHATLEKAR, S.; FIELDS, J. Z.; BOMAN, B. M. HOX genes and their role in the development of human cancers. **Journal of Molecular Medicine**, v. 92, n. 8, p. 811-823, 2014.

BIE, Q. *et al.* IL-17B: A new area of study in the IL-17 family. **Molecular Immunology**, v. 90, p. 50-56, 2017.

BLACKLEDGE, N. P.; ROSE, N. R.; KLOSE, R. J. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. **Nature Reviews Molecular Cell Biology**, v. 16, p. 643-649, 2015.

BOURDEANU, L.; LIU, E. A. Systemic treatment for breast cancer: chemotherapy and biotherapy agents. **Seminars in Oncology Nursing**, v. 31, n. 2, p. 156-162, 2015.

BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, v. 72, p. 248-254, 1976.

BRAY, S. J. Notch Signalling in Context. **Nature Reviews Molecular Cell Biology**, v. 17, n. 11, p. 722-735, 2016.

BROCK, H. W.; VAN LOHUIZEN, M. The Polycomb group - no longer an exclusive club?. **Current Opinion in Genetics and Development**, v. 11, n. 2, p. 175-181, 2001.

CAMPBELL, M. J.; TURNER, B. M. Altered Histone Modifications in Cancer. In: KARP, A (Ed.). Epigenetic Alterations in Oncogenesis, **Advances in Experimental Medicine and Biology**, v. 754, p. 81-108. Nova Iorque: Springer, 2013.

CARRICK, S. *et al.* Single agent versus combination chemotherapy for metastatic breast cancer. **The Cochrane Database of Systematic Reviews**, v. 15, n. 2, 2009.

CASTILLO, S. D.; SANCHEZ-CESPEDES, M. The SOX family of genes in cancer development: biological relevance and opportunities for therapy. **Expert Opinion on Therapeutic Targets**, v. 16, n. 9, p. 903-919, 2012.

CECH, T. R.; STEITZ, J. A. The noncoding RNA revolution – trashing old rules to forge new ones. **Cell**, v. 157, n. 1, p. 77-94, 2014.

CECILIO, A. P. *et al.* Breast Cancer in Brazil: epidemiology and treatment challenges. **Breast Cancer: Targets and Therapy**, v. 7, p. 43-49, 2015.

CHANG, W. J. *et al.* Inflammation-related factors predicting prognosis of gastric cancer. **World Journal of Gastroenterology**, v. 20, n. 16, p. 4586-4596, 2014.

CHIEN, Y. *et al.* EZH2 promotes migration and invasion of triple-negative breast cancer cells via regulating TIMP2-MMP-2/-9 pathway. **American Journal of Cancer Research**, v. 8, n. 3, p. 422-434, 2018.

CHITALIA, V. C. *et al.* Jade-1 inhibits Wnt signaling by ubiquitinating  $\beta$ -catenin and mediates Wnt pathway inhibition by pVHL. **Nature Cell Biology**, v. 10, n. 10, p. 1208-1216, 2008.

CHOI, Y. J. *et al.* Expression of Bmi-1 protein in tumor tissues is associated with favorable prognosis in breast cancer patients. **Breast Cancer Research and Treatment**, v. 113, n. 1, p. 83-93, 2009.

CORADINI, D.; ORIANA, S. The role of maintenance proteins in the preservation of epithelial cell identity during mammary gland remodeling and breast cancer initiation. **Chinese Journal of Cancer**, v. 33, n. 2, p. 51-67, 2014.



CREA, F. *et al.* Polycomb genes and cancer: Time for clinical application?. **Critical Reviews in Oncology/Hematology**, v. 83, p. 184-193, 2012.

CREEKMORE, A. *et al.* The Role of Retinoblastoma Associated Proteins 46 and 48 in Estrogen Receptor  $\alpha$  Mediated Gene Expression. **Molecular and Cellular Endocrinology**, v. 291, n. 1-2, p. 79-86, 2008.

D'OTO, A. *et al.* Histone demethylases and their roles in cancer epigenetics. **Journal of Medical Oncology and Therapeutics**, v. 1, n. 2, p. 34-40, 2016.

DAI, X. *et al.* WDR5 Expression Is Prognostic of Breast Cancer Outcome. **PLoS One**, v. 10, n. 9, 2015.

DAOU, S. *et al.* The BAP1/ASXL2 Histone H2A Deubiquitinase Complex Regulates Cell Proliferation and Is Disrupted in Cancer. **The Journal of Biological Chemistry**, v. 290, n. 48, p. 28643-28663, 2015.

DAWSON, M. A.; KOUZARIDES, T. Cancer Epigenetics: From Mechanism to Therapy. **Cell**, n. 150, p. 12-27, 2012.

DECRISTOFARO, M. F. *et al.* Characterization of SWI/SNF Protein Expression in Human Breast Cancer Cell Lines and Other Malignancies. **Journal of Cellular Physiology**, n. 186, p. 136-145, 2001.

DE LEENEER, K.; CLAES, K. Non Coding RNA Molecules as Potential Biomarkers in Breast Cancer. In: SCATENA, R. (Ed.). Advances in Cancer Biomarkers, **Advances in Experimental Medicine and Biology**, v. 867, p. 263-275. Nova Iorque: Springer, 2015.

DIRENZO, J. *et al.* BRG-1 Is Recruited to Estrogen-Responsive Promoters and Cooperates with Factors Involved in Histone Acetylation. **Molecular and Cellular Biology**, v. 20, n. 20, p. 7541-7549, 2000.

DONG, C.; WILHELM, D.; KOOPMAN, P. Sox genes and cancer. **Cytogenetic and Genomic Research**, v. 105, p. 442-447, 2004.

DOR, Y.; CEDAR, H. Principles of DNA methylation and their implications for biology and medicine. **Lancet**, v. 392, n. 10149, p. 777-786, 2018.

DUTTA, A. *et al.* Swi/Snf dynamics on stress-responsive genes is governed by competitive bromodomain interactions. **Genes and Development**, v. 28, p. 2314-2330, 2014.

ECKEY, M. *et al.* Nucleosome Remodeler SNF2L Suppresses Cell Proliferation and Migration and Attenuates Wnt Signaling. **Molecular and Cellular Biology**, v. 32, n. 13, p. 2359-2371, 2012.

EHRlich, M.; LACEY, M. DNA Hypomethylation and Hemimethylation in Cancer. In: KARPF, A (Ed.). Epigenetic Alterations in Oncogenesis, **Advances in Experimental Medicine and Biology**, v. 754, p. 31-56. Nova Iorque: Springer, 2013.

EISENBERG, A. L. A.; KOIFMAN, S. Câncer de Mama: Marcadores Tumorais (Revisão de Literatura). **Revista Brasileira de Cancerologia**, v. 47, n. 4, p. 377-388, 2001.

ENCINAS, G. *et al.* Somatic mutations in early onset luminal breast cancer. **Oncotarget**, v. 9, n. 32, p. 22460-22479, 2018.

ENGEL, C.; FISCHER, C. Breast Cancer Risks and Risk Prediction Models. **Breast Care**, v. 10, p. 7-12, 2015.

ESCRIVÁ-DE-ROMANÍ, S. *et al.* HER2-positive breast cancer: Current and new therapeutic strategies. **Breast**, v. 39, p. 80-88, 2018.

FERRARI, N. *et al.* RUNX2 in mammary gland development and breast cancer. **Journal of Cellular Physiology**, v. 228, n. 6, p. 1137-1142, 2013.

GANGARAJU, V. K.; BARTHOLOMEW, B. Mechanisms of ATP Dependent Chromatin Remodeling. **Mutation Research**, v. 618, n. 1-2, p. 3-17, 2007.

GELSI-BOYER, V. *et al.* Mutations in *ASXL1* are associated with poor prognosis across the spectrum of malignant myeloid diseases. **Journal of Hematology and Oncology**, v. 5, n. 12, 2012.

GILDEA, J. J.; LOPEZ, R. C.; SHEARN, A. A Screen for New Trithorax Group Genes Identified *little imaginal discs*, the *Drosophila melanogaster* Homologue of Human Retinoblastoma Binding Protein 2. **Genetics**, v. 156, p. 645-663, 2000.

GOLDHIRSCH, A. *et al.* Strategies for subtypes – dealing with the diversity of breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. **Annals of Oncology**, v. 22, p. 1736-1747, 2011.

GWAK, J. M. *et al.* Expression of embryonal stem cell transcription factors in breast cancer: Oct4 as an indicator for poor clinical outcome and tamoxifen resistance. **Oncotarget**, v. 8, n. 22, p. 36305–36318, 2017.

HAN, J. *et al.* The scaffolding protein JADE1 physically links the acetyltransferase subunit HBO1 with its histone H3-H4 substrate. **Journal of Biological Chemistry**, v. 293, n. 12, p. 4498-4509, 2018.

HANAHAN, D.; WEINBERG, R. A. Hallmarks of Cancer: The Next Generation. **Cell**, v. 144, p. 646-674, 2011.

HANAHAN, D.; WEINBERG, R. A. The Hallmarks of Cancer. **Cell**, v. 100, p. 57-70, 2000.

HE, H. *et al.* Prognostic significance of low DICER expression regulated by miR-130a in cervical cancer. **Cell Death and Disease**, v. 5, n. 5, 2014.

HO, L.; CRABTREE, G. R. Chromatin remodelling during development. **Nature**, v. 463, n. 7280, p. 474-484, 2010.

HOLM, K. *et al.* Global H3K27 trimethylation and EZH2 abundance in breast tumor subtypes. **Molecular Oncology**, v. 6, p. 494-506, 2012.

HOU, J. *et al.* Genomic amplification and a role in drug-resistance for the KDM5A histone demethylase in breast cancer. **American Journal of Translational Research**, v. 4, n. 3, -. 247-256, 2012.

HU, J. *et al.* TRPS1 Suppresses Breast Cancer Epithelial-mesenchymal Transition Program as a Negative Regulator of SUZ12. **Translational Oncology**, v.11, n. 2, p. 416-425, 2018.

HUANG, X. *et al.* Enhancers of Polycomb EPC1 and EPC2 sustain the oncogenic potential of MLL leukemia stem cells. **Leukemia**, v. 28, n. 5, p. 1081-1091, 2014.

INSTITUTO NACIONAL DO CÂNCER. Estimativa 2016: Incidência de Câncer no Brasil. Rio de Janeiro: INCA, 2015.

INSTITUTO NACIONAL DO CÂNCER. Estimativa 2018: Incidência de Câncer no Brasil. Rio de Janeiro: INCA, 2017.

INSTITUTO NACIONAL DO CÂNCER. Tipos de câncer: mama. Disponível em: <<http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/mama>>. Acesso em: 23 mai. 2015, 14h.

JENUWEIN, T; ALLIS, C. D. Translating the Histone Code. **Science**, v. 293, p. 1074-1080, 2001.

JEONG, K. W *et al.* Establishment of active chromatin structure at enhancer elements by mixed-lineage leukemia 1 to initiate estrogen-dependent gene expression. **Nucleic Acids Research**, v. 42, n. 4, p. 2245-2256, 2014.

JEONG, K. W. *et al.* Recognition of enhancer element–specific histone methylation by TIP60 in transcriptional activation. **Nature Structural and Molecular Biology**, v. 18, n. 12, p. 1358-1366, 2011.

JIAN, Y. *et al.* Jade family PHD finger 3 (JADE3) increases cancer stem cell-like properties and tumorigenicity in colon cancer. **Cancer Letters**, v. 428, p. 1-11, 2018.

JIANG, T. *et al.* Prognostic value of high EZH2 expression in patients with different types of cancer: a systematic review with meta-analysis. **Oncotarget**, v. 7, n. 4, p. 4584-4597, 2015.

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

KADOCH, C. *et al.* Proteomic and Bioinformatic Analysis of mSWI/SNF (BAF) Complexes Reveals Extensive Roles in Human Malignancy. **Nature Genetics**, v. 45, n. 6, p. 592-601, 2013.

KATOH, M. Functional and cancer genomics of ASXL family members. **British Journal of Cancer**, v. 109, p. 299-306, 2013.

KATOH, M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. **Expert Reviews Proteomics**, v. 12, n. 3, p. 317-328, 2015.

KATZ, L. H. *et al.* Targeting TGF- $\beta$  signaling in cancer. **Expert Opinion on Therapeutic Targets**, v. 17, n. 7, p. 763-740, 2013.

KAUFHOLD, S.; GARBÁN, H.; BONAVIDA, B. Yin Yang 1 is associated with cancer stem cell transcription factors (SOX2, OCT4, BMI1) and clinical implication. **Journal of Experimental and Clinical Cancer Research**, v. 35, n. 84, 2016.

KENEMANS, P.; VERSTRAETEN, R. A.; VERHEIJEN, R. H. M. Oncogenic pathways in hereditary and sporadic breast cancer. **Maturitas**, v. 61, n. 1-2, p. 141-150, 2008.

KHORDADMEHR, M. *et al.* Key microRNAs in the biology of breast cancer; emerging evidence in the last decade. **Journal of Cellular Physiology**, 2018.

KHOSHNAW, E. M. *et al.* Loss of Dicer expression is associated with breast cancer progression and recurrence. **Breast Cancer Research and Treatment**, v. 135, p. 403-413, 2012.

KIM, K. H.; ROBERTS, C. W. M. Targeting EZH2 in Cancer. **Nature Medicine**, v. 22, n. 2, p. 128-134, 2016.

KINGSTON, R. E.; TAMKUN, J. W. Transcriptional Regulation by Trithorax-Group Proteins. **Cold Spring Harbor Perspectives in Biology**, v. 6, n. 10, 2014.

KOISTINEN, P.; HEINO, J. Integrins in Cancer Cell Invasion. In: Madame Curie Bioscience Database. Austin, Texas: Landes Bioscience, 2000-2013.

KONTOMANOLIS, E. N. *et al.* The Notch Pathway in Breast Cancer Progression. **Scientific World Journal**, 2018.

LAM, S. W. *et al.* Breast cancer classification by proteomic technologies: Current state of knowledge. **Cancer Treatment Reviews**, v. 40, p. 129-138, 2014.

LEE, E. Y. H. P.; MULLER, W. J. Oncogenes and Tumor Suppressor Genes. **Cold Spring Harbor Perspectives in Biology**, v. 2, n. 10, 2010.

LEONG, D. T. *et al.* Cancer-related ectopic expression of the bone-related transcription factor RUNX2 in non-osseous metastatic tumor cells is linked to cell proliferation and motility. **Breast Cancer Research**, v. 12, n. 5, 2010.

LESTER, J. Local Treatment of Breast Cancer. **Seminars in Oncology Nursing**, v. 31, n. 2, p. 122-133, 2015.

LI, J. *et al.* Loss of *Asxl2* leads to myeloid malignancies in mice. **Nature Communications**, v. 8, n. 15456, 2017.

LI, T. *et al.* *Additional sex combs* interacts with enhancer of zeste and trithorax and modulates levels of trimethylation on histone H3K4 and H3K27 during transcription of *hsp70*. **Epigenetics and Chromatin**, v. 10, n. 43, 2017.

LIN, C.; HE, L. Noncoding RNAs in Cancer Development. **Annual Review of Cancer Biology**, v. 1, p.163-184, 2017.

LIU, H. *et al.* EZH2-mediated Puma gene repression regulates non-small cell lung cancer cell proliferation and cisplatin-induced apoptosis. **Oncotarget**, v. 7, n. 35, p. 56338-56354, 2016.

LIU, X. *et al.* The EZH2-H3K27me3-DNMT1 complex orchestrates epigenetic silencing of the *wwc1* gene, a Hippo/YAP pathway upstream effector, in breast cancer epithelial cells. **Cellular Signalling**, n. 51, p. 243-256, 2018.

MAERTENS, G. N. *et al.* Several Distinct Polycomb Complexes Regulate and Co-Localize on the *INK4a* Tumor Suppressor Locus. **PLoS One**, v. 4, n. 7, 2009.

MAKKI, J. Diversity of Breast Carcinoma: Histological Subtypes and Clinical Relevance. **Clinical Medicine Insights: Pathology**, v. 8, p. 23-31, 2015.

MASOUD, V.; PAGES, G. Targeted therapies in breast cancer: New challenges to fight against resistance. **World Journal of Clinical Oncology**, v. 8, n. 2, p. 120-134, 2017.

MEISEL, J. L. *et al.* Evolution of Targeted Therapy in Breast Cancer: Where Precision Medicine Began. **American Society of Clinical Oncology Educational Book**, v. 38, p. 78-86, 2018.

MICOL, J.; ABDEL-WAHAB, O. The Role of Additional Sex-Combs Like Proteins in Cancer. **Cold Spring Harbor Perspectives in Medicine**, v. 6, n. 10, 2016.

MICOL, J. *et al.* ASXL2 is essential for haematopoiesis and acts as a haploinsufficient tumour suppressor in leukemia. **Nature Communications**, v. 8, n. 15429, 2017.

MILLS, A. A. Throwing the Cancer Switch: Reciprocal Roles of Polycomb and Trithorax Proteins. **Nature Reviews Cancer**, v. 10, n. 10, p.669-682, 2010.

MOMPARLER, R. L.; CÔTÉ, S. Targeting of cancer stem cells by inhibitors of DNA and histone methylation. **Expert Opinion on Investigational Drugs**, v. 24, n. 8, p. 1031-1043, 2015.

MOODY, R. R. *et al.* Probing the interaction between the histone methyltransferase/deacetylase subunit RBBP4/7 and the transcription factor BCL11A in epigenetic complexes. **Journal of Biological Chemistry**, v. 293, n. 6, p. 2125-2136, 2018.

NABESHIMA, K. *et al.* Matrix metalloproteinases in tumor invasion: Role for cell migration. **Pathology International**, v. 52, n. 4, p. 255-264, 2002.

NARLIKAR, G. J.; SUNDARAMOORTHY, R.; OWEN-HUGHES, T. Mechanisms and Functions of ATP-Dependent Chromatin-Remodeling Enzymes. **Cell**, v. 154, p. 490-503, 2013.

NATIONAL CANCER INSTITUTE. Breast Cancer. 2015. Disponível em: <<http://www.cancer.gov/types/breast>>. Acesso em: 23 mai. 2015, 14h.

NATIONAL CANCER INSTITUTE. Chemotherapy – PubMed Health Glossary. 2016. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmedhealth/PMHT0024232/>>. Acesso em 19 mar. 2016, 17h.

NERVI, C.; DE MARINIS, E.; CODACCI-PISANELLI, G. Epigenetic treatment of solid tumours: a review of clinical trials. **Clinical Epigenetics**, v. 7, n. 127, 2015.

NEUSQUEN, L. P. D. G. *et al.* EZH2 Protein Expression and Tumor Response to Neoadjuvant Chemotherapy in Locally Advanced Breast Cancer. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 38, n. 6, 2016.

NOUNOU, M. I. *et al.* Breast Cancer: Conventional Diagnosis and Treatment Modalities and Recent Patents and Technologies. **Breast Cancer: Basic and Clinical Research**, v. 9, n. 2, p. 17-34, 2015.

ONODERA, A.; NAKAYAMA, T. Epigenetics of T cells regulated by Polycomb/Trithorax molecules. **Trends in Molecular Medicine**, v.21, n. 5, p.330-340, 2015.

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS). Cancer. 2018. Disponível em: <<http://www.who.int/en/news-room/fact-sheets/detail/cancer>>. Acesso em: 06 jun. 2018, 15h.

PANCHENKO, M. V. Structure, function and regulation of Jade Family PHD Finger 1 (JADE1). **Gene**, v. 589, n. 1, p. 1-11, 2016.

PARK, U. H. *et al.* ASXL2 promotes proliferation of breast cancer cells by linking ER $\alpha$  to histone methylation. **Oncogene**, v. 35, n. 28, p. 3742-3752, 2016.

PASKA, V.; HUDLER, P. Aberrant methylation patterns in cancer: a clinical view. **Biochimica Medica**, v. 25, n. 2, p. 161-176, 2015.

PELLETIER, N. *et al.* MOZ and MORF histone acetyltransferases interact with the Runt-domain transcription factor Runx2. **Oncogene**, v. 21, p. 27-29-2740, 2002.

PERNAS, S.; BARROSO-SOUSA, R.; TOLANEY, S. M. Optimal treatment of early stage HER2-positive breast cancer. **Cancer**, v. 124, n. 23, p. 4455-4466, 2018.

PEROU, C. M. *et al.* Molecular portraits of human breast tumours. **Letters to Nature**, v. 406, p. 747-752, 2000.

PIETERSEN, A. M. *et al.* EZH2 and BMI1 inversely correlate with prognosis and TP53 mutation in breast cancer. **Breast Cancer Research**, v. 10, n. 6, 2008.



PIZZATTI, L. et al. SUZ12 is a candidate target of the non-canonical WNT pathway in the progression of chronic myeloid leukemia. **Genes Chromosomes Cancer**, v. 49, n. 2, p. 107-118, 2010.

POHL, S. et al. Wnt signaling in triple-negative breast cancer. **Oncogenesis**, v. 6, n. 4, 2017.

POYNTER, S. T.; KADOCH, C. Polycomb and trithorax opposition in development and disease. **Wiley Interdisciplinary Reviews: Developmental Biology**, v. 5, n. 6, p. 659-688, 2016.

PROIETTI, S. et al. Increase in motility and invasiveness of MCF7 cancer cells induced by nicotine is abolished by melatonin through inhibition of ERK phosphorylation. **Journal of Pineal Research**, v. 64, n. 4, 2018.

RABELLO, D. A. et al. Altered expression of MLL methyltransferase family genes in breast cancer. **International Journal of Oncology**, v. 42, n. 2, p. 653-660, 2013.

RAMAIAH, M. J.; VAISHNAVE, S. BMI1 and PTEN are key determinants of breast cancer therapy: A plausible therapeutic target in breast cancer. **Gene**, v. 678, p. 302-311, 2018.

REINBOLT, R. E. et al. Endocrine Therapy in Breast Cancer: The Neoadjuvant, Adjuvant and Metastatic Approach. **Seminars in Oncology Nursing**, v. 31, n. 2, p. 146-155, 2015.

RICHTIG, G. et al. SOX9 is a proliferation and stem cell factor in hepatocellular carcinoma and possess widespread prognostic significance in different cancer types. **PLoS One**, v. 12, n. 11, 2017.

RODRIGUEZ-BRAVO, V. et al. The role of GATA2 in lethal prostate cancer aggressiveness. **Nature Reviews Urology**, v. 14, n. 1, p. 38-48, 2017.

ROMAGNOLO, D. F. et al. Epigenetics of breast cancer: modifying role of environmental and bioactive food compounds. **Molecular Nutrition and Food Research**, v. 60, n. 6, p. 1310-1329, 2016.

ROUSSET-JABLONSKI, C.; GOMPEL, A. Screening for familial cancer risk: Focus on breast cancer. **Maturitas**, v. 105, p. 69-77, 2017.

SACHAN, M.; KAUR, M. Epigenetic Modifications: Therapeutic Potential in Cancer. **Brazilian Archives of Biology and Technology**, v. 58, n. 4, p. 526-539, 2015.

SAKSOUK, N. *et al.* HBO1 HAT COMPLEXES TARGET CHROMATIN THROUGHOUT GENE CODING REGIONS VIA MULTIPLE PHD FINGER INTERACTIONS WITH HISTONE H3 TAIL. **Molecular Cell**, v. 33, n. 2, p. 257-265, 2009.

SAN, K. *et al.* Deregulated expression of microRNA-200b/c and SUZ12, a Polycomb repressive complex 2 subunit, in chemoresistant colorectal cancer cells. **Genes and Cancer**, v. 8, n. 7-8, p. 673-681, 2017.

SCALIA-WILBUR, J. *et al.* Breast Cancer Risk Assessment: Moving Beyond BRCA 1 and 2. **Seminars in Radiation Oncology**, v. 26, p. 3-8, 2016.

SCELFO, A.; PIUNTI, A; PASINI, D. The controversial role of the Polycomb group proteins in transcription and cancer: how much do we not understand Polycomb proteins?. **The FEBS Journal**, v. 282, p.1703-1722, 2014.

SCHUETTENGRUBER, B. *et al.* Genome Regulation by Polycomb and Trithorax Proteins. **Cell**, n. 128, p. 735-745, 2007.

SCHUETTENGRUBER, B. *et al.* Trithorax group proteins: switching genes on and keeping them active. **Nature Reviews**, v. 12, p. 799-814, 2011.

SHAH, N.; SUKUMAR, S. The Hox genes and their roles in oncogenesis. **Nature Reviews Cancer**, v. 10, n. 5, p. 361-371, 2010.

SHAIN, A. H.; POLLACK, J. R. The Spectrum of SWI/SNF Mutations, Ubiquitous in Human Cancers. **PLoS One**, v. 8, n. 1, 2013.

SHEN, J. *et al.* The E3 Ligase RING1 Targets p53 for Degradation and Promotes Cancer Cell Proliferation and Survival. **Cancer Research**, v. 78, n. 2, p. 359-371, 2018.

SEYMOUR, T.; TWIGGER, A.; KAKULAS, F. Pluripotency Genes and Their Functions in the Normal and Aberrant Breast and Brain. **International Journal of Molecular Sciences**, v. 16, n. 11, p. 27288-27301, 2015.

SINGLETARY, S. E. Rating the Risk Factors for Breast Cancer. **Annals of Surgery**, v. 237, n. 4, p. 474-482, 2003.

SINN, H.; KREIPE, H. A Brief Overview of the WHO Classification of Breast Tumours, 4<sup>th</sup> Edition, Focusing on Issues and Updates from the 3<sup>rd</sup> Edition. **Breast Care**, v. 8, p. 149-154, 2013.

SKOL, A. D.; SASAKI, M. M.; ONEL, K. The genetics of breast cancer risk in the post-genome era: thoughts on study design to move past *BRCA* and towards clinical relevance. **Breast Cancer Research**, v. 18, n. 99, 2016.

SONG, X. *et al.* CBX8 and CD96 Are Important Prognostic Biomarkers of Colorectal Cancer. **Medical Science Monitor**, v. 24, p. 7820-7827, 2018.

STEFFEN, P. A.; RINGROSE, L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. **Nature Reviews**, v. 15, p. 340-356, 2014.

SUN, Y. *et al.* Protein arginine methyltransferase 6 enhances ligand-dependent and -independent activity of estrogen receptor  $\alpha$  via distinct mechanisms. **Biochimica et Biophysica Acta**, v. 1843, p. 2067-2078, 2014.

TABARIES, S.; SIEGEL, P. M. The role of claudins in cancer metastasis. **Oncogene**, v. 36, p. 1176-1190, 2017.

TAN, J. *et al.* EZH2: biology, disease and structure-based drug discovery. **Acta Pharmacologica Sinica**, v. 35, n. 2, p. 161-174, 2013.

TANDON, M. *et al.* The role of Runx2 in facilitating autophagy in metastatic breast cancer cells. **Journal of Cellular Physiology**, v. 233, n. 1, p. 559-571, 2017.

TAVASSOLI, F. A.; DEVILEE, P. Pathology and Genetics of Tumours of the Breast and Female Genital Organs. Lyon: International Agency for Research on Cancer, 2003.

TOLLERVEY, J. R.; LUNYAK, V. V. Epigenetics: Judge, jury and executioner of stem cell fate. **Epigenetics**, v. 7, n. 8, p. 823-840, 2012.

ULLAH, M. *et al.* Molecular Architecture of Quartet MOZ/MORF Histone Acetyltransferase Complexes. **Molecular and Cellular Biology**, v. 28, n. 22, p. 6828-6843, 2008.

VISHAL. M. *et al.* Role of Runx2 in breast cancer-mediated bone metastasis. **International Journal of Biological Macromolecules**, v. 99, p. 608-614, 2017.

VUONG, D. *et al.* Molecular Classification of Breast Cancer. **Virchows Archiv**, v. 465, n. 1, p. 1-14, 2014.

WANG, L. *et al.* Interleukin 21 and Its Receptor Play a Role in Proliferation, Migration and Invasion of Breast Cancer Cells. **Cancer Genomics and Proteomics**, v. 12, p. 211-222, 2015.

WAN, M. *et al.* Yin Yang 1 Plays an Essential Role in Breast Cancer and Negatively Regulates p27. **The American Journal of Pathology**, v. 180, n. 5, p. 2120-2133, 2012.

WANG, H. *et al.* Estrogen receptor  $\alpha$ -coupled Bmi1 regulation pathway in breast cancer and its clinical implications. **BMC Cancer**, v. 14, n. 122, 2014.

WANG, X. *et al.* Clinical and prognostic relevance of EZH2 in breast cancer: A meta-analysis. **Biomedicine and Pharmacotherapy**, v. 75, p. 218-225, 2015.

WASSEF, M.; MARGUERON, R. The Multiple Facets of PRC2 Alterations in Cancers. **Journal of Molecular Biology**, 2016.

WINSHIP, I.; SOUTHEY, M. C. Gene panel testing for hereditary breast cancer. **Medical Journal of Australia**, v. 204, n. 5, p. 188-190, 2016.

WONG, E.; REBELO, J. Breast Cancer Pathogenesis and Histologic vs. Molecular Subtypes. 2012. Disponível em: <<http://www.pathophys.org/breast-cancer/breastcancer-copy/>>. Acesso em: 06 jun. 2018, 15h.

WU, K. *et al.* Extracellular vesicles as emerging targets in cancer: recent development from bench to bedside. **Biochimica et Biophysica Acta**, v. 1868, n. 2, p. 538-563, 2017.

WU, Q. *et al.* The SWI/SNF ATPases are Required for Triple Negative Breast Cancer Cell Proliferation. **Journal of Cellular Physiology**, v. 230, n. 11, p. 2683-2694, 2015.

WU, T. *et al.* Increased MMP-21 expression is associated with poor overall survival of patients with gastric cancer. **Medical Oncology**, v. 30, n. 1, 2013.

XIA, R. *et al.* SUZ12 promotes gastric cancer cell proliferation and metastasis by regulating KLF2 and E-cadherin. **Tumor Biology**, v. 36, n. 7, p. 5341-5351, 2015.

XIAO-FEN, W. *et al.* Correlation analysis of VHL and Jade-1 gene expression in human renal cell carcinoma. **Open Medicine**, v. 11, n. 1, p. 226-230, 2016.

XIE, F. *et al.* TGF- $\beta$  signaling in cancer metastasis. **Acta Biochimica et Biophysica Sinica**, v. 50, n. 1, p. 121-132, 2018.

XU, Y.; YANG, W. SOX-mediated molecular crosstalk during the progression of tumorigenesis. **Seminars in Cell and Developmental Biology**, v. 63, p. 23-34, 2017.

YAMAGUCHI, H.; HUNG, M. Regulation and Role of EZH2 in Cancer. **Cancer Research and Treatment**, v. 46, n. 3, p. 209-222, 2014.

YANG, L. *et al.* Reduced Expression of TET1, TET2, TET3 and TDG mRNAs Are Associated with Poor Prognosis of Patients with Early Breast Cancer. **PloS One**, v. 10, n. 7, 2015.

YE, Y. *et al.* Inhibition of Expression of the Chromatin Remodelling Gene, *SNF2L*, Selectively Leads to DNA Damage, Growth Inhibition, and Cancer Cell Death. **Molecular Cancer Research**, v. 7, n. 12, p. 1984-1999, 2009.

YOO, K. H.; HENNIGHAUSEN, L. EZH2 Methyltransferase and H3K27 Methylation in Breast Cancer. **International Journal of Biological Sciences**, v. 8, n. 1, p. 59-65, 2012.

YOU, J. S.; JONES, P. A. Cancer Genetics and Epigenetics: Two Sides of the Same Coin?. **Cancer Cell**, v. 22, n. 1, p. 9-20, 2012.

YUAN, G. *et al.* Chromobox homolog 8 is a predictor of muscle invasive bladder cancer and promotes cell proliferation by repressing the p53 pathway. **Cancer Science**, v. 108, n. 11, p. 2166-2175, 2017.

ZAHA, D. C. Significance of immunohistochemistry in breast cancer. **World Journal of Clinical Oncology**, v. 5, n. 3, p. 382-392, 2014.

ZHAN, T.; RINDTORFF, N.; BOUTROS, M. Wnt signaling in cancer. **Oncogene**, v. 36, p. 1461-1473, 2017.

ZHANG, C. Z. *et al.* CBX8 Exhibits Oncogenic Activity via AKT/ $\beta$ -Catenin Activation in Hepatocellular Carcinoma. **Cancer Research**, v. 78, n. 1, 2018.

ZHANG, J. *et al.* Overexpression of MMP21 and MMP28 is associated with gastric cancer progression and poor prognosis. **Oncology Letters**, v. 15, n. 5, p. 7776-7782, 2018.

ZHANG, Q. *et al.* Yin Yang 1 promotes mTORC2-mediated AKT phosphorylation. **Journal of Molecular Cell Biology**, v. 8, n. 3, p. 232-243, 2016.

ZHANG, W.; XU, J. DNA methyltransferases and their roles in tumorigenesis. **Biomarker Research**, v. 5, n. 1, 2017.

ZHANG, Y. *et al.* The roles of ZEB1 in tumorigenic progression and epigenetic modifications. **Biomedicine and Pharmacotherapy**, v. 110, p. 400-408, 2019.

ZIOGAS, D. E. Risk factors and early detection of breast cancer: facts, questions and genome-based perspectives. **Current Oncology**, v. 19, n. 3, p. 131, 2012.

## 9. ANEXOS

**Anexo A - Tabela dos genes diferencialmente expressos após silenciamento de *ASXL2* na linhagem MCF-7 (fold change 2,0)**

Símbolo do Gene	Fold change	Descrição
NPCDR1	-123.75873	nasopharyngeal carcinoma, down-regulated 1
LCN10	-63.39992	lipocalin 10
IL31	-41.254425	interleukin 31
SOX2	-41.1441	SRY (sex determining region Y)-box 2
SEMA3A	-40.813164	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A
ERICH1-AS1	-37.765213	ERICH1 antisense RNA 1
OR9I3P	-32.65475	Olfactory Receptor Family 9 Subfamily I Member 3 Pseudogene
ZNF280D	-28.953295	Zinc Finger Protein 280D
AMTN	-22.891617	amelotin
ACADL	-19.530342	acyl-CoA dehydrogenase, long chain
PDCL2P1	-19.079407	Phosducin-Like 2 Pseudogene 1
ZNF382	-18.184149	zinc finger protein 382
AC009495.2	-17.778114	ENSG00000232411
CCDC169	-16.235737	coiled-coil domain containing 169
LOC727799	-14.559099	uncharacterized LOC727799
TRIM48	-13.471879	tripartite motif containing 48
VN1R5	-13.416082	vomer nasal 1 receptor 5 (gene/pseudogene)
CDH18	-12.898405	cadherin 18, type 2
C10orf107	-11.661091	chromosome 10 open reading frame 107
NGF	-11.349169	nerve growth factor (beta polypeptide)
GJA1	-10.535743	gap junction protein, alpha 1, 43kDa
TLR3	-9.694027	toll-like receptor 3
LINC00993	-9.111276	long intergenic non-protein coding RNA 993
LIPH	-9.047566	lipase, member H

LOC100130992	-8.810407	uncharacterized LOC100130992
LINC00705	-8.515787	long intergenic non-protein coding RNA 705
OR5AK2	-8.377168	olfactory receptor, family 5, subfamily AK, member 2
HPCAL4	-8.037397	hippocalcin like 4
CD36	-7.9715915	CD36 molecule (thrombospondin receptor)
PIK3AP1	-7.725448	phosphoinositide-3-kinase adaptor protein 1
lnc-GGCT-1	-7.562059	lnc-GGCT-1:19
LMCD1-AS1	-7.3416224	LMCD1 antisense RNA 1 (head to head)
IPMK	-7.331536	inositol polyphosphate multikinase
TGFB2-AS1	-7.2378874	TGFB2 antisense RNA 1 (head to head)
ARMC4	-7.0739975	armadillo repeat containing 4
lnc-GMDS-2	-7.010199	lnc-GMDS-2:1
C9orf66	-6.9864497	chromosome 9 open reading frame 66
ADM	-6.886885	adrenomedullin
LOC644070	-6.8614125	germ cell-specific gene 1-like protein 2-like
LINC00251	-6.7820215	long intergenic non-protein coding RNA 251
FAM150B	-6.711595	family with sequence similarity 150, member B
MYBPC1	-6.5448303	myosin binding protein C, slow type
CNBD1	-6.446674	cyclic nucleotide binding domain containing 1
LINC02256	-6.419352	Long Intergenic Non-Protein Coding RNA 2256
GRIK1-AS2	-6.314477	GRIK1 antisense RNA 2
LOC399715	-6.252729	uncharacterized LOC399715
SEMA3E	-6.240925	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3E
TRAV2	-6.226599	T Cell Receptor Alpha Variable 2
ITGB8	-6.1380067	integrin, beta 8
IGHV3-33	-6.0942554	Immunoglobulin Heavy Variable 3-33
SULF1	-5.8851485	sulfatase 1
KCNQ1DN	-5.81574	KCNQ1 downstream neighbor (non-protein coding)
STATH	-5.7976437	statherin



PCDH20	-5.6965065	protocadherin 20
TCN1	-5.4938774	transcobalamin I (vitamin B12 binding protein, R binder family)
MEF2C	-5.3927875	myocyte enhancer factor 2C
AGR2	-5.349051	anterior gradient 2
BCAS1	-5.3167877	breast carcinoma amplified sequence 1
ABCC6P1	-5.282704	ATP-binding cassette, sub-family C, member 6 pseudogene 1 (functional)
ARHGAP15	-5.2819448	Rho GTPase activating protein 15
CALN1	-5.167098	calneuron 1
PCGEM1	-5.1611977	PCGEM1, prostate-specific transcript (non-protein coding)
ARHGEF38	-5.078607	Rho guanine nucleotide exchange factor (GEF) 38
SPINK8	-5.0273833	serine peptidase inhibitor, Kazal type 8 (putative)
HCG4	-5.026003	HLA complex group 4 (non-protein coding)
ANKRD30A	-4.8528476	ankyrin repeat domain 30A
SMIM17	-4.690255	small integral membrane protein 17
MACC1	-4.667976	metastasis associated in colon cancer 1
CKM	-4.6628237	creatine kinase, muscle
MLNR	-4.611054	motilin receptor
JAZF1	-4.5987644	JAZF zinc finger 1
FIGN	-4.568086	fidgetin
CD74	-4.546843	CD74 molecule, major histocompatibility complex, class II invariant chain
ODAM	-4.544613	odontogenic, ameloblast associated
lnc-TCTN3-1	-4.505629	lnc-TCTN3-1:5
GPR87	-4.4872737	G protein-coupled receptor 87
MUM1L1	-4.4369626	melanoma associated antigen (mutated) 1-like 1
ANKRD22	-4.3951907	ankyrin repeat domain 22
DOK6	-4.3798842	docking protein 6
ANKRD20A19P	-4.363639	ankyrin repeat domain 20 family, member A19, pseudogene
CARD18	-4.288469	caspase recruitment domain family, member 18
lnc-RP3-377D14.1.1-3	-4.2286286	lnc-RP3-377D14.1.1-3:19

MRPL23-AS1	-4.153438	MRPL23 antisense RNA 1
ASIC5	-4.1422772	acid sensing (proton gated) ion channel family member 5
lnc-ZIC1-1	-4.125367	lnc-ZIC1-1:1
TMEM229A	-4.117371	transmembrane protein 229A
C1QTNF4	-4.10462	C1q and tumor necrosis factor related protein 4
LOC729696	-4.0964427	uncharacterized LOC729696
DKFZp547J222	-4.071082	uncharacterized LOC84237
COLCA1	-4.0656037	colorectal cancer associated 1
LOC154761	-4.0451684	family with sequence similarity 115, member C pseudogene
OR2T5	-4.0312157	olfactory receptor, family 2, subfamily T, member 5
LOC100133616	-4.031004	uncharacterized LOC100133616
POU1F1	-4.0309057	POU class 1 homeobox 1
DSG1-AS1	-3.9788427	DSG1 antisense RNA 1
ADRA2A	-3.9501812	adrenoceptor alpha 2A
AC142086.6	-3.947688	ENSG00000277304
KIAA1661	-3.9347138	KIAA1661 protein
CXorf51A	-3.9283714	chromosome X open reading frame 51A
ADD3	-3.913287	adducin 3 (gamma)
SLC7A13	-3.907685	solute carrier family 7 (anionic amino acid transporter), member 13
CACNG5	-3.8804634	calcium channel, voltage-dependent, gamma subunit 5
SPATA31A3	-3.8754	SPATA31 subfamily A, member 3
CLCN4	-3.866986	chloride channel, voltage-sensitive 4
RIMS1	-3.8456025	regulating synaptic membrane exocytosis 1
TOX3	-3.8285537	TOX high mobility group box family member 3
CDH16	-3.8088932	cadherin 16, KSP-cadherin
MID1	-3.7985866	midline 1
CST1	-3.7941787	cystatin SN
IL18	-3.778779	interleukin 18
LOC100996844	-3.761605	mucin-3A-like

BANK1	-3.7359896	B-cell scaffold protein with ankyrin repeats 1
DMD	-3.7229211	dystrophin
TRBV11-2	-3.718282	T Cell Receptor Beta Variable 11-2
OR8G1	-3.7130637	olfactory receptor, family 8, subfamily G, member 1 (gene/pseudogene)
LBX1-AS1	-3.6774473	LBX1 antisense RNA 1 (head to head)
AMER3	-3.633627	APC membrane recruitment protein 3
NIM1K	-3.6161332	NIM1 serine/threonine protein kinase
LOC100653133	-3.6099875	golgin subfamily A member 6-like protein 1-like
SVOPL	-3.5838976	SVOP-like
FLJ31104	-3.5698533	uncharacterized LOC441072
LOC255654	-3.5694003	uncharacterized LOC255654
LOC100240734	-3.5586352	uncharacterized LOC100240734
CCDC36	-3.5468314	coiled-coil domain containing 36
ETV5	-3.5462937	ets variant 5
SPRR2B	-3.512968	small proline-rich protein 2B
NAT16	-3.5039	N-acetyltransferase 16 (GCN5-related, putative)
lnc-C10orf71-1	-3.4989753	lnc-C10orf71-1:1
LINC00622	-3.4623203	long intergenic non-protein coding RNA 622
LRRC29	-3.4559464	leucine rich repeat containing 29
BPIFB1	-3.4558759	BPI fold containing family B, member 1
LOC100131910	-3.4521847	uncharacterized LOC100131910
SLC25A30-AS1	-3.4418206	SLC25A30 antisense RNA 1
DAZL	-3.4252834	deleted in azoospermia-like
OACYLP	-3.4181762	O-acyltransferase like, pseudogene
LINC00112	-3.4007068	long intergenic non-protein coding RNA 112
XLOC_12_012147	-3.3979187	XLOC_12_012147
MIR3945HG	-3.381886	MIR3945 Host Gene
CCDC176	-3.3812752	coiled-coil domain containing 176
NRK	-3.3777907	Nik related kinase

VAV3	-3.3765666	vav 3 guanine nucleotide exchange factor
SOX9-AS1	-3.356618	SOX9 antisense RNA 1
ASXL2	-3.338976	additional sex combs like transcriptional regulator 2
CXCR2	-3.3235612	chemokine (C-X-C motif) receptor 2
CRYBG3	-3.3093016	beta-gamma crystallin domain containing 3
LOC101927133	-3.3065112	uncharacterized LOC101927133
LINC00575	-3.2917268	long intergenic non-protein coding RNA 575
L1TD1	-3.2883554	LINE-1 type transposase domain containing 1
OR1B1	-3.2851021	olfactory receptor, family 1, subfamily B, member 1 (gene/pseudogene)
LOC101929689	-3.2705748	uncharacterized LOC101929689
PRUNE2	-3.2552016	prune homolog 2 (Drosophila)
SMG6-IT1	-3.237643	SMG6 intronic transcript 1 (non-protein coding)
PAPSS2	-3.236283	3'-phosphoadenosine 5'-phosphosulfate synthase 2
UGT2B15	-3.2161543	UDP glucuronosyltransferase 2 family, polypeptide B15
RNF17	-3.215066	ring finger protein 17
IGHG1	-3.2109084	Immunoglobulin Heavy Constant Gamma 1 (G1m Marker)
CR1	-3.2010217	complement component (3b/4b) receptor 1 (Knops blood group)
KCNJ2	-3.197802	potassium channel, inwardly rectifying subfamily J, member 2
ZNF140	-3.1913328	zinc finger protein 140
EPB41L4A-AS2	-3.1597147	EPB41L4A antisense RNA 2 (head to head)
ACPP	-3.1264398	acid phosphatase, prostate
FGA	-3.1217034	fibrinogen alpha chain
CXorf65	-3.117644	chromosome X open reading frame 65
AC083843.1	-3.1175497	Non-Protein Coding RNA 250
SLITRK6	-3.1138883	SLIT and NTRK-like family, member 6
VN1R64P	-3.109577	Vomer nasal 1 Receptor 64 Pseudogene
MIR99AHG	-3.1086273	mir-99a-let-7c cluster host gene (non-protein coding)
HTR1A	-3.102446	5-hydroxytryptamine (serotonin) receptor 1A, G protein-coupled
OR4N4	-3.1001768	olfactory receptor, family 4, subfamily N, member 4

ZC2HC1A	-3.0942771	zinc finger, C2HC-type containing 1A
MYO15A	-3.0873234	myosin XVA
IL20RA	-3.0827854	interleukin 20 receptor, alpha
GALNT15	-3.0754542	polypeptide N-acetylgalactosaminyltransferase 15
ERICH6B	-3.0735195	glutamate-rich 6B
DHRS2	-3.0635772	dehydrogenase/reductase (SDR family) member 2
ESM1	-3.0459487	endothelial cell-specific molecule 1
PLAG1	-3.0456839	pleiomorphic adenoma gene 1
SERP1	-3.030902	Stress Associated Endoplasmic Reticulum Protein 1
AC137767.1	-3.0162122	ENSG00000256092
TPK1	-3.0132005	thiamin pyrophosphokinase 1
NRAP	-3.0100963	nebulin-related anchoring protein
HOXD8	-3.0096323	homeobox D8
RGS9BP	-3.0072646	regulator of G protein signaling 9 binding protein
GAF1	-2.9995105	Gene associated with FGF-2 activity protein 1
AC099489.1	-2.994577	Vitellogenin
LOC101927609	-2.9830132	uncharacterized LOC101927609
PDK1	-2.9809108	pyruvate dehydrogenase kinase, isozyme 1
GUCY2D	-2.9800496	guanylate cyclase 2D, membrane (retina-specific)
LINC00992	-2.9778848	long intergenic non-protein coding RNA 992
GNGT2	-2.973342	guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 2
ARSI	-2.9709535	arylsulfatase family, member I
HOXC11	-2.970404	homeobox C11
B3GNT5	-2.962279	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5
KLF7	-2.9586227	Kruppel-like factor 7 (ubiquitous)
TARP	-2.9390974	TCR gamma alternate reading frame protein
SAA3P	-2.9280667	serum amyloid A3 pseudogene
SPANXN1	-2.9259684	SPANX family, member N1
LOC101929408	-2.9254322	uncharacterized LOC101929408

RIN2	-2.92293	Ras and Rab interactor 2
PTPN21	-2.9102151	protein tyrosine phosphatase, non-receptor type 21
LOC730257	-2.909387	uncharacterized LOC730257
SGK2	-2.8938558	serum/glucocorticoid regulated kinase 2
LOC729970	-2.8924851	hCG2028352-like
TTC9	-2.880361	tetratricopeptide repeat domain 9
RASGRF2	-2.8783567	Ras protein-specific guanine nucleotide-releasing factor 2
KCNC1	-2.8753872	potassium channel, voltage gated Shaw related subfamily C, member 1
PP14571	-2.8719742	uncharacterized LOC100130449
SGMS2	-2.867989	sphingomyelin synthase 2
TSPAN12	-2.8610172	tetraspanin 12
CYP4Z1	-2.8606782	cytochrome P450, family 4, subfamily Z, polypeptide 1
MIR31HG	-2.8600032	MIR31 host gene (non-protein coding)
AL671710.1	-2.8598657	ENSG00000273192
COLEC12	-2.8512475	collectin sub-family member 12
CDH10	-2.8459563	cadherin 10, type 2 (T2-cadherin)
LOC646976	-2.8398182	uncharacterized LOC646976
MDS2	-2.8346741	myelodysplastic syndrome 2 translocation associated
CEP41	-2.8345625	centrosomal protein 41kDa
LOC100132215	-2.8343651	uncharacterized LOC100132215
SLC27A6	-2.8216655	solute carrier family 27 (fatty acid transporter), member 6
SLC28A3	-2.8193686	solute carrier family 28 (concentrative nucleoside transporter), member 3
PCDH10	-2.8058162	protocadherin 10
KANSL1L	-2.7906559	KAT8 regulatory NSL complex subunit 1-like
PDE11A	-2.789881	phosphodiesterase 11A
VWF	-2.7819245	von Willebrand factor
ZNF610	-2.780598	zinc finger protein 610
lnc-DPP4-2	-2.7795396	lnc-DPP4-2:3
SNAR-C3	-2.7758226	small ILF3/NF90-associated RNA C3

L3MBTL3	-2.7691915	l(3)mbt-like 3 (Drosophila)
C8orf49	-2.756866	chromosome 8 open reading frame 49
BPIFA4P	-2.756223	BPI fold containing family A, member 4, pseudogene
SATB2-AS1	-2.754845	SATB2 antisense RNA 1
CDIPT-AS1	-2.75018	CDIPT antisense RNA 1 (head to head)
ACSM1	-2.7490077	acyl-CoA synthetase medium-chain family member 1
TMEM246	-2.7464721	transmembrane protein 246
lnc-AGPAT6-1	-2.7388988	lnc-AGPAT6-1:1
TARS2	-2.738811	Threonyl-TRNA Synthetase 2, Mitochondrial
LINC00869	-2.7387252	long intergenic non-protein coding RNA 869
MGAT3	-2.7354965	mannosyl (beta-1,4-)-glycoprotein beta-1,4-N-acetylglucosaminyltransferase
LINC00324	-2.734783	long intergenic non-protein coding RNA 324
JHDM1D-AS1	-2.7294033	JHDM1D antisense RNA 1 (head to head)
PYGM	-2.7293098	phosphorylase, glycogen, muscle
EPB41L4A	-2.721495	erythrocyte membrane protein band 4.1 like 4A
ILDR2	-2.7208924	immunoglobulin-like domain containing receptor 2
RFTN2	-2.717974	raftlin family member 2
ERC2	-2.7161617	ELKS/RAB6-interacting/CAST family member 2
PTPRB	-2.7148323	protein tyrosine phosphatase, receptor type, B
RRAGD	-2.7141478	Ras-related GTP binding D
MYRIP	-2.7136672	myosin VIIA and Rab interacting protein
XAGE5	-2.7115486	X antigen family, member 5
FRRS1	-2.7105439	ferric-chelate reductase 1
IKZF2	-2.696631	IKAROS family zinc finger 2 (Helios)
FANK1	-2.6964366	fibronectin type III and ankyrin repeat domains 1
ATL1	-2.6872709	atlastin GTPase 1
ERVV-2	-2.682988	endogenous retrovirus group V, member 2
NR3C2	-2.6785064	nuclear receptor subfamily 3, group C, member 2
FRMD6-AS1	-2.6738346	FRMD6 antisense RNA 1

LURAPIL	-2.6674347	leucine rich adaptor protein 1-like
LOC727944	-2.6645691	uncharacterized LOC727944
HOXA1	-2.6626186	homeobox A1
GPR101	-2.6539292	G protein-coupled receptor 101
DSC1	-2.6522083	desmocollin 1
Inc-HRASLS-3	-2.6488466	Inc-HRASLS-3:1
TRBV28	-2.648472	T Cell Receptor Beta Variable 28
AC068594.1	-2.6482255	Uncharacterized LOC105371907
LOC256880	-2.6435494	uncharacterized LOC256880
ERP27	-2.6431258	endoplasmic reticulum protein 27
SPATA24	-2.641468	spermatogenesis associated 24
MEIS1	-2.6333952	Meis homeobox 1
C3orf80	-2.633157	chromosome 3 open reading frame 80
CST6	-2.6273677	cystatin E/M
APOBEC3H	-2.621867	apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3H
DMRTC1	-2.6213064	DMRT-like family C1
PLSCR4	-2.6103082	phospholipid scramblase 4
AC104131.1	-2.606947	Chromosome 18 Open Reading Frame 25 Pseudogene
PDLIM3	-2.597701	PDZ and LIM domain 3
MYPN	-2.5966706	myopalladin
VEGFC	-2.5815227	vascular endothelial growth factor C
SPOCK3	-2.571649	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3
SCN9A	-2.5699434	sodium channel, voltage gated, type IX alpha subunit
CHRNA4	-2.5692933	cholinergic receptor, nicotinic, beta 4 (neuronal)
WFIKKN1	-2.5671601	WAP, follistatin/kazal, immunoglobulin, kunitz and netrin domain containing 1
NT5DC4	-2.56638	5'-nucleotidase domain containing 4
HEPHL1	-2.5642421	hephaestin-like 1
TRBV10-1	-2.563931	T Cell Receptor Beta Variable 10-1(Gene/Pseudogene)
SOX9	-2.5610116	SRY (sex determining region Y)-box 9



SPERT	-2.5565577	spermatid associated
lnc-TAF1C-1	-2.5520089	lnc-TAF1C-1:1
CDK6	-2.5444384	cyclin-dependent kinase 6
ADRB1	-2.5411649	adrenoceptor beta 1
CXADR	-2.5366535	coxsackie virus and adenovirus receptor
SCGB1D1	-2.5355453	secretoglobin, family 1D, member 1
FAM170A	-2.5325513	family with sequence similarity 170, member A
BAALC-AS1	-2.5293975	BAALC antisense RNA 1
VEPH1	-2.5254462	ventricular zone expressed PH domain-containing 1
ADCY10P1	-2.5217886	adenylate cyclase 10 (soluble) pseudogene 1
LRRC53	-2.5198328	Leucine Rich Repeat Containing 53
LOC646588	-2.5187998	uncharacterized LOC646588
LOC100288911	-2.5173683	uncharacterized LOC100288911
KLHL4	-2.517003	kelch-like family member 4
TTN	-2.5136673	titin
ASCL1	-2.510043	achaete-scute family bHLH transcription factor 1
AP000619.1	-2.5093768	ENSG00000225678
AC147055.1	-2.5092094	UDP Glucuronosyltransferase Family 2 Member B15 Pseudogene
TMEM45A	-2.50886	transmembrane protein 45A
CSRP2	-2.506942	cysteine and glycine-rich protein 2
DENN2C	-2.5064251	DENN/MADD domain containing 2C
BCL2L10	-2.501072	BCL2-like 10 (apoptosis facilitator)
PARP8	-2.4989424	poly (ADP-ribose) polymerase family, member 8
UPK2	-2.4951823	uroplakin 2
ITGB3	-2.494289	integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)
SLC1A1	-2.490597	solute carrier family 1 (neuronal/epithelial high affinity glutamate transporter, system Xag), member 1
IZUMO4	-2.4883935	IZUMO family member 4
lnc-SEC61G-7	-2.4881968	lnc-SEC61G-7:3
UBTFL7	-2.4880576	UBTF Like 7 (Pseudogene)

LOC285638	-2.4879115	uncharacterized LOC285638
lnc-KIAA0087-2	-2.4870787	lnc-KIAA0087-2:2
LOC101928561	-2.4869678	uncharacterized LOC101928561
lnc-TMEM144-1	-2.4865322	lnc-TMEM144-1:1
FGR	-2.4838986	FGR proto-oncogene, Src family tyrosine kinase
NTN4	-2.4782388	netrin 4
MATN3	-2.4746945	matrilin 3
lnc-VIM-1	-2.464861	lnc-VIM-1:2
LOC441052	-2.4642277	uncharacterized LOC441052
LINC00883	-2.463057	long intergenic non-protein coding RNA 883
STAMBPL1	-2.4616065	STAM binding protein-like 1
ANXA1	-2.4614851	annexin A1
KCTD7	-2.461423	potassium channel tetramerization domain containing 7
FOXN4	-2.4607227	forkhead box N4
EIF4E1B	-2.460196	eukaryotic translation initiation factor 4E family member 1B
GJD2	-2.4581933	gap junction protein, delta 2, 36kDa
CYP7B1	-2.4550698	cytochrome P450, family 7, subfamily B, polypeptide 1
PITPNM3	-2.4545138	PITPNM family member 3
COX8C	-2.453088	cytochrome c oxidase subunit VIIIc
ASCC1	-2.4523354	Activating Signal Cointegrator 1 Complex Subunit 1
ROBO1	-2.4517956	roundabout, axon guidance receptor, homolog 1 (Drosophila)
COLCA2	-2.4511552	colorectal cancer associated 2
PPP1R3G	-2.4502091	protein phosphatase 1, regulatory subunit 3G
HAS2	-2.449627	hyaluronan synthase 2
CEP19	-2.4468424	centrosomal protein 19kDa
IDI2	-2.443956	isopentenyl-diphosphate delta isomerase 2
LGI2	-2.4433775	leucine-rich repeat LGI family, member 2
IL15	-2.4421966	interleukin 15
MARCKS	-2.4386268	myristoylated alanine-rich protein kinase C substrate

LINC00113	-2.4327028	long intergenic non-protein coding RNA 113
DCLK1	-2.4231474	doublecortin-like kinase 1
SLC9B1	-2.4175556	solute carrier family 9, subfamily B (NHA1, cation proton antiporter 1), member 1
GUCY2EP	-2.4171367	guanylate cyclase 2E, pseudogene
INSIG1	-2.4164548	insulin induced gene 1
STX19	-2.415008	syntaxin 19
RARRES1	-2.4142537	retinoic acid receptor responder (tazarotene induced) 1
lnc-KAT7-4	-2.4139585	lnc-KAT7-4:1
CALB1	-2.4129834	calbindin 1, 28kDa
CASC15	-2.4126408	cancer susceptibility candidate 15 (non-protein coding)
FAM13A	-2.4084349	family with sequence similarity 13, member A
TRBV25/OR9-2	-2.4079347	T Cell Receptor Beta Variable 25/OR9-2 (Pseudogene)
CRISP2	-2.407612	cysteine-rich secretory protein 2
GJB6	-2.4067166	gap junction protein, beta 6, 30kDa
TFAP2B	-2.4063478	transcription factor AP-2 beta (activating enhancer binding protein 2 beta)
PLEKHH2	-2.4030464	pleckstrin homology domain containing, family H (with MyTH4 domain) member 2
S100A10	-2.400737	S100 calcium binding protein A10
ANKRD37	-2.4003022	ankyrin repeat domain 37
CH25H	-2.3995223	cholesterol 25-hydroxylase
ADPRH	-2.397176	ADP-ribosylarginine hydrolase
LRRC3B	-2.3942668	leucine rich repeat containing 3B
ADORA3	-2.392193	adenosine A3 receptor
NOS2P3	-2.3904614	Nitric Oxide Synthase 2 Pseudogene 3
MARK1	-2.3896084	MAP/microtubule affinity-regulating kinase 1
LOC102724747	-2.3817718	ovostatin homolog 2-like
GLRX3P2	-2.3755026	Glutaredoxin 3 Pseudogene 2
SLC23A1	-2.374391	solute carrier family 23 (ascorbic acid transporter), member 1
P4HA3	-2.372404	prolyl 4-hydroxylase, alpha polypeptide III
HTRA4	-2.371826	HtrA serine peptidase 4

AHR	-2.3716617	aryl hydrocarbon receptor
SASH1	-2.370147	SAM and SH3 domain containing 1
ADAMTS6	-2.3695931	ADAM metallopeptidase with thrombospondin type 1 motif, 6
SNTN	-2.369444	sentan, cilia apical structure protein
FLJ16734	-2.3687541	uncharacterized LOC641928
RPL22L1	-2.3662717	ribosomal protein L22-like 1
CICP26	-2.3646078	Capicua Transcriptional Repressor Pseudogene 26
LINC01949	-2.362599	Long Intergenic Non-Protein Coding RNA 1949
GSTA5	-2.3604467	glutathione S-transferase alpha 5
PTCRA	-2.354879	pre T-cell antigen receptor alpha
FOXB2	-2.3538966	forkhead box B2
CTTNBP2	-2.3511505	cortactin binding protein 2
SNX10	-2.3490207	sorting nexin 10
CCDC148	-2.3479834	coiled-coil domain containing 148
ANKRD2	-2.3433995	ankyrin repeat domain 2 (stretch responsive muscle)
GAREM	-2.3428102	GRB2 associated, regulator of MAPK1
ASCL4	-2.3416426	achaete-scute family bHLH transcription factor 4
LINC01376	-2.3390777	Long Intergenic Non-Protein Coding RNA 1376
LOC280665	-2.3365679	anti-CNG alpha 1 cation channel translation product-like
GPNMB	-2.3352685	glycoprotein (transmembrane) nmb
KCNJ8	-2.3352308	potassium channel, inwardly rectifying subfamily J, member 8
UBE2E1	-2.3321662	ubiquitin-conjugating enzyme E2E 1
PSG1	-2.3319519	pregnancy specific beta-1-glycoprotein 1
LINC01566	-2.329088	long intergenic non-protein coding RNA 1566
CXorf21	-2.3290591	chromosome X open reading frame 21
LOC643770	-2.3275898	uncharacterized LOC643770
LCT	-2.3199584	lactase
TMEM210	-2.315965	transmembrane protein 210
SLIT1	-2.3150249	slit homolog 1 (Drosophila)

RAG1	-2.3136356	recombination activating gene 1
lnc-AC073343.1-1	-2.3129108	lnc-AC073343.1-1:3
KIAA0408	-2.3126662	KIAA0408
XLOC_12_013861	-2.3123956	XLOC_12_013861
IGHV3/OR16-9	-2.3098862	Immunoglobulin Heavy Variable 3/OR16-9 (Non-Functional)
SPATA18	-2.3071349	spermatogenesis associated 18
SH3BGRL	-2.3063474	SH3 domain binding glutamate-rich protein like
CNRIP1	-2.304862	cannabinoid receptor interacting protein 1
KIAA1107	-2.3001034	KIAA1107
RAB39A	-2.2999814	RAB39A, member RAS oncogene family
CCDC42	-2.298699	coiled-coil domain containing 42
PLAC1	-2.2970617	placenta-specific 1
GC	-2.2873378	group-specific component (vitamin D binding protein)
PRR23C	-2.2872024	proline rich 23C
MED15P3	-2.28449	Mediator Complex Subunit 15 Pseudogene 3
PRR34	-2.2828898	proline rich 34
FAM153A	-2.2825751	family with sequence similarity 153, member A
DDX25	-2.2814333	DEAD (Asp-Glu-Ala-Asp) box helicase 25
CD101	-2.2796173	CD101 molecule
GPR32P1	-2.2751806	G Protein-Coupled Receptor 32, Pseudogene 1
lnc-DIO2-5	-2.2729633	lnc-DIO2-5:1
CHRNA2	-2.27244	cholinergic receptor, nicotinic, alpha 2 (neuronal)
DSG1	-2.27123	desmoglein 1
SYNPO2	-2.2708611	synaptopodin 2
SLC5A7	-2.2665963	solute carrier family 5 (sodium/choline cotransporter), member 7
ZNF385B	-2.2659285	zinc finger protein 385B
LINC01136	-2.2630522	long intergenic non-protein coding RNA 1136
OR9M1P	-2.26231	Olfactory Receptor Family 9 Subfamily M Member 1, Pseudogene
DCANP1	-2.2616718	dendritic cell-associated nuclear protein

ZG16	-2.26131	zymogen granule protein 16
CCDC30	-2.2567255	coiled-coil domain containing 30
MORN5	-2.255523	MORN repeat containing 5
SCGB1D2	-2.2536964	secretoglobin, family 1D, member 2
LRIT2	-2.2531989	leucine-rich repeat, immunoglobulin-like and transmembrane domains 2
BTBD17	-2.2504578	BTB (POZ) domain containing 17
NR2F2	-2.2502139	nuclear receptor subfamily 2, group F, member 2
TRGV10	-2.2498207	T Cell Receptor Gamma Variable 10 (Non-Functional)
CAB39L	-2.249247	calcium binding protein 39-like
MARCKSL1P1	-2.2479997	MARCKS Like 1 Pseudogene 1
SLC9A2	-2.247657	solute carrier family 9, subfamily A (NHE2, cation proton antiporter 2), member 2
WNT2B	-2.2416053	wingless-type MMTV integration site family, member 2B
lnc-GRHL2-9	-2.2413034	lnc-GRHL2-9:1
KSR1	-2.2391357	kinase suppressor of ras 1
GAB3	-2.2344475	GRB2-associated binding protein 3
LEF1	-2.2321026	lymphoid enhancer-binding factor 1
lnc-RP11-351M8.1.1-1	-2.2316155	lnc-RP11-351M8.1.1-1:1
NKX2-4	-2.2271833	NK2 homeobox 4
CXorf57	-2.2264786	chromosome X open reading frame 57
DPP4	-2.2226737	dipeptidyl-peptidase 4
LOC101930611	-2.2218208	uncharacterized LOC101930611
PCDHGB3	-2.2205837	protocadherin gamma subfamily B, 3
PKD1L2	-2.2204745	polycystic kidney disease 1-like 2 (gene/pseudogene)
ARHGEF37	-2.2194898	Rho guanine nucleotide exchange factor (GEF) 37
TRGV9	-2.2183743	T Cell Receptor Gamma Variable 9
LOC100129973	-2.2170699	uncharacterized LOC100129973
SLCO6A1	-2.2122095	solute carrier organic anion transporter family, member 6A1
LOC440173	-2.2120786	uncharacterized LOC440173
DIAPH2	-2.210474	diaphanous-related formin 2

FAM114A1	-2.2103674	family with sequence similarity 114, member A1
UTRN	-2.2062137	utrophin
OR2J4P	-2.2051702	Olfactory Receptor Family 2 Subfamily J Member 4 Pseudogene
GEMIN8P4	-2.200502	gem (nuclear organelle) associated protein 8 pseudogene 4
CYP8B1	-2.199091	cytochrome P450, family 8, subfamily B, polypeptide 1
ZC3H6	-2.1959553	zinc finger CCCH-type containing 6
RIMKLA	-2.194034	ribosomal modification protein rimK-like family member A
ZBTB20	-2.1899486	zinc finger and BTB domain containing 20
EOGT	-2.189835	EGF domain-specific O-linked N-acetylglucosamine (GlcNAc) transferase
HEY2	-2.1865172	hes-related family bHLH transcription factor with YRPW motif 2
LOC100506136	-2.1839283	uncharacterized LOC100506136
AKR1C1	-2.1803653	aldo-keto reductase family 1, member C1
RHCE	-2.1789634	Rh blood group, CcEe antigens
SAG	-2.1788225	S-antigen; retina and pineal gland (arrestin)
AC105345.1	-2.1786578	AC105345.1
CCDC159	-2.1776695	coiled-coil domain containing 159
SPANXN4	-2.1774166	SPANX Family Member N4
DBX1	-2.1754782	developing brain homeobox 1
FAM53C	-2.1747234	Family With Sequence Similarity 53 Member C
SNORA26	-2.1734424	small nucleolar RNA, H/ACA box 26
EFHC2	-2.171055	EF-hand domain (C-terminal) containing 2
lnc-INTS9-1	-2.1683593	lnc-INTS9-1:3
RHOT1	-2.1664162	Ras Homolog Family Member T1
ID1	-2.1653028	inhibitor of DNA binding 1, dominant negative helix-loop-helix protein
TMSB4Y	-2.163002	thymosin beta 4, Y-linked
LOC101927345	-2.161723	uncharacterized LOC101927345
ARAP2	-2.1614745	ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 2
PFKFB4	-2.1595592	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 4
TDRD5	-2.157759	tudor domain containing 5

FAM134B	-2.155006	family with sequence similarity 134, member B
LOC653602	-2.1537354	uncharacterized LOC653602
C18orf65	-2.1526196	chromosome 18 open reading frame 65
C3orf79	-2.152419	chromosome 3 open reading frame 79
LOC646522	-2.151052	uncharacterized LOC646522
LRRC15	-2.1456609	leucine rich repeat containing 15
lnc-C17orf80-3	-2.145608	lnc-C17orf80-3:1
SCOC-AS1	-2.1434355	SCOC antisense RNA 1
IFI44L	-2.1428895	interferon-induced protein 44-like
AL138828.1	-2.1388829	ENSG00000237596
DCDC2C	-2.1373134	doublecortin domain containing 2C
PYDC2	-2.1367388	pyrin domain containing 2
DOCK4	-2.134058	dedicator of cytokinesis 4
CCNT2-AS1	-2.1325383	CCNT2 antisense RNA 1
SCNN1B	-2.1320093	sodium channel, non voltage gated 1 beta subunit
LOC101928424	-2.1317048	uncharacterized LOC101928424
C15orf48	-2.1308193	chromosome 15 open reading frame 48
LCA5	-2.130489	Leber congenital amaurosis 5
GLYATL2	-2.130282	glycine-N-acyltransferase-like 2
RPGRIP1	-2.1262658	retinitis pigmentosa GTPase regulator interacting protein 1
LOC100129900	-2.1254933	P008-1
SH2D1A	-2.1243927	SH2 domain containing 1A
GAB4	-2.123146	GRB2-associated binding protein family, member 4
GBP3	-2.1214976	guanylate binding protein 3
LINC00943	-2.1183646	long intergenic non-protein coding RNA 943
AC138915.2	-2.118059	Solute Carrier Family 9 Member B1 Pseudogene 5
BTN3A1	-2.1178076	butyrophilin, subfamily 3, member A1
LOC102724484	-2.1171832	uncharacterized LOC102724484
ZMAT1	-2.1136978	zinc finger, matrin-type 1



CFHR1	-2.1109679	complement factor H-related 1
CASP4	-2.11034	caspase 4, apoptosis-related cysteine peptidase
TEX13B	-2.1075878	testis expressed 13B
LOC729867	-2.1067257	uncharacterized LOC729867
CPEB1-AS1	-2.1063194	CPEB1 antisense RNA 1
ZBBX	-2.1053224	zinc finger, B-box domain containing
ACRV1	-2.1011152	acrosomal vesicle protein 1
SCEL	-2.0999708	sciellin
SPATA19	-2.099343	spermatogenesis associated 19
OR5M13P	-2.0991492	Olfactory Receptor Family 5 Subfamily M Member 13 Pseudogene
KRT86	-2.0977058	keratin 86, type II
CDKL5	-2.0971935	cyclin-dependent kinase-like 5
LOC100128001	-2.0962875	uncharacterized LOC100128001
CCDC83	-2.0952272	coiled-coil domain containing 83
PTGS1	-2.0950146	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
PPP1R14D	-2.0944214	protein phosphatase 1, regulatory (inhibitor) subunit 14D
AKR1C4	-2.092993	aldo-keto reductase family 1, member C4
ZNF385C	-2.0908427	zinc finger protein 385C
ZNF608	-2.08972	zinc finger protein 608
FBXL22	-2.0895307	F-box and leucine-rich repeat protein 22
NR2F2-AS1	-2.0889757	NR2F2 antisense RNA 1
SVOP	-2.0885773	SV2 related protein homolog (rat)
RNFT1	-2.0870488	ring finger protein, transmembrane 1
AC005524.1	-2.0860622	ENSG00000283269
SERPINB9	-2.0850844	serpin peptidase inhibitor, clade B (ovalbumin), member 9
RAET1E	-2.0841994	retinoic acid early transcript 1E
TMEM8B	-2.0819373	transmembrane protein 8B
ELF5	-2.0801022	E74-like factor 5 (ets domain transcription factor)
PRSS56	-2.0798163	protease, serine, 56

DAO	-2.0793798	D-amino-acid oxidase
KCNJ13	-2.0793736	potassium channel, inwardly rectifying subfamily J, member 13
AC073544.1	-2.0793166	Zinc Finger Protein 430 Pseudogene
TNIK	-2.0785332	TRAF2 and NCK interacting kinase
MDGA2	-2.0772984	MAM domain containing glycosylphosphatidylinositol anchor 2
LIMA1	-2.0772614	LIM domain and actin binding 1
KRT24	-2.076892	keratin 24, type I
KRTAP29-1	-2.0750272	keratin associated protein 29-1
GVINP1	-2.0740194	GTPase, very large interferon inducible pseudogene 1
LOC100049716	-2.071537	uncharacterized LOC100049716
LINC00116	-2.0702412	long intergenic non-protein coding RNA 116
GPC5	-2.0702345	glypican 5
LINC01565	-2.069532	long intergenic non-protein coding RNA 1565
CEP112	-2.067293	centrosomal protein 112kDa
EPB41L2	-2.0671673	erythrocyte membrane protein band 4.1-like 2
EPHA4	-2.0667682	EPH receptor A4
GATM	-2.061233	glycine amidinotransferase (L-arginine:glycine amidinotransferase)
GANC	-2.0604727	glucosidase, alpha; neutral C
OR5T3	-2.0593913	olfactory receptor, family 5, subfamily T, member 3
USP40	-2.057611	ubiquitin specific peptidase 40
CCDC171	-2.0568116	coiled-coil domain containing 171
LOC728752	-2.055559	uncharacterized LOC728752
CAMTA1	-2.0543704	calmodulin binding transcription activator 1
SPATA4	-2.0539405	spermatogenesis associated 4
SMC1B	-2.0536504	structural maintenance of chromosomes 1B
BNIP3	-2.0536153	BCL2/adenovirus E1B 19kDa interacting protein 3
lnc-ACOT12-2	-2.0535836	lnc-ACOT12-2:1
XAGE2	-2.0535572	X antigen family, member 2
IGHE	-2.052961	Immunoglobulin Heavy Constant Epsilon

TMPRSS11F	-2.0522578	transmembrane protease, serine 11F
MECOM	-2.0522401	MDS1 and EVI1 complex locus
LOC100130849	-2.049444	phosphorylase kinase, gamma 1 (muscle) pseudogene
LPAR5	-2.0494196	lysophosphatidic acid receptor 5
CCDC175	-2.0491514	coiled-coil domain containing 175
DNAJB8	-2.0482128	DnaJ (Hsp40) homolog, subfamily B, member 8
LOC400682	-2.048133	zinc finger protein 100-like
CARD6	-2.0475347	caspase recruitment domain family, member 6
SLC19A2	-2.0453782	solute carrier family 19 (thiamine transporter), member 2
LOC100129098	-2.045084	WAS/WASL-interacting protein family member 1-like
SYT8	-2.0420125	synaptotagmin VIII
FOXP2	-2.0412843	forkhead box P2
CYP26A1	-2.0407674	cytochrome P450, family 26, subfamily A, polypeptide 1
CFHR4	-2.0407655	complement factor H-related 4
HESX1	-2.0396428	HESX homeobox 1
PDYN	-2.038495	prodynorphin
LHFPL3-AS1	-2.037224	LHFPL3 antisense RNA 1
FSD1L	-2.036571	fibronectin type III and SPRY domain containing 1-like
LOC388692	-2.0364928	uncharacterized LOC388692
FGF13	-2.036028	fibroblast growth factor 13
CHGB	-2.0358768	chromogranin B (secretogranin 1)
FKBP7	-2.034468	FK506 binding protein 7
DPY19L2	-2.0342727	dpy-19-like 2 (C. elegans)
TFF2	-2.0340376	trefoil factor 2
TCP11L2	-2.0317914	t-complex 11, testis-specific-like 2
FAM183B	-2.031095	acyloxyacyl hydrolase (neutrophil)
SLC25A27	-2.0301917	solute carrier family 25, member 27
HMCN1	-2.0298898	hemicentin 1
LOC100499484	-2.029861	SUGT1-1300002K09Rik pseudogene

GOLM1	-2.0296898	golgi membrane protein 1
PDE7A	-2.0292478	phosphodiesterase 7A
MERTK	-2.0256848	MER proto-oncogene, tyrosine kinase
PCDH11Y	-2.024386	protocadherin 11 Y-linked
MAGEA4	-2.0242004	melanoma antigen family A, 4
lnc-BMP6-10	-2.0236595	lnc-BMP6-10:2
KIAA1731NL	-2.0225582	KIAA1731 N-terminal like
lnc-BCL9L-1	-2.020285	lnc-BCL9L-1:1
HYDIN	-2.0188794	HYDIN, Axonemal Central Pair Apparatus Protein
GSTA3	-2.0176287	glutathione S-transferase alpha 3
LRRC63	-2.0170274	leucine rich repeat containing 63
MPZ	-2.0133653	myelin protein zero
TNFRSF9	-2.0132413	tumor necrosis factor receptor superfamily, member 9
PASD1	-2.0111144	PAS domain containing 1
RAB19	-2.009727	RAB19, member RAS oncogene family
TTYH1	-2.009539	tweety family member 1
SLC16A2	-2.0092819	solute carrier family 16, member 2 (thyroid hormone transporter)
lnc-SART3-1	-2.0090303	lnc-SART3-1:1
LL22NC01-81G9.3	-2.0070598	uncharacterized protein FLJ39582-like
LOC339685	-2.006569	uncharacterized LOC339685
SPINT5P	-2.004472	Serine Peptidase Inhibitor, Kunitz Type 5, Pseudogene
WIPF1	-2.0011406	WAS/WASL interacting protein family, member 1
AC020907.1	-2.0008237	ENSG00000179066
MLN	-2.0003953	motilin
CMA1	2.0026698	chymase 1, mast cell
LOC100130417	2.003604	uncharacterized LOC100130417
DIO1	2.0054278	deiodinase, iodothyronine, type I
ALX1	2.005629	ALX homeobox 1
SULT6B1	2.0060012	sulfotransferase family, cytosolic, 6B, member 1

ABCG1	2.0096438	ATP-binding cassette, sub-family G (WHITE), member 1
EMX2OS	2.0102398	EMX2 opposite strand/antisense RNA
SOX15	2.0106773	SRY (sex determining region Y)-box 15
CATSPER1	2.0114899	cation channel, sperm associated 1
LOC100129322	2.0121307	uncharacterized LOC100129322
LOC653581	2.0121307	uncharacterized LOC653581
PSG5	2.0128365	pregnancy specific beta-1-glycoprotein 5
FLG2	2.013536	filaggrin family member 2
CR2	2.0142176	complement component (3d/Epstein Barr virus) receptor 2
RBPMS2	2.0144074	RNA binding protein with multiple splicing 2
LOC403312	2.0146835	putative uncharacterized protein MGC39545
CD1D	2.0146897	CD1d molecule
MMP13	2.016876	matrix metalloproteinase 13 (collagenase 3)
AC133919.3	2.019001	Capicua Transcriptional Repressor Pseudogene 25
RBM38	2.0192475	RNA binding motif protein 38
LOC101060157	2.0199258	uncharacterized LOC101060157
PDE4A	2.0218012	phosphodiesterase 4A, cAMP-specific
CLCN2	2.0220144	chloride channel, voltage-sensitive 2
PRAMEF4	2.023252	PRAME family member 4
ZNF556	2.0256019	zinc finger protein 556
HOMER1	2.025751	homer homolog 1 (Drosophila)
LL22NC03-75H12.2	2.0273976	uncharacterized LOC101927722
HSPA5	2.0292647	heat shock 70kDa protein 5 (glucose-regulated protein, 78kDa)
GPR162	2.0340734	G protein-coupled receptor 162
LINC01018	2.0345361	long intergenic non-protein coding RNA 1018
lnc-FMNL2-2	2.0351572	lnc-FMNL2-2:1
RNF186	2.035799	ring finger protein 186
RGS8	2.0364842	regulator of G-protein signaling 8
FAM27L	2.038072	family with sequence similarity 27-like

SMAD2	2.038735	SMAD family member 2
IL27RA	2.04133	interleukin 27 receptor, alpha
TPTE2P5	2.0414226	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 5
LINC00587	2.041582	long intergenic non-protein coding RNA 587
HMHA1	2.0433297	histocompatibility (minor) HA-1
RNASE8	2.044617	ribonuclease, RNase A family, 8
XLOC_12_011265	2.0455163	XLOC_12_011265
LYPLAL1-AS1	2.0471637	LYPLAL1 antisense RNA 1 (head to head)
TEC	2.0480711	tec protein tyrosine kinase
WTH3DI	2.0487106	RAB6C-like
TYW3	2.0498822	tRNA-yW synthesizing protein 3 homolog ( <i>S. cerevisiae</i> )
LOC101929256	2.0507653	uncharacterized LOC101929256
OR2G2	2.0510457	olfactory receptor, family 2, subfamily G, member 2
ARSE	2.0516338	arylsulfatase E (chondrodysplasia punctata 1)
AC103563.3	2.053849	ENSG00000231331
APOL5	2.0564404	apolipoprotein L, 5
CCDC88B	2.056901	Coiled-Coil Domain Containing 88B
SEPT3	2.059271	septin 3
LOC100130370	2.0604162	uncharacterized LOC100130370
BMPER	2.0609262	BMP binding endothelial regulator
TMEM52	2.0626042	transmembrane protein 52
C12orf71	2.0636206	chromosome 12 open reading frame 71
LINC00636	2.064627	long intergenic non-protein coding RNA 636
CHST1	2.0650167	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1
SIM2	2.0651157	single-minded family bHLH transcription factor 2
LOC101928259	2.065749	uncharacterized LOC101928259
C1QL2	2.0674286	complement component 1, q subcomponent-like 2
RIMBP3	2.0681062	RIMS binding protein 3
LINC00961	2.0691564	long intergenic non-protein coding RNA 961

JSRP1	2.0696719	junctional sarcoplasmic reticulum protein 1
RSPH6A	2.0703928	radial spoke head 6 homolog A (Chlamydomonas)
ABHD16B	2.0734308	abhydrolase domain containing 16B
STC2	2.0737312	stanniocalcin 2
LINC02054	2.0747328	Long Intergenic Non-Protein Coding RNA 2054
GPX2	2.0752795	glutathione peroxidase 2 (gastrointestinal)
KCNJ6	2.0765245	potassium channel, inwardly rectifying subfamily J, member 6
LOC102723367	2.0772305	uncharacterized LOC102723367
AC010601.1	2.0808012	ENSG00000259929
C20orf173	2.081429	chromosome 20 open reading frame 173
QRSL1	2.0829375	glutaminyl-tRNA synthase (glutamine-hydrolyzing)-like 1
MUC16	2.0854402	mucin 16, cell surface associated
KRT18P46	2.085656	Keratin 18 Pseudogene 46
lnc-KBTBD5-2	2.086837	lnc-KBTBD5-2:2
EFR3B	2.0879989	EFR3 homolog B ( <i>S. cerevisiae</i> )
PRO2958	2.0880976	uncharacterized LOC100128329
CLMP	2.0884545	CXADR-like membrane protein
POLR3E	2.0885766	polymerase (RNA) III (DNA directed) polypeptide E (80kD)
ARRDC5	2.0904121	arrestin domain containing 5
THG1L	2.092266	tRNA-histidine guanylyltransferase 1-like ( <i>S. cerevisiae</i> )
SPTLC1	2.093359	serine palmitoyltransferase, long chain base subunit 1
FENDRR	2.095391	FOXF1 adjacent non-coding developmental regulatory RNA
RBPJL	2.100045	recombination signal binding protein for immunoglobulin kappa J region-like
EFCC1	2.101288	EF-hand and coiled-coil domain containing 1
ITFG1-AS1	2.1022878	ITFG1 antisense RNA 1
C10orf113	2.1028519	chromosome 10 open reading frame 113
HSPA12A	2.102858	heat shock 70kDa protein 12A
AL354984.1	2.1057603	ENSG00000203971
LINC00928	2.1058078	long intergenic non-protein coding RNA 928

IGLV4-69	2.106601	Immunoglobulin Lambda Variable 4-69
SUN3	2.1068633	Sad1 and UNC84 domain containing 3
ATE1	2.1104755	arginyltransferase 1
CLEC4A	2.1117175	C-type lectin domain family 4, member A
PROCR	2.112896	protein C receptor, endothelial
AP003532.1	2.1148198	ENSG00000254591
KHDRBS2	2.1164927	KH domain containing, RNA binding, signal transduction associated 2
SUSD2	2.1173687	sushi domain containing 2
NMNAT2	2.117993	nicotinamide nucleotide adenyltransferase 2
FAM71D	2.1185145	family with sequence similarity 71, member D
UROCI	2.118572	urocanate hydratase 1
ANKRD1	2.1203427	ankyrin repeat domain 1 (cardiac muscle)
AC092376.1	2.1222975	ENSG00000261722
GALNT6	2.12269	polypeptide N-acetylgalactosaminyltransferase 6
LGR6	2.1236591	Leucine Rich Repeat Containing G Protein-Coupled Receptor 6
SPOPL	2.124439	speckle-type POZ protein-like
RSU1P2	2.1265194	Ras Suppressor Protein 1 Pseudogene 2
Inc-POLR2B-1	2.127944	Inc-POLR2B-1:6
TAS2R16	2.1279643	taste receptor, type 2, member 16
SLC18A2	2.135613	solute carrier family 18 (vesicular monoamine transporter), member 2
TTY12	2.1357312	testis-specific transcript, Y-linked 12 (non-protein coding)
DNMT1	2.1365294	DNA (cytosine-5-)-methyltransferase 1
C12orf50	2.137667	chromosome 12 open reading frame 50
LOC101929983	2.1381125	PRAME family member 9/15-like
LMOD2	2.138213	leiomodrin 2 (cardiac)
FGF13-AS1	2.1385088	FGF13 antisense RNA 1
LSP1P4	2.1387422	LSP1 pseudogene 4
KCNMB2	2.1396081	potassium channel subfamily M regulatory beta subunit 2
CRMP1	2.1400263	collapsin response mediator protein 1



KLRF1	2.1405075	killer cell lectin-like receptor subfamily F, member 1
ABO	2.1425943	ABO blood group (transferase A, alpha 1-3-N-acetylgalactosaminyltransferase; transferase B, alpha 1-3-galactosyltransferase)
PACS2	2.144385	phosphofurin acidic cluster sorting protein 2
ZNF528	2.1454134	zinc finger protein 528
lnc-WDR1-1	2.1467495	lnc-WDR1-1:2
KIF5C	2.1472135	kinesin family member 5C
NANP	2.1482003	N-acetylneuraminic acid phosphatase
CES4A	2.1486518	carboxylesterase 4A
C7orf31	2.1510017	chromosome 7 open reading frame 31
OR52E4	2.1522288	olfactory receptor, family 52, subfamily E, member 4
LINC01387	2.153437	long intergenic non-protein coding RNA 1387
MMP9	2.1543064	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)
AC245177.1	2.1546817	ENSG00000179028
SYCP2L	2.1549795	synaptonemal complex protein 2-like
MIR4435-2HG	2.1551373	MIR4435-2 Host Gene
TBX18	2.1562092	T-box 18
EHD3	2.1577976	EH-domain containing 3
POU3F2	2.1597662	POU class 3 homeobox 2
CCR1	2.1630352	chemokine (C-C motif) receptor 1
CAMK4	2.16412	calcium/calmodulin-dependent protein kinase IV
FAM95B1	2.1652844	Family With Sequence Similarity 95 Member B1
BPIFB6	2.1657777	BPI fold containing family B, member 6
KHK	2.1686044	ketoheokinase (fructokinase)
GP2	2.1706214	glycoprotein 2 (zymogen granule membrane)
VWA3B	2.172986	Von Willebrand Factor A Domain Containing 3B
PRR16	2.1743631	proline rich 16
OTOL1	2.177846	otolin 1
RBFOX1	2.1810198	RNA binding protein, fox-1 homolog (C. elegans) 1

USP6	2.1829119	ubiquitin specific peptidase 6
LOC100507477	2.18294	uncharacterized LOC100507477
MS4A8	2.1861427	membrane-spanning 4-domains, subfamily A, member 8
IPW	2.1883228	imprinted in Prader-Willi syndrome (non-protein coding)
HOXD11	2.1899595	homeobox D11
PXYLP1	2.190882	2-phosphoxylose phosphatase 1
NECAB2	2.1926785	N-terminal EF-hand calcium binding protein 2
PAGE1	2.1934347	P antigen family, member 1 (prostate associated)
DKK1	2.1944706	dickkopf WNT signaling pathway inhibitor 1
lnc-C5orf42-2	2.199013	lnc-C5orf42-2:1
KRTAP5-8	2.2025204	keratin associated protein 5-8
PNMA6A	2.2045634	paraneoplastic Ma antigen family member 6A
LUZP6	2.206725	leucine zipper protein 6
IBSP	2.2070994	integrin-binding sialoprotein
ZNF469	2.2073762	zinc finger protein 469
LOC101928307	2.2104259	uncharacterized LOC101928307
REEP2	2.2130547	receptor accessory protein 2
LMO7DN	2.2139325	LMO7 downstream neighbor
DEFB126	2.214954	defensin, beta 126
ANXA10	2.2154438	annexin A10
LTB	2.215605	lymphotoxin beta (TNF superfamily, member 3)
PCDHGC3	2.2162292	protocadherin gamma subfamily C, 3
CREM	2.2179036	cAMP responsive element modulator
PKD1L3	2.2206604	polycystic kidney disease 1-like 3
C2CD4C	2.221196	C2 calcium-dependent domain containing 4C
CYP26B1	2.225383	cytochrome P450, family 26, subfamily B, polypeptide 1
MTPN	2.225995	myotrophin
UCN2	2.2266338	urocortin 2
GP5	2.2266889	glycoprotein V (platelet)

C9orf171	2.227039	chromosome 9 open reading frame 171
FAM230A	2.2275848	family with sequence similarity 230, member A
LOC102723428	2.2283447	uncharacterized LOC102723428
GABRA1	2.2290423	gamma-aminobutyric acid (GABA) A receptor, alpha 1
PXT1	2.229243	peroxisomal, testis specific 1
GRID1	2.2310371	glutamate receptor, ionotropic, delta 1
OR8J3	2.2323005	olfactory receptor, family 8, subfamily J, member 3
KRTAP4-6	2.232969	keratin associated protein 4-6
VWC2L	2.2350984	von Willebrand factor C domain containing protein 2-like
TIMP3	2.2361803	TIMP metalloproteinase inhibitor 3
PMP22	2.2393243	Peripheral Myelin Protein 22
MAP1B	2.2409215	microtubule-associated protein 1B
ZNF510	2.2416155	zinc finger protein 510
BICD1	2.251384	bicaudal D homolog 1 (Drosophila)
LOC389895	2.2521865	chromosome 16 open reading frame 72-like
NRXN2	2.2552783	neurexin 2
LOC644554	2.256851	uncharacterized LOC644554
lnc-MGAT1-1	2.2600577	lnc-MGAT1-1:1
BSX	2.2659612	brain-specific homeobox
lnc-SFTA2-3	2.2666779	lnc-SFTA2-3:1
RSPO1	2.2689	R-spondin 1
PNLIPRP2	2.2697227	pancreatic lipase-related protein 2
ROM1	2.2706683	retinal outer segment membrane protein 1
OR2AJ1	2.2708023	Olfactory Receptor Family 2 Subfamily AJ Member 1
FLRT1	2.2731624	fibronectin leucine rich transmembrane protein 1
NT5DC2	2.2757065	5'-nucleotidase domain containing 2
ADCY5	2.2826872	adenylate cyclase 5
FAM102B	2.282945	family with sequence similarity 102, member B
ZFAT-AS1	2.2831697	ZFAT antisense RNA 1

S100A14	2.283446	S100 calcium binding protein A14
ANO4	2.295878	anoctamin 4
OR5D14	2.2959998	olfactory receptor, family 5, subfamily D, member 14
SPATA3-AS1	2.2970345	SPATA3 antisense RNA 1 (head to head)
CHST11	2.2981315	carbohydrate (chondroitin 4) sulfotransferase 11
PNCK	2.2990637	pregnancy up-regulated nonubiquitous CaM kinase
KL	2.300797	klotho
NCKAP5	2.3071089	NCK-associated protein 5
TOP1MT	2.3072004	topoisomerase (DNA) I, mitochondrial
SRD5A2	2.3082194	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)
GPRC5B	2.31141	G protein-coupled receptor, class C, group 5, member B
KIF1A	2.3180897	kinesin family member 1A
FLT3	2.3226464	fms-related tyrosine kinase 3
GAL	2.324516	galanin/GMAP prepropeptide
ELAVL4	2.3249402	ELAV like neuron-specific RNA binding protein 4
PTBP3	2.327853	polypyrimidine tract binding protein 3
TEX36	2.328838	testis expressed 36
PCDH11X	2.3291488	protocadherin 11 X-linked
GMNC	2.3329053	geminin coiled-coil domain containing
GNB1	2.3335688	guanine nucleotide binding protein (G protein), beta polypeptide 1
ANKRD26P3	2.3351762	ankyrin repeat domain 26 pseudogene 3
LRG1	2.3365355	leucine-rich alpha-2-glycoprotein 1
C1orf87	2.336948	chromosome 1 open reading frame 87
FAM47C	2.3391237	family with sequence similarity 47, member C
RPGRIP1L	2.3403993	RPGRIP1-like
LOC100129617	2.3407204	uncharacterized LOC100129617
CADM4	2.3425558	cell adhesion molecule 4
UOX	2.3431818	urate oxidase, pseudogene
SOCS4	2.34477	suppressor of cytokine signaling 4

AC243960.3	2.3484504	ENSG00000268355
PRSS35	2.3490057	protease, serine, 35
AP001341.1	2.3504252	ENSG00000222042
GPR82	2.3504531	G protein-coupled receptor 82
XLOC_12_009136	2.3523579	XLOC_12_009136
FAM224A	2.354712	family with sequence similarity 224, member A (non-protein coding)
TMOD2	2.3560264	tropomodulin 2 (neuronal)
SEPT9	2.356624	Septin 9
ATP1B1	2.3604069	ATPase, Na <sup>+</sup> /K <sup>+</sup> transporting, beta 1 polypeptide
LOC400541	2.361697	uncharacterized LOC400541
ABCC3	2.3617663	ATP-binding cassette, sub-family C (CFTR/MRP), member 3
LOC645752	2.3678148	golgin A6 family, member A pseudogene
IGFBP2	2.369269	insulin-like growth factor binding protein 2, 36kDa
KIF12	2.370345	kinesin family member 12
DUX4L45	2.3763254	Double Homeobox 4 Like 45 (Pseudogene)
IGKV2D-24	2.3810787	Immunoglobulin Kappa Variable 2D-24 (Non-Functional)
HHIPL2	2.381656	HHIP-like 2
ADAM32	2.390578	ADAM metallopeptidase domain 32
ADRA2B	2.392629	adrenoceptor alpha 2B
B4GALNT1	2.3946254	beta-1,4-N-acetyl-galactosaminyl transferase 1
LOC100128063	2.3984218	uncharacterized LOC100128063
OR10A7	2.4031937	olfactory receptor, family 10, subfamily A, member 7
LRRC2-AS1	2.40521	LRRC2 antisense RNA 1
TREM2	2.4056334	triggering receptor expressed on myeloid cells 2
LOC339975	2.4062605	uncharacterized LOC339975
ZBTB8B	2.406318	zinc finger and BTB domain containing 8B
SPATA31D3	2.410885	SPATA31 subfamily D, member 3
DISC1	2.4117823	disrupted in schizophrenia 1
CA10	2.413012	carbonic anhydrase X

GDF11	2.4174278	growth differentiation factor 11
CYP3A5	2.4178944	cytochrome P450, family 3, subfamily A, polypeptide 5
DMRT3	2.4210348	doublesex and mab-3 related transcription factor 3
ALDH8A1	2.4213111	aldehyde dehydrogenase 8 family, member A1
lnc-SLC29A4-1	2.4234056	lnc-SLC29A4-1:1
MS4A6E	2.4245882	membrane-spanning 4-domains, subfamily A, member 6E
ARL10	2.4291015	ADP-ribosylation factor-like 10
lnc-PDZD7-4	2.4295208	lnc-PDZD7-4:3
CHPF	2.4333563	chondroitin polymerizing factor
SCAMP1	2.434801	secretory carrier membrane protein 1
INGX	2.4370558	inhibitor of growth family, X-linked, pseudogene
CYB5R4	2.439856	cytochrome b5 reductase 4
ERAS	2.4400206	ES cell expressed Ras
HIGD2B	2.4457924	HIG1 hypoxia inducible domain family, member 2B
SLC25A48	2.4468052	solute carrier family 25, member 48
CCT5	2.4522703	chaperonin containing TCP1, subunit 5 (epsilon)
C1orf145	2.4544528	chromosome 1 open reading frame 145
KIAA1524	2.455378	KIAA1524
COPG2IT1	2.4596655	COPG2 imprinted transcript 1 (non-protein coding)
FGF19	2.461473	fibroblast growth factor 19
TDRD12	2.4639251	tudor domain containing 12
IGHV1-58	2.4654121	Immunoglobulin Heavy Variable 1-58
IGHV3OR16-13	2.4671266	Immunoglobulin Heavy Variable 3/OR16-13 (Non-Functional)
FGF17	2.467152	fibroblast growth factor 17
LINC00208	2.4736137	long intergenic non-protein coding RNA 208
PCDH15	2.4895597	protocadherin-related 15
OR10S1	2.492242	olfactory receptor, family 10, subfamily S, member 1
GJE1	2.4991117	Gap Junction Protein Epsilon 1
TMEM100	2.499928	transmembrane protein 100

CPNE7	2.5013273	copine VII
LOXL3	2.5017722	lysyl oxidase-like 3
KLHDC8A	2.502091	kelch domain containing 8A
AADACP1	2.5055757	arylacetamide deacetylase pseudogene 1
CRYGB	2.5056334	crystallin, gamma B
lnc-TRIM29-1	2.5058002	lnc-TRIM29-1:2
lnc-GUSB-5	2.5106711	lnc-GUSB-5:1
lnc-RP11-712L6.5.1-2	2.5107296	lnc-RP11-712L6.5.1-2:1
KRT6C	2.5126727	keratin 6C, type II
LOC388882	2.5163863	uncharacterized LOC388882
LOC642943	2.5172894	SPATA31 subfamily A pseudogene
KCP	2.5175488	Kielin/Chordin-Like Protein
USF1	2.5222697	upstream transcription factor 1
LY6D	2.523226	lymphocyte antigen 6 complex, locus D
LOC100131180	2.5261338	uncharacterized LOC100131180
RBM15B	2.5263112	RNA binding motif protein 15B
lnc-FBRSL1-4	2.5286815	lnc-FBRSL1-4:2
SLC22A4	2.538493	solute carrier family 22 (organic cation/zwitterion transporter), member 4
ARHGAP26	2.5427175	Rho GTPase activating protein 26
NTN1	2.5442114	netrin 1
GCSAML	2.5454612	germinal center-associated, signaling and motility-like
SLC26A2	2.5507267	solute carrier family 26 (anion exchanger), member 2
KRT1	2.5519822	keratin 1, type II
DNAJC7	2.553171	DnaJ (Hsp40) homolog, subfamily C, member 7
SH3PXD2B	2.5656044	SH3 and PX domains 2B
TMEM71	2.5678475	transmembrane protein 71
BNIP3P41	2.5713983	BCL2 Interacting Protein 3 Pseudogene 41
WEE2	2.575728	WEE1 homolog 2 ( <i>S. pombe</i> )
TGFB1	2.5760267	transforming growth factor, beta 1

LOC284191	2.5810828	uncharacterized LOC284191
LRRC4C	2.582104	leucine rich repeat containing 4C
TAS2R64P	2.5836565	Taste 2 Receptor Member 64 Pseudogene
ELL3	2.5870457	elongation factor RNA polymerase II-like 3
SOAT2	2.5910513	sterol O-acyltransferase 2
TSPY26P	2.5915108	testis specific protein, Y-linked 26, pseudogene
PSG10P	2.594838	pregnancy specific beta-1-glycoprotein 10, pseudogene
DGKE	2.6053314	diacylglycerol kinase, epsilon 64kDa
SOWAHB	2.605907	sosondowah ankyrin repeat domain family member B
IGHV1-46	2.609493	Immunoglobulin Heavy Variable 1-46
LINC00870	2.6101665	long intergenic non-protein coding RNA 870
SLC41A1	2.6122396	solute carrier family 41 (magnesium transporter), member 1
ART3	2.6141174	ADP-ribosyltransferase 3
lnc-PDE6B-1	2.6206648	lnc-PDE6B-1:8
SCN10A	2.6291199	sodium channel, voltage gated, type X alpha subunit
SAR1B	2.6482036	secretion associated, Ras related GTPase 1B
MAGEB16	2.6508121	melanoma antigen family B, 16
PDE6C	2.652548	phosphodiesterase 6C, cGMP-specific, cone, alpha prime
IL24	2.6553617	interleukin 24
IGSF22	2.6594489	immunoglobulin superfamily, member 22
LOC729950	2.6602106	uncharacterized LOC729950
LRRC55	2.665658	leucine rich repeat containing 55
LOC101930426	2.6683354	double homeobox protein 4-like protein 4-like
UGT2B28	2.6689186	UDP glucuronosyltransferase 2 family, polypeptide B28
CCDC155	2.6776154	coiled-coil domain containing 155
ADH4	2.6808534	alcohol dehydrogenase 4 (class II), pi polypeptide
lnc-WNT8A-2	2.6816318	lnc-WNT8A-2:1
TMEM50B	2.6822236	transmembrane protein 50B
EFNB3	2.6870594	ephrin-B3



OR2AT4	2.6897566	olfactory receptor, family 2, subfamily AT, member 4
CCKBR	2.6915002	cholecystokinin B receptor
C14orf178	2.7020187	chromosome 14 open reading frame 178
ACTG2	2.7104626	actin, gamma 2, smooth muscle, enteric
TMCO5B	2.7110493	transmembrane and coiled-coil domains 5B, pseudogene
LOC100101478	2.7116027	H2B histone family, member M pseudogene
GPR4	2.7214937	G protein-coupled receptor 4
LOC102725453	2.7233675	uncharacterized LOC102725453
OR4F6	2.7320786	olfactory receptor, family 4, subfamily F, member 6
LOC90768	2.7417893	uncharacterized LOC90768
SLC4A8	2.7530103	solute carrier family 4, sodium bicarbonate cotransporter, member 8
IGLV	2.7580826	Immunoglobulin Lambda Variable Cluster
DNAL4	2.7680147	Dynein Axonemal Light Chain 4
CRTAM	2.770724	cytotoxic and regulatory T cell molecule
DIO3OS	2.7742946	DIO3 opposite strand/antisense RNA (head to head)
FAM135A	2.7794063	family with sequence similarity 135, member A
ATP6AP1	2.7868638	ATPase, H <sup>+</sup> transporting, lysosomal accessory protein 1
lnc-C3orf37-1	2.7887564	lnc-C3orf37-1:3
DUSP5	2.791439	dual specificity phosphatase 5
KIAA0509	2.8247924	uncharacterized LOC57242
CCDC67	2.829276	coiled-coil domain containing 67
POU2AF1	2.8366208	POU class 2 associating factor 1
MED28	2.8424103	mediator complex subunit 28
LILRB5	2.8440514	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 5
LDLRAD4	2.849208	low density lipoprotein receptor class A domain containing 4
CCL14	2.8677616	chemokine (C-C motif) ligand 14
TTY5	2.869913	testis-specific transcript, Y-linked 5 (non-protein coding)
EVI2A	2.8714283	ecotropic viral integration site 2A
BFSP2	2.8739474	beaded filament structural protein 2, phakinin

LOC100240728	2.8739607	uncharacterized LOC100240728
LYZL6	2.882558	lysozyme-like 6
LOC613266	2.885565	uncharacterized LOC613266
SCUBE2	2.887476	signal peptide, CUB domain, EGF-like 2
AFAP1L2	2.8898363	actin filament associated protein 1-like 2
lnc-HIBADH-1	2.895715	lnc-HIBADH-1:1
lnc-ABCF2-1	2.8968334	lnc-ABCF2-1:1
KRTAP26-1	2.9088323	keratin associated protein 26-1
SHC4	2.9122944	SHC (Src homology 2 domain containing) family, member 4
GSG1L	2.9224064	GSG1-like
DEFB119	2.922585	defensin, beta 119
FAT2	2.9231904	FAT atypical cadherin 2
AL161938.1	2.9269836	ENSG00000261411
TRPA1	2.9349	transient receptor potential cation channel, subfamily A, member 1
OR6C72P	2.9350586	Olfactory Receptor Family 6 Subfamily C Member 72 Pseudogene
CXCR4	2.96424	chemokine (C-X-C motif) receptor 4
PRR31	2.9664736	proline rich 31
SSTR5-AS1	2.967686	SSTR5 antisense RNA 1
MINDY4B	2.9880085	MINDY Family Member 4B
LINC00304	2.9936476	long intergenic non-protein coding RNA 304
CPHL1P	3.002053	Ceruloplasmin And Hephaestin-Like 1 Pseudogene
NEUROG1	3.0033967	neurogenin 1
LINC00158	3.0051095	long intergenic non-protein coding RNA 158
LINC00598	3.0157802	Long Intergenic Non-Protein Coding RNA 598
BTNL3	3.0198433	butyrophilin-like 3
GPR157	3.0292544	G protein-coupled receptor 157
POTEG	3.033503	POTE Ankyrin Domain Family Member G
CCDC27	3.0371072	coiled-coil domain containing 27
TEX22	3.0568066	testis expressed 22

CCR3	3.063796	chemokine (C-C motif) receptor 3
LINC01896	3.0892072	Long Intergenic Non-Protein Coding RNA 1896
YWHAH	3.0977626	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta
PCMTD1	3.102361	protein-L-isoaspartate (D-aspartate) O-methyltransferase domain containing 1
OR9G9	3.1038601	olfactory receptor, family 9, subfamily G, member 9
TRBV3	3.1074624	T Cell Receptor Beta Variable 3
WDR88	3.1091418	WD repeat domain 88
NTN3	3.1299903	netrin 3
AL731563.3	3.139426	ENSG00000272630
LINC01483	3.1580334	long intergenic non-protein coding RNA 1483
CDH12	3.1711087	cadherin 12, type 2 (N-cadherin 2)
DCAF8L2	3.1899385	DDB1 And CUL4 Associated Factor 8 Like 2
IGKV1D-13	3.2129903	immunoglobulin kappa variable 1D-13
FAM155A	3.2135317	family with sequence similarity 155, member A
lnc-ZNF454-1	3.217579	lnc-ZNF454-1:1
LINC01342	3.218105	long intergenic non-protein coding RNA 1342
TRIM51	3.2204735	tripartite motif-containing 51
SSMEM1	3.2411435	serine-rich single-pass membrane protein 1
CLDN5	3.2413135	claudin 5
OR5P2	3.2599745	olfactory receptor, family 5, subfamily P, member 2
LINC01552	3.2712939	long intergenic non-protein coding RNA 1552
FAM184B	3.2849174	family with sequence similarity 184, member B
LOC284669	3.2850525	uncharacterized LOC284669
PCDHAC2	3.299465	protocadherin alpha subfamily C, 2
MICALCL	3.3260322	MICAL C-terminal like
PPARGC1A	3.3341908	peroxisome proliferator-activated receptor gamma, coactivator 1 alpha
NUDT15	3.337782	nudix (nucleoside diphosphate linked moiety X)-type motif 15
RANBP3L	3.3445399	RAN binding protein 3-like
LHX4	3.351661	LIM homeobox 4

ZFP82	3.3921645	ZFP82 zinc finger protein
HSPA12B	3.4032884	heat shock 70kD protein 12B
LOC100128477	3.4106548	uncharacterized LOC100128477
MYBPC2	3.4146407	myosin binding protein C, fast type
ABCC4	3.4447567	ATP-binding cassette, sub-family C (CFTR/MRP), member 4
FOXL2NB	3.4573803	FOXL2 neighbor
BLACE	3.4814036	B-cell acute lymphoblastic leukemia expressed
HTR1E	3.486106	5-hydroxytryptamine (serotonin) receptor 1E, G protein-coupled
XYLT1	3.5179684	xylosyltransferase I
LOC100127972	3.5257618	uncharacterized LOC100127972
NOX3	3.533292	NADPH oxidase 3
ZNF534	3.5417032	zinc finger protein 534
CHSY3	3.5673075	chondroitin sulfate synthase 3
KCNJ4	3.569996	potassium channel, inwardly rectifying subfamily J, member 4
LOC387720	3.5722103	uncharacterized LOC387720
EMP2	3.5753117	epithelial membrane protein 2
FRRS1L	3.578982	ferric-chelate reductase 1-like
INHBA-AS1	3.6039424	INHBA antisense RNA 1
KCNJ11	3.6050255	potassium channel, inwardly rectifying subfamily J, member 11
GRPR	3.6106408	gastrin-releasing peptide receptor
Inc-OR10AB1P-1	3.632497	Inc-OR10AB1P-1:1
TIFAB	3.6523893	TRAF-interacting protein with forkhead-associated domain, family member B
SUPT20HL2	3.6576176	suppressor of Ty 20 homolog ( <i>S. cerevisiae</i> )-like 2
CYCSP52	3.6641204	cytochrome c, somatic pseudogene 52
HIPK3	3.670444	homeodomain interacting protein kinase 3
NFATC4	3.6864524	nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 4
AC022748.2	3.69952	Golgin A6 Family Member D Pseudogene
LPAR3	3.7035255	lysophosphatidic acid receptor 3
FAM150A	3.7219892	family with sequence similarity 150, member A

FAM65C	3.7630498	family with sequence similarity 65, member C
LINGO3	3.7858822	leucine rich repeat and Ig domain containing 3
FAM182A	3.7911813	family with sequence similarity 182, member A
IPO9	3.822707	importin 9
lnc-GSDMD-2	3.8439574	lnc-GSDMD-2:1
AC009097.3	3.8565602	ENSG00000260612
C1GALT1P1	3.8900714	Core 1 Synthase, Glycoprotein-N-Acetylgalactosamine 3-Beta-Galactosyltransferase 1 Pseudogene 1
KIRREL3-AS3	3.9024515	KIRREL3 antisense RNA 3
TLR8-AS1	4.000094	TLR8 antisense RNA 1
KRT18P24	4.1879582	Keratin 18 Pseudogene 24
LINC00242	4.2036963	long intergenic non-protein coding RNA 242
lnc-AL359736.1-2	4.248621	lnc-AL359736.1-2:1
ADARB2-AS1	4.2789326	ADARB2 antisense RNA 1
C16orf82	4.3380547	chromosome 16 open reading frame 82
CT83	4.433506	cancer/testis antigen 83
STK33	4.4345155	serine/threonine kinase 33
KISS1R	4.5108356	KISS1 receptor
OR8D2	4.596958	olfactory receptor, family 8, subfamily D, member 2 (gene/pseudogene)
SGCG	4.6387944	sarcoglycan, gamma (35kDa dystrophin-associated glycoprotein)
PRSS29P	4.667375	Serine Protease 29, Pseudogene
LINC00661	4.7172217	long intergenic non-protein coding RNA 661
lnc-TGFBRAP1-5	4.7654047	lnc-TGFBRAP1-5:3
lnc-SERINC5-1	4.7655835	lnc-SERINC5-1:2
CNTNAP2	4.8478794	contactin associated protein-like 2
LINC01344	4.8994393	long intergenic non-protein coding RNA 1344
STMN2	4.9121447	stathmin 2
FLJ45872	5.130464	FLJ45872 protein
DMGDH	5.132734	dimethylglycine dehydrogenase
FAM159B	5.281743	family with sequence similarity 159, member B

ADORA1	5.384696	adenosine A1 receptor
NKAIN1	5.5092096	Na <sup>+</sup> /K <sup>+</sup> transporting ATPase interacting 1
RTL1	5.7464585	retrotransposon-like 1
FLJ16171	5.881972	FLJ16171 protein
NEK8	5.903838	NIMA-related kinase 8
LOC100506688	6.011282	uncharacterized LOC100506688
LINC00261	9.743585	long intergenic non-protein coding RNA 261
LINC01905	10.101725	Long Intergenic Non-Protein Coding RNA 1905
TKTL2	11.317514	transketolase-like 2
FBXL21	11.521111	F-box and leucine-rich repeat protein 21 (gene/pseudogene)
SLC5A8	11.797434	solute carrier family 5 (sodium/monocarboxylate cotransporter), member 8
IRF8	14.167899	interferon regulatory factor 8

**Anexo B – Tabela dos genes diferencialmente expressos após silenciamento de *ASXL2* na linhagem MDA-MB-231 (*fold change* 2,0)**

<b>Símbolo do Gene</b>	<b>Fold change</b>	<b>Descrição</b>
RPL31	-281.16086	Ribosomal Protein L31
LOC100129697	-202.64232	uncharacterized LOC100129697
TMEM116	-200.88628	Transmembrane Protein 116
KLRB1	-158.67833	killer cell lectin-like receptor subfamily B, member 1
PROL1	-138.55702	proline rich, lacrimal 1
MMP21	-104.0049	matrix metalloproteinase 21
PROZ	-93.55351	protein Z, vitamin K-dependent plasma glycoprotein
LOC100131373	-45.956795	uncharacterized LOC100131373
CRTAM	-42.833443	cytotoxic and regulatory T cell molecule
TLR8	-35.58521	toll-like receptor 8
LINC00935	-35.410217	long intergenic non-protein coding RNA 935
MRVI1	-23.082567	murine retrovirus integration site 1 homolog
CLCA3P	-18.820498	chloride channel accessory 3, pseudogene
TEDDM1	-18.477453	transmembrane epididymal protein 1
GUCY2D	-14.071935	guanylate cyclase 2D, membrane (retina-specific)
PAGE1	-11.933665	P antigen family, member 1 (prostate associated)
MINDY4B	-11.799685	MINDY Family Member 4B
ZDHHC15	-11.664928	zinc finger, DHHC-type containing 15
TRIM15	-11.333998	tripartite motif containing 15
GLIPR1L1	-11.301301	GLI pathogenesis-related 1 like 1
SPRR2B	-10.82919	small proline-rich protein 2B
MEIKIN	-9.750172	Meiotic Kinetochores Factor
ENPP7	-9.466041	ectonucleotide pyrophosphatase/phosphodiesterase 7
MTMR7	-9.198176	myotubularin related protein 7
OLAH	-8.989962	oleoyl-ACP hydrolase
THAP2	-8.748788	THAP domain containing, apoptosis associated protein 2
TNIP3	-8.709629	TNFAIP3 interacting protein 3
LOC102724467	-8.510047	uncharacterized LOC102724467
LACRT	-8.021921	lacrigin
AKR1B15	-7.6223197	aldo-keto reductase family 1, member B15
GPR87	-7.4427037	G protein-coupled receptor 87
C8orf4	-6.8457	chromosome 8 open reading frame 4
LOC100127994	-6.8203754	uncharacterized LOC100127994

AKR1B10	-6.685944	aldo-keto reductase family 1, member B10 (aldose reductase)
MARCH8	-6.528383	membrane-associated ring finger (C3HC4) 8, E3 ubiquitin protein ligase
C17orf47	-6.294415	chromosome 17 open reading frame 47
VLDLR-AS1	-6.1904035	VLDLR antisense RNA 1
KLHL10	-6.1461096	kelch-like family member 10
SLCO1B1	-6.1411304	solute carrier organic anion transporter family, member 1B1
MBD3L1	-5.7783756	methyl-CpG binding domain protein 3-like 1
DPY19L2P2	-5.768069	DPY19L2 pseudogene 2
NWD1	-5.6257877	NACHT and WD repeat domain containing 1
ARL14	-5.5554314	ADP-ribosylation factor-like 14
ZIC3	-5.5059495	Zic family member 3
ASXL2	-5.464731	additional sex combs like transcriptional regulator 2
ANXA10	-5.364992	annexin A10
THEMIS	-5.3639665	thymocyte selection associated
KRT25	-5.333258	keratin 25, type I
HOXB1	-5.203106	homeobox B1
PCDH20	-5.152076	protocadherin 20
PPM1E	-4.998681	protein phosphatase, Mg <sup>2+</sup> /Mn <sup>2+</sup> dependent, 1E
HUS1B	-4.9525523	HUS1 checkpoint homolog b ( <i>S. pombe</i> )
GADL1	-4.8990984	glutamate decarboxylase-like 1
TMEM257	-4.858957	transmembrane protein 257
TMC5	-4.852636	transmembrane channel-like 5
PRKCQ	-4.6936245	protein kinase C, theta
FRY	-4.646479	furry homolog ( <i>Drosophila</i> )
PHACTR1	-4.5357757	phosphatase and actin regulator 1
XIRP2	-4.497737	xin actin-binding repeat containing 2
LOC100289120	-4.4834194	uncharacterized LOC100289120
FZD10	-4.456769	frizzled class receptor 10
GAP43	-4.4265013	growth associated protein 43
HOXA2	-4.325738	homeobox A2
HPSE2	-4.294909	heparanase 2 (inactive)
TRHDE-AS1	-4.259429	TRHDE antisense RNA 1
OR11G2	-4.167533	olfactory receptor, family 11, subfamily G, member 2
GSC	-4.1435847	gooseoid homeobox
AC007952.2	-4.1431026	Uncharacterized Protein ENSP00000382042-Like



lnc-WNT8A-2	-4.109492	lnc-WNT8A-2:1
KIAA0125	-4.092319	KIAA0125
RNF39	-4.055148	ring finger protein 39
KLRAP1	-4.0133123	killer cell lectin-like receptor subfamily A pseudogene 1
SLC5A7	-3.9955974	solute carrier family 5 (sodium/choline cotransporter), member 7
LOC644662	-3.9916995	uncharacterized LOC644662
ZG16	-3.9844325	zymogen granule protein 16
BDNF-AS	-3.9773748	BDNF antisense RNA
TARM1	-3.9596457	T cell-interacting, activating receptor on myeloid cells 1
CTH	-3.9503496	cystathionine gamma-lyase
SYT1	-3.948368	synaptotagmin I
ASNS	-3.9055524	asparagine synthetase (glutamine-hydrolyzing)
TVP23A	-3.9020543	trans-golgi network vesicle protein 23 homolog A (S. cerevisiae)
LRRC71	-3.9015636	leucine rich repeat containing 71
KBTBD8	-3.8806937	kelch repeat and BTB (POZ) domain containing 8
SYTL3	-3.8606913	synaptotagmin-like 3
SUPT20HL2	-3.859383	suppressor of Ty 20 homolog (S. cerevisiae)-like 2
IL21	-3.8507588	interleukin 21
LINC00575	-3.8422809	long intergenic non-protein coding RNA 575
AWAT2	-3.788634	acyl-CoA wax alcohol acyltransferase 2
FLJ43315	-3.7464106	asparagine synthetase pseudogene
C8orf34	-3.7459438	chromosome 8 open reading frame 34
AC005699.1	-3.7430773	Uncharacterized LOC105374483
MYCT1	-3.727794	myc target 1
LOC729930	-3.7224405	chromosome 11 open reading frame 74 pseudogene
TMEM100	-3.7032528	transmembrane protein 100
FAM9B	-3.693763	family with sequence similarity 9, member B
LOC643623	-3.6858752	uncharacterized LOC643623
GPR123	-3.6691608	G protein-coupled receptor 123
lnc-FMNL2-2	-3.6644995	lnc-FMNL2-2:1
EPN3	-3.6376112	epsin 3
FRMPD2	-3.616224	FERM and PDZ domain containing 2
AOX2P	-3.591784	aldehyde oxidase 2 pseudogene
LOC729867	-3.571011	uncharacterized LOC729867
LOC344887	-3.5515888	NmrA-like family domain containing 1 pseudogene

ANKRD11	-3.5344968	Ankyrin Repeat Domain 11
LOC100131180	-3.5131254	uncharacterized LOC100131180
LOC151760	-3.5077493	uncharacterized LOC151760
ARHGEF1	-3.5068915	Rho Guanine Nucleotide Exchange Factor 1
TCP11	-3.4729621	t-complex 11, testis-specific
ZNF385C	-3.4635181	zinc finger protein 385C
SPATA4	-3.4570305	spermatogenesis associated 4
LRRN4	-3.4555004	leucine rich repeat neuronal 4
ZC3H8	-3.4428377	zinc finger CCCH-type containing 8
SCARNA1	-3.4254675	small Cajal body-specific RNA 1
lnc-C17orf97-4	-3.4242003	lnc-C17orf97-4:1
C9orf152	-3.3921235	chromosome 9 open reading frame 152
MYPN	-3.3871162	myopalladin
UBL5P1	-3.3862188	Ubiquitin Like 5 Pseudogene 1
lnc-PRPF18-5	-3.3766627	lnc-PRPF18-5:1
LRRC7	-3.3702738	leucine rich repeat containing 7
SORCS3-AS1	-3.3628638	SORCS3 antisense RNA 1
LINC02310	-3.3582408	Long Intergenic Non-Protein Coding RNA 2310
LOC100128035	-3.342496	VGSA5840
DNAJB8	-3.3273256	DnaJ (Hsp40) homolog, subfamily B, member 8
SCN1A	-3.3243666	sodium channel, voltage gated, type I alpha subunit
EREG	-3.3182366	epiregulin
GDPD1	-3.3124788	glycerophosphodiester phosphodiesterase domain containing 1
MYSM1	-3.2958186	Myb-like, SWIRM and MPN domains 1
FAM26D	-3.2948601	family with sequence similarity 26, member D
GEMIN8P4	-3.292253	gem (nuclear organelle) associated protein 8 pseudogene 4
CCDC60	-3.2910175	coiled-coil domain containing 60
LOC79999	-3.285422	uncharacterized LOC79999
TSPAN19	-3.2438164	tetraspanin 19
LARP4P	-3.2168877	La Ribonucleoprotein Domain Family Member 4 Pseudogene
CEP41	-3.2138026	centrosomal protein 41kDa
GALNT15	-3.2080188	polypeptide N-acetylgalactosaminyltransferase 15
LINC02354	-3.2041416	Long Intergenic Non-Protein Coding RNA 2354
LINC00238	-3.200247	long intergenic non-protein coding RNA 238
KCNS1	-3.198589	potassium voltage-gated channel, modifier subfamily S, member 1

ZNF366	-3.1985054	zinc finger protein 366
KRTAP5-8	-3.1958737	keratin associated protein 5-8
CHAC1	-3.1864676	ChaC glutathione-specific gamma-glutamylcyclotransferase 1
IRF8	-3.1827266	interferon regulatory factor 8
IGLV10-54	-3.1634881	Immunoglobulin Lambda Variable 10-54
TEMN3-AS1	-3.155491	TENM3 Antisense Transcript 1
OTX2	-3.1524556	orthodenticle homeobox 2
LMLN	-3.1492012	leishmanolysin-like (metallopeptidase M8 family)
LOC392452	-3.1390643	mitochondrial fission factor pseudogene
C10orf99	-3.1305375	chromosome 10 open reading frame 99
lnc-ANKRD40-1	-3.12207	lnc-ANKRD40-1:3
CXCL6	-3.1216354	C-X-C Motif Chemokine Ligand 6
LOC101929256	-3.1122096	uncharacterized LOC101929256
CDX2	-3.10912	caudal type homeobox 2
NBAT1	-3.095125	neuroblastoma associated transcript 1
CDH6	-3.0944178	cadherin 6, type 2, K-cadherin (fetal kidney)
KLRC3	-3.0913546	killer cell lectin-like receptor subfamily C, member 3
XLOC_12_013192	-3.0857718	XLOC_12_013192
CCDC173	-3.0780947	coiled-coil domain containing 173
LINC01710	-3.0776322	Long Intergenic Non-Protein Coding RNA 1710
DGKG	-3.0683055	diacylglycerol kinase, gamma 90kDa
TCHH	-3.0632203	trichohyalin
ECM2	-3.054624	extracellular matrix protein 2, female organ and adipocyte specific
PRR16	-3.04715	proline rich 16
LINC00710	-3.0413578	long intergenic non-protein coding RNA 710
LY6K	-3.0330234	lymphocyte antigen 6 complex, locus K
IGKV2D-24	-3.0230556	Immunoglobulin Kappa Variable 2D-24 (Non-Functional)
HIST1H2AL	-3.0131288	histone cluster 1, H2al
lnc-LSMD1-1	-2.9927928	lnc-LSMD1-1:2
LOC100289650	-2.992645	uncharacterized LOC100289650
FBXL13	-2.9900343	F-box and leucine-rich repeat protein 13
TTC29	-2.9878032	tetratricopeptide repeat domain 29
MTCP1	-2.9830968	mature T-cell proliferation 1
IGHV2-26	-2.9815736	Immunoglobulin Heavy Variable 2-26
IFNA2	-2.981211	interferon, alpha 2

TAS2R46	-2.9775982	taste receptor, type 2, member 46
LOC100128477	-2.9719257	uncharacterized LOC100128477
IL3	-2.9630442	interleukin 3
NPM1P52	-2.9626846	Nucleophosmin 1 Pseudogene 52
SLC16A1	-2.9317317	solute carrier family 16 (monocarboxylate transporter), member 1
KRT27	-2.9281898	keratin 27, type I
lnc-NCOA3-1	-2.9205937	lnc-NCOA3-1:1
GPR4	-2.920463	G protein-coupled receptor 4
ZNF816-ZNF321P	-2.9196987	ZNF816-ZNF321P readthrough
NBPF7	-2.9178798	neuroblastoma breakpoint family, member 7
IZUMO1	-2.898958	izumo sperm-egg fusion 1
GNAT2	-2.897654	guanine nucleotide binding protein (G protein), alpha transducing activity polypeptide 2
CLEC18C	-2.8826826	C-type lectin domain family 18, member C
PRAMEF11	-2.8807433	PRAME family member 11
LRRC18	-2.8683105	leucine rich repeat containing 18
MYOZ2	-2.8616285	myozenin 2
BTNL3	-2.8564548	butyrophilin-like 3
WNT3	-2.8501976	Wnt Family Member 3
AC096577.1	-2.8482676	ENSG00000248373
LOC731157	-2.8474593	uncharacterized LOC731157
AREG	-2.8446257	amphiregulin
CHRNA9	-2.8379111	cholinergic receptor, nicotinic, alpha 9 (neuronal)
ZFP69B	-2.8164418	ZFP69 zinc finger protein B
ADRB1	-2.8109949	adrenoceptor beta 1
IGHV3-66	-2.808459	Immunoglobulin Heavy Variable 3-66
FGF16	-2.7990065	fibroblast growth factor 16
SLITRK3	-2.7924826	SLIT and NTRK-like family, member 3
LOC100506128	-2.791305	LOC100506128
PXDNL	-2.7828014	peroxidasin-like
IGHV1-58	-2.7788918	Immunoglobulin Heavy Variable 1-58
CLDN10	-2.7759411	claudin 10
CD247	-2.775924	CD247 molecule
MATN4	-2.7729888	matrilin 4
DDIT4L	-2.7692418	DNA-damage-inducible transcript 4-like
MYF5	-2.769107	myogenic factor 5

PLAC1	-2.7651498	placenta-specific 1
KERA	-2.7627804	keratocan
IGHV3-20	-2.757845	Immunoglobulin Heavy Variable 3-20
ITGB2	-2.7495766	integrin, beta 2 (complement component 3 receptor 3 and 4 subunit)
BCAN	-2.7455027	brevican
LOC100128885	-2.7399123	uncharacterized LOC100128885
MED15P3	-2.7334828	Mediator Complex Subunit 15 Pseudogene 3
LINC00392	-2.7310224	long intergenic non-protein coding RNA 392
DSCR4	-2.7300951	Down syndrome critical region 4
LOC553103	-2.6914766	uncharacterized LOC553103
TMEM212	-2.6885426	transmembrane protein 212
OBP2A	-2.683777	odorant binding protein 2A
ONECUT3	-2.681639	one cut homeobox 3
AC138701.1	-2.6810598	ENSG00000215567
OR2S2	-2.6792948	olfactory receptor, family 2, subfamily S, member 2 (gene/pseudogene)
AL691403.1	-2.6740384	ENSG00000258740
LOC400768	-2.6739032	uncharacterized LOC400768
P2RY12	-2.6734052	purinergic receptor P2Y, G-protein coupled, 12
LOC101928569	-2.665205	uncharacterized LOC101928569
AC147055.3	-2.6640563	UDP Glucuronosyltransferase Family 2 Member B15 Pseudogene
CSPG4P1Y	-2.6577268	chondroitin sulfate proteoglycan 4 pseudogene 1, Y-linked
LOC100506747	-2.6574173	alpha-1,3-mannosyl-glycoprotein 4-beta-N-acetylglucosaminyltransferase-like protein LOC641515 homolog
NRP2	-2.6544886	neuropilin 2
CYP2A7	-2.6489425	cytochrome P450, family 2, subfamily A, polypeptide 7
FMO2	-2.6437624	flavin containing monooxygenase 2 (non-functional)
SLC8A3	-2.6387997	solute carrier family 8 (sodium/calcium exchanger), member 3
OR52N2	-2.6377976	olfactory receptor, family 52, subfamily N, member 2
FAM187B2P	-2.6349075	Family With Sequence Similarity 187 Member B2, Pseudogene
FADS6	-2.6132915	fatty acid desaturase 6
LOC440742	-2.6103642	uncharacterized LOC440742
CLCA4	-2.608583	chloride channel accessory 4
ZNF829	-2.6067896	zinc finger protein 829
LOC727944	-2.6027746	uncharacterized LOC727944
ZNF674	-2.5976887	zinc finger protein 674
MIR99AHG	-2.597596	mir-99a-let-7c cluster host gene (non-protein coding)

MAGEC2	-2.5974658	melanoma antigen family C, 2
WDR11-AS1	-2.5937486	WDR11 antisense RNA 1
CLDN11	-2.5932221	claudin 11
COL18A1-AS1	-2.591548	COL18A1 antisense RNA 1
DGCR9	-2.5846038	DiGeorge syndrome critical region gene 9 (non-protein coding)
LINC01387	-2.5800128	long intergenic non-protein coding RNA 1387
FLJ37201	-2.5779622	tigger transposable element derived 2 pseudogene
CSN2	-2.5701842	casein beta
ZCWPW2	-2.5696175	zinc finger, CW type with PWWP domain 2
OR9M1P	-2.5681288	Olfactory Receptor Family 9 Subfamily M Member 1, Pseudogene
SERPINB4	-2.5677724	serpin peptidase inhibitor, clade B (ovalbumin), member 4
TRPA1	-2.5645323	transient receptor potential cation channel, subfamily A, member 1
RNF212B	-2.5587308	ring finger protein 212B
PYDC2	-2.5564678	pyrin domain containing 2
CYSLTR2	-2.5562065	cysteinyl leukotriene receptor 2
lnc-AC021066.1-2	-2.5539775	lnc-AC021066.1-2:1
LOC389641	-2.5524745	uncharacterized LOC389641
PHF7	-2.5514476	PHD finger protein 7
TM4SF19	-2.5443559	transmembrane 4 L six family member 19
CRTAC1	-2.5368283	cartilage acidic protein 1
BOLL	-2.5338712	boule-like RNA-binding protein
NPPC	-2.5333922	natriuretic peptide C
CABP1	-2.5317912	calcium binding protein 1
ANKRD19P	-2.530889	ankyrin repeat domain 19, pseudogene
MYOC	-2.5275161	myocilin, trabecular meshwork inducible glucocorticoid response
STAB1	-2.5218832	stabilin 1
PROSER2-AS1	-2.521832	PROSER2 antisense RNA 1
LINC01021	-2.5214603	long intergenic non-protein coding RNA 1021
CHNRG	-2.5140324	cholinergic receptor, nicotinic, gamma (muscle)
ST6GALNAC3	-2.5086935	ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3)-N-acetylgalactosaminide alpha-2,6-sialyltransferase 3
KCNN3	-2.5076854	potassium channel, calcium activated intermediate/small conductance subfamily N alpha, member 3
JHDM1D-AS1	-2.503528	JHDM1D antisense RNA 1 (head to head)
LINC01270	-2.4950156	long intergenic non-protein coding RNA 1270
PRAMEF14	-2.493613	PRAME family member 14
LOC100128325	-2.4923074	uncharacterized LOC100128325

LCE4A	-2.4921248	late cornified envelope 4A
SYT6	-2.4905148	synaptotagmin VI
Inc-FCGR3B-2	-2.4889174	Inc-FCGR3B-2:1
LOC100507431	-2.488195	uncharacterized LOC100507431
GIP	-2.4878943	gastric inhibitory polypeptide
RTL1	-2.4843657	retrotransposon-like 1
SOX9-AS1	-2.4830287	SOX9 antisense RNA 1
C9	-2.4816408	complement component 9
OR9H1P	-2.4811845	olfactory receptor, family 9, subfamily H, member 1 pseudogene
SGK494	-2.4799948	uncharacterized serine/threonine-protein kinase SgK494
LOC100287477	-2.4796436	zinc finger protein 420-like
PIGR	-2.4793642	polymeric immunoglobulin receptor
CACNA1D	-2.478911	calcium channel, voltage-dependent, L type, alpha 1D subunit
TTY5	-2.477264	testis-specific transcript, Y-linked 5 (non-protein coding)
KCNE1	-2.477044	potassium channel, voltage gated subfamily E regulatory beta subunit 1
PSAT1	-2.4762287	phosphoserine aminotransferase 1
SNORA28	-2.468744	small nucleolar RNA, H/ACA box 28
RDH12	-2.4680007	retinol dehydrogenase 12 (all-trans/9-cis/11-cis)
WSCD1	-2.4676986	WSC domain containing 1
CCDC179	-2.4665582	coiled-coil domain containing 179
CD34	-2.4651518	CD34 molecule
COL5A2	-2.4620247	collagen, type V, alpha 2
SDR16C5	-2.4591103	short chain dehydrogenase/reductase family 16C, member 5
IPO7P1	-2.458269	Importin 7 Pseudogene 1
IGFL3	-2.455133	IGF-like family member 3
AC105345.1	-2.4514253	ENSG00000205959
C1orf145	-2.4497344	chromosome 1 open reading frame 145
LRRC37A9P	-2.4351356	Leucine Rich Repeat Containing 37 Member A9, Pseudogene
LCA5L	-2.4288645	Leber congenital amaurosis 5-like
CYP2A7P1	-2.4246907	Cytochrome P450 Family 2 Subfamily A Member 7 Pseudogene 1
HS6ST3	-2.423058	heparan sulfate 6-O-sulfotransferase 3
HSPB8	-2.4211757	heat shock 22kDa protein 8
ACOT4	-2.4130197	acyl-CoA thioesterase 4
CSMD3	-2.4129934	CUB and Sushi multiple domains 3
DTHD1	-2.412136	death domain containing 1

KCNT2	-2.4120457	potassium channel, sodium activated subfamily T, member 2
INSRR	-2.4111083	insulin receptor-related receptor
lnc-TRIM29-1	-2.4091737	lnc-TRIM29-1:2
MIR3945HG	-2.402031	MIR3945 Host Gene
LOC101927418	-2.3951902	uncharacterized LOC101927418
DUSP19	-2.3939314	dual specificity phosphatase 19
TSPYL5	-2.3911462	TSPY-like 5
FRRS1	-2.3898401	ferric-chelate reductase 1
AC243772.3	-2.3893158	Phosphodiesterase 4D Interacting Protein Pseudogene
PP2D1	-2.3884552	protein phosphatase 2C-like domain containing 1
PRR23C	-2.386624	proline rich 23C
OTC	-2.3854084	ornithine carbamoyltransferase
RTP2	-2.3813365	receptor (chemosensory) transporter protein 2
KLHL15	-2.380131	kelch-like family member 15
GFM1	-2.3756583	G elongation factor, mitochondrial 1
FAM27B	-2.3755214	Family With Sequence Similarity 27 Member B
DNMT3L	-2.3747373	DNA (cytosine-5-)-methyltransferase 3-like
DRD2	-2.3679843	dopamine receptor D2
GNRHR	-2.365903	gonadotropin-releasing hormone receptor
ADPRH	-2.363053	ADP-ribosylarginine hydrolase
AC021106.3	-2.3592439	ENSG00000279386
CYP2C18	-2.3578296	cytochrome P450, family 2, subfamily C, polypeptide 18
COL6A4P2	-2.3562474	collagen, type VI, alpha 4 pseudogene 2
SVIP	-2.351962	small VCP/p97-interacting protein
B3GNT6	-2.3488932	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 6 (core 3 synthase)
SCML2	-2.3475616	sex comb on midleg-like 2 (Drosophila)
SERPINB9	-2.339751	serpin peptidase inhibitor, clade B (ovalbumin), member 9
DAW1	-2.3385203	dynein assembly factor with WDR repeat domains 1
AL356056.3	-2.3367286	Uncharacterized LOC102724350
PSAT1P1	-2.3295646	Phosphoserine Aminotransferase 1 Pseudogene 1
C15orf54	-2.3272905	chromosome 15 open reading frame 54
NEK10	-2.3263438	NIMA-related kinase 10
ARL14EPL	-2.324215	ADP-ribosylation factor-like 14 effector protein-like
LOC400965	-2.3219798	uncharacterized LOC400965
CLEC2L	-2.3217273	C-type lectin domain family 2, member L



NCF1B	-2.316666	neutrophil cytosolic factor 1B pseudogene
CLEC4GP1	-2.3162944	C-type lectin domain family 4, member G pseudogene 1
PRORS1P	-2.3105888	prolyl-tRNA synthetase associated domain containing 1, pseudogene
DPY19L2P1	-2.3079991	DPY19L2 pseudogene 1
LOC727710	-2.3058712	uncharacterized LOC727710
OR1Q1	-2.301167	olfactory receptor, family 1, subfamily Q, member 1
CCRN4L	-2.3010912	CCR4 carbon catabolite repression 4-like ( <i>S. cerevisiae</i> )
C2orf80	-2.300891	chromosome 2 open reading frame 80
IFRD1	-2.30067	interferon-related developmental regulator 1
RAMP2	-2.291362	receptor (G protein-coupled) activity modifying protein 2
RGS1	-2.2861326	regulator of G-protein signaling 1
CHST4	-2.285194	carbohydrate (N-acetylglucosamine 6-O) sulfotransferase 4
LOC399815	-2.2831886	chromosome 10 open reading frame 88 pseudogene
MKX	-2.2826734	mohawk homeobox
EMR3	-2.2783551	egf-like module containing, mucin-like, hormone receptor-like 3
REXO1L1P	-2.2780008	REX1, RNA exonuclease 1 homolog ( <i>S. cerevisiae</i> )-like 1, pseudogene
LOC145837	-2.2762978	uncharacterized LOC145837
C9orf135-AS1	-2.2757378	C9orf135 antisense RNA 1 (head to head)
OR8H2	-2.2755935	olfactory receptor, family 8, subfamily H, member 2
AL359073.1	-2.2682607	Ribosomal Protein L32 Pseudogene 22
FAM27C	-2.2669218	family with sequence similarity 27, member C
LOC100129473	-2.2619536	uncharacterized LOC100129473
MMP20	-2.2602592	matrix metalloproteinase 20
TRBV11-1	-2.259834	T Cell Receptor Beta Variable 11-1
LOC100129042	-2.2571225	uncharacterized LOC100129042
PCK2	-2.2560358	phosphoenolpyruvate carboxykinase 2 (mitochondrial)
CLDN22	-2.2511299	claudin 22
PLEKHG7	-2.2437692	pleckstrin homology domain containing, family G (with RhoGef domain) member 7
PSAT1P3	-2.2434425	Phosphoserine Aminotransferase 1 Pseudogene 3
PREX1	-2.239248	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1
LOC644189	-2.238853	acyl-CoA thioesterase 4 pseudogene
TPTE2	-2.2353804	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2
TINCR	-2.234793	tissue differentiation-inducing non-protein coding RNA
ADAMTSL3	-2.233128	ADAMTS-like 3
ZNF487	-2.2285419	zinc finger protein 487

CYBB	-2.228439	cytochrome b-245, beta polypeptide
CYP1A1	-2.226935	cytochrome P450, family 1, subfamily A, polypeptide 1
SNAR-C3	-2.2255492	small ILF3/NF90-associated RNA C3
AC131532.1	-2.2225075	Uncharacterized LOC105369876
DICER1	-2.2216923	dicer 1, ribonuclease type III
TMEM252	-2.2207592	transmembrane protein 252
IL36RN	-2.219776	interleukin 36 receptor antagonist
PNLDC1	-2.216669	poly(A)-specific ribonuclease (PARN)-like domain containing 1
lnc-IL4R-2	-2.2123532	lnc-IL4R-2:1
LOC256880	-2.2118056	uncharacterized LOC256880
OLFML3	-2.2112303	olfactomedin-like 3
PIWIL4	-2.2089279	piwi-like RNA-mediated gene silencing 4
TLL1	-2.2089143	tolloid-like 1
LOC644135	-2.2078285	uncharacterized LOC644135
RNF186	-2.2049737	ring finger protein 186
PABPC5-AS1	-2.199422	PABPC5 antisense RNA 1
CNRIP1	-2.1939406	cannabinoid receptor interacting protein 1
LINC00662	-2.1844711	long intergenic non-protein coding RNA 662
TRBV6-7	-2.1809034	T Cell Receptor Beta Variable 6-7 (Non-Functional)
RASSF8	-2.1804261	Ras association (RalGDS/AF-6) domain family (N-terminal) member 8
OR51J1	-2.1803448	Olfactory Receptor Family 51 Subfamily J Member 1 (Gene/Pseudogene)
FAM110C	-2.178237	family with sequence similarity 110, member C
KDM7A	-2.1773365	lysine (K)-specific demethylase 7A
lnc-OBFC2A-3	-2.175781	lnc-OBFC2A-3:1
OR13C8	-2.1716413	olfactory receptor, family 13, subfamily C, member 8
SYP	-2.1714075	synaptophysin
PIP5K1A	-2.169327	phosphatidylinositol-4-phosphate 5-kinase, type I, alpha
TRAF3IP2-AS1	-2.1686668	TRAF3IP2 antisense RNA 1
LILRA5	-2.1685545	leukocyte immunoglobulin-like receptor, subfamily A (with TM domain), member 5
PCAT18	-2.1674225	prostate cancer associated transcript 18 (non-protein coding)
AARS	-2.1672459	alanyl-tRNA synthetase
IFNA7	-2.1670535	interferon, alpha 7
LOC57399	-2.1662261	uncharacterized gastric protein ZA52P
LOC339468	-2.1660213	uncharacterized LOC339468
DDIT3	-2.1659942	DNA-damage-inducible transcript 3

OR5H15	-2.1657534	olfactory receptor, family 5, subfamily H, member 15
HELT	-2.1657166	helt bHLH transcription factor
INSC	-2.1655333	inscuteable homolog (Drosophila)
PDE4D	-2.1651402	phosphodiesterase 4D, cAMP-specific
LOC101927762	-2.1637878	uncharacterized LOC101927762
LOC100133746	-2.1631846	LOC100133746
HHIPL2	-2.1625032	HHIP-like 2
CYP4F30P	-2.1624901	cytochrome P450, family 4, subfamily F, polypeptide 30, pseudogene
A2M-AS1	-2.162399	A2M antisense RNA 1 (head to head)
ZNF678	-2.1598961	zinc finger protein 678
LOC102723895	-2.1577277	uncharacterized LOC102723895
RASA2	-2.155661	RAS p21 protein activator 2
KHDC3L	-2.153843	KH domain containing 3-like, subcortical maternal complex member
lnc-KLHL36-1	-2.1538281	lnc-KLHL36-1:1
ST6GAL2	-2.1531007	ST6 beta-galactosamide alpha-2,6-sialyltransferase 2
CASC16	-2.15193	cancer susceptibility candidate 16 (non-protein coding)
PPP1R1C	-2.1495845	protein phosphatase 1, regulatory (inhibitor) subunit 1C
HSPB2	-2.1483684	heat shock 27kDa protein 2
CCL13	-2.1459339	chemokine (C-C motif) ligand 13
LINC00112	-2.1454082	long intergenic non-protein coding RNA 112
GABRR1	-2.1453052	gamma-aminobutyric acid (GABA) A receptor, rho 1
C3P1	-2.1434548	complement component 3 precursor pseudogene
LRRC74A	-2.141863	leucine rich repeat containing 74A
TAT	-2.1374996	tyrosine aminotransferase
TDRP	-2.1354275	testis development related protein
FLJ33534	-2.1341054	uncharacterized LOC285150
VAT1L	-2.1316483	vesicle amine transport 1-like
XLOC_12_013630	-2.1268854	XLOC_12_013630
RNF157	-2.1254447	ring finger protein 157
lnc-RARS-3	-2.1250527	lnc-RARS-3:1
TLR6	-2.12368	toll-like receptor 6
BTBD8	-2.122725	BTB (POZ) domain containing 8
TCPI0L2	-2.119468	t-complex 10-like 2
PLBD1-AS1	-2.1188817	PLBD1 antisense RNA 1
SPATA31D1	-2.1171114	SPATA31 subfamily D, member 1

LINC00899	-2.1169066	long intergenic non-protein coding RNA 899
RFPL3S	-2.1162522	RFPL3 antisense
PCDHGC3	-2.115468	protocadherin gamma subfamily C, 3
AC019294.1	-2.1118772	Golgin-Like Pseudogene
ZNF582	-2.1101718	zinc finger protein 582
HIST4H4	-2.109737	histone cluster 4, H4
MIDN	-2.1080885	midnolin
MIMT1	-2.1071138	MER1 repeat containing imprinted transcript 1 (non-protein coding)
HVCN1	-2.1063035	hydrogen voltage gated channel 1
lnc-MORF4L1-2	-2.1058302	lnc-MORF4L1-2:1
CLIP4	-2.1027663	CAP-GLY domain containing linker protein family, member 4
CH25H	-2.1019819	cholesterol 25-hydroxylase
TMEM79	-2.0994995	transmembrane protein 79
SEMA6D	-2.0990427	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D
SYT9	-2.0982685	synaptotagmin IX
JSRP1	-2.0971942	junctional sarcoplasmic reticulum protein 1
FSCN1	-2.09557	fascin actin-bundling protein 1
CYP3A5	-2.0955021	cytochrome P450, family 3, subfamily A, polypeptide 5
LOC285638	-2.095179	uncharacterized LOC285638
TMEM45A	-2.0933132	transmembrane protein 45A
IGHV3-48	-2.0923207	Immunoglobulin Heavy Variable 3-48
OR52R1	-2.0919168	olfactory receptor, family 52, subfamily R, member 1 (gene/pseudogene)
KCNMB2	-2.0905595	potassium channel subfamily M regulatory beta subunit 2
GCNT7	-2.090399	glucosaminyl (N-acetyl) transferase family member 7
TNRC6C-AS1	-2.08997	TNRC6C antisense RNA 1
EAF2	-2.089841	ELL associated factor 2
IGHV3OR16-9	-2.0892134	immunoglobulin Heavy Variable 3/OR16-9 (Non-Functional)
PAFAH2	-2.0867212	platelet-activating factor acetylhydrolase 2, 40kDa
LOC100129675	-2.082527	uncharacterized LOC100129675
SRXN1	-2.0802445	sulfiredoxin 1
BX640514.1	-2.0767567	ENSG00000264490
ACTL7B	-2.076592	actin-like 7B
FGF17	-2.0756817	fibroblast growth factor 17
PAK3	-2.074804	p21 protein (Cdc42/Rac)-activated kinase 3
PLA2G12B	-2.074616	phospholipase A2, group XIIB

HMG5	-2.0734668	high mobility group nucleosome binding domain 5
CNDP1	-2.0721865	carnosine dipeptidase 1 (metallopeptidase M20 family)
TMEM132E	-2.0711339	transmembrane protein 132E
ANKRD33B	-2.0660956	ankyrin repeat domain 33B
TPD52L1	-2.0660143	tumor protein D52-like 1
TOMM22P3	-2.0658379	TOMM22 Pseudogene 3
ZNF471	-2.065386	zinc finger protein 471
XLOC_12_006584	-2.0638125	XLOC_12_006584
MYT1L	-2.063422	myelin transcription factor 1-like
LOC100128950	-2.062293	uncharacterized LOC100128950
UGT2B28	-2.059823	UDP glucuronosyltransferase 2 family, polypeptide B28
SOX21	-2.0597258	SRY (sex determining region Y)-box 21
SDC4P	-2.0595484	syndecan 4 pseudogene
GTSF1L	-2.058977	gametocyte specific factor 1-like
KIAA1467	-2.0588825	KIAA1467
GAS7	-2.0556405	growth arrest-specific 7
LOC100128482	-2.048197	uncharacterized LOC100128482
LOC284395	-2.0477247	uncharacterized LOC284395
MAGEB6	-2.046788	melanoma antigen family B, 6
LOC389332	-2.0449262	uncharacterized LOC389332
CYP11B1	-2.0442162	cytochrome P450, family 11, subfamily B, polypeptide 1
ZNF861P	-2.0424047	Zinc Finger Protein 861, Pseudogene
PPP4R1L	-2.042042	Protein Phosphatase 4 Regulatory Subunit 1 Like (Pseudogene)
KIAA1324	-2.0401902	KIAA1324
NUAK2	-2.0397897	NUAK family, SNF1-like kinase, 2
CASP16	-2.0371568	caspase 16, apoptosis-related cysteine peptidase (putative)
LETM1P2	-2.0368752	Leucine Zipper And EF-Hand Containing Transmembrane Protein 1 Pseudogene 2
CST6	-2.0351977	cystatin E/M
GPCPD1	-2.0348997	glycerophosphocholine phosphodiesterase GDE1 homolog (S. cerevisiae)
LOC102723428	-2.0317357	uncharacterized LOC102723428
OTOR	-2.0315487	otoraplin
MBL1P	-2.0308983	mannose-binding lectin (protein A) 1, pseudogene
UPP1	-2.030459	uridine phosphorylase 1
AC138761.3	-2.030284	ENSG00000264811
MTNR1B	-2.0278683	melatonin receptor 1B

TRBV21-1	-2.0271962	T Cell Receptor Beta Variable 21-1 (Pseudogene)
ADTRP	-2.02518	androgen-dependent TFPI-regulating protein
TCL1B	-2.0248997	T-cell leukemia/lymphoma 1B
MMP2	-2.0198925	matrix metalloproteinase 2 (gelatinase A, 72kDa gelatinase, 72kDa type IV collagenase)
LOC100506731	-2.0195012	uncharacterized LOC100506731
TRPC7	-2.018477	transient receptor potential cation channel, subfamily C, member 7
GLRA4	-2.0161967	glycine receptor, alpha 4
CIB4	-2.015913	calcium and integrin binding family member 4
LOC102724747	-2.0127802	ovostatin homolog 2-like
YARS	-2.0126896	tyrosyl-tRNA synthetase
CFHR2	-2.0122042	complement factor H-related 2
DEFB134	-2.01195	defensin, beta 134
CASP14	-2.011831	caspase 14, apoptosis-related cysteine peptidase
WDR63	-2.011361	WD repeat domain 63
LOC729296	-2.0112276	uncharacterized LOC729296
SPHKAP	-2.0109065	SPHK1 interactor, AKAP domain containing
ODAM	-2.0088298	odontogenic, ameloblast associated
CNGA1	-2.0079389	cyclic nucleotide gated channel alpha 1
SPANXA1	-2.0079176	sperm protein associated with the nucleus, X-linked, family member A1
BNIP3P41	-2.0071497	BCL2 Interacting Protein 3 Pseudogene 41
GCLC	-2.006998	glutamate-cysteine ligase, catalytic subunit
NIPAL2	-2.0054862	NIPA-like domain containing 2
PROKR1	-2.0020928	prokineticin receptor 1
CADM2	-2.0013175	cell adhesion molecule 2
RASGRF2	-2.001037	Ras protein-specific guanine nucleotide-releasing factor 2
TDG	-2.0001163	thymine-DNA glycosylase
TMEM132A	2.0015051	transmembrane protein 132A
ZSWIM8	2.0015554	zinc finger, SWIM-type containing 8
MAST1	2.0016096	microtubule associated serine/threonine kinase 1
KIAA1659	2.0030422	KIAA1659
ARHGAP20	2.0036383	Rho GTPase activating protein 20
TRBV27	2.0043232	T Cell Receptor Beta Variable 27
TERF2	2.0051012	telomeric repeat binding factor 2
LINC01011	2.0051343	long intergenic non-protein coding RNA 1011
SKIL	2.0065238	SKI-like proto-oncogene

FAM227A	2.006528	family with sequence similarity 227, member A
TBC1D22A-AS1	2.0073602	TBC1D22A antisense RNA 1
ZNF365	2.0094752	zinc finger protein 365
FLJ42627	2.0098758	uncharacterized LOC645644
LINC01431	2.0101607	long intergenic non-protein coding RNA 1431
FECH	2.0103176	ferrochelatase
DHRS3	2.0114355	dehydrogenase/reductase (SDR family) member 3
KPNA7	2.0122201	karyopherin alpha 7 (importin alpha 8)
TRIM43	2.0140567	tripartite motif containing 43
NANOG	2.014277	Nanog homeobox
LINC01582	2.0156546	long intergenic non-protein coding RNA 1582
ADAMTS20	2.0159485	ADAM metallopeptidase with thrombospondin type 1 motif, 20
IL4I1	2.016042	interleukin 4 induced 1
C15orf45	2.0172098	chromosome 15 open reading frame 45
RPL23AP32	2.0191908	ribosomal protein L23a pseudogene 32
IVNS1ABP	2.0201755	influenza virus NS1A binding protein
ARMC12	2.020448	armadillo repeat containing 12
HEXA-AS1	2.0207748	HEXA antisense RNA 1
NOL4L	2.0222137	nucleolar protein 4-like
LOC388210	2.0222802	uncharacterized LOC388210
IL17C	2.0225925	interleukin 17C
lnc-DUSP4-1	2.0230298	lnc-DUSP4-1:1
LOXL2	2.0246472	lysyl oxidase-like 2
MMP9	2.0247858	matrix metallopeptidase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)
PYCARD	2.0247905	PYD and CARD domain containing
FLJ39080	2.0257049	uncharacterized LOC441355
LOC100128573	2.0257392	uncharacterized LOC100128573
OR2A12	2.0258412	olfactory receptor, family 2, subfamily A, member 12
SORBS1	2.0259922	sorbin and SH3 domain containing 1
CD83	2.0261269	CD83 molecule
DHTKD1	2.0268192	dehydrogenase E1 and transketolase domain containing 1
TBC1D21	2.0281217	TBC1 domain family, member 21
SATL1	2.0301776	spermidine/spermine N1-acetyl transferase-like 1
TRAV5	2.030955	T Cell Receptor Alpha Variable 5
PLEKHM1	2.0318081	pleckstrin homology domain containing, family M (with RUN domain) member 1

TGFBR1	2.0343745	transforming growth factor, beta receptor 1
FAM90A27P	2.034594	family with sequence similarity 90, member A27, pseudogene
SDPR	2.0346556	serum deprivation response
ATP11C	2.0350134	ATPase, class VI, type 11C
OVOL3	2.036501	ovo-like zinc finger 3
DGKI	2.0366096	diacylglycerol kinase, iota
SETBP1	2.036987	SET binding protein 1
GP5	2.038918	glycoprotein V (platelet)
ANKRD6	2.03962	ankyrin repeat domain 6
LCTL	2.0409658	lactase-like
LOC100129617	2.0428674	uncharacterized LOC100129617
TMEM170B	2.0429115	transmembrane protein 170B
PDPR	2.0434713	pyruvate dehydrogenase phosphatase regulatory subunit
PCK1	2.0436368	phosphoenolpyruvate carboxykinase 1 (soluble)
IL17B	2.0438595	interleukin 17B
PYHIN5P	2.0447752	Pyrin And HIN Domain Family Member 5, Pseudogene
TMIE	2.0454764	transmembrane inner ear
C4orf50	2.045897	chromosome 4 open reading frame 50
ZFAND6	2.0478947	zinc finger, AN1-type domain 6
RAC3	2.0479069	ras-related C3 botulinum toxin substrate 3 (rho family, small GTP binding protein Rac3)
CAMK4	2.0482013	calcium/calmodulin-dependent protein kinase IV
DOK4	2.0496242	docking protein 4
ST8SIA3	2.0496798	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltransferase 3
NRG2	2.0497997	neuregulin 2
M1AP	2.0513692	meiosis 1 associated protein
SEC63	2.0534513	SEC63 homolog ( <i>S. cerevisiae</i> )
POLN	2.0552058	polymerase (DNA directed) nu
RGS9	2.0557566	regulator of G-protein signaling 9
SRPK2	2.056244	SRSF protein kinase 2
CDKN3	2.0571318	cyclin-dependent kinase inhibitor 3
IGHV3-64	2.0583572	Immunoglobulin Heavy Variable 3-64
AGO4	2.062529	argonaute RISC catalytic component 4
PLIN1	2.0641873	perilipin 1
POLR2KP2	2.0646002	RNA Polymerase II Subunit K Pseudogene 2
RHOU	2.0656455	ras homolog family member U



SOCS4	2.0657945	suppressor of cytokine signaling 4
ZFP28	2.066997	ZFP28 zinc finger protein
AC092902.3	2.067679	ENSG00000272840
SEC24A	2.0681102	SEC24 family member A
DNAAF1	2.0691888	dynein, axonemal, assembly factor 1
CPED1	2.0697966	cadherin-like and PC-esterase domain containing 1
STON1	2.070138	stonin 1
KCNK17	2.0709302	potassium channel, two pore domain subfamily K, member 17
PTPN22	2.071416	protein tyrosine phosphatase, non-receptor type 22 (lymphoid)
NEXN	2.0718403	nexilin (F actin binding protein)
MAGEB3	2.0730135	melanoma antigen family B, 3
lnc-C5orf42-2	2.0731332	lnc-C5orf42-2:1
lnc-HADH-1	2.0738688	lnc-HADH-1:7
SOST	2.0739162	sclerostin
TRBV11-2	2.0744898	T Cell Receptor Beta Variable 11-2
HS6ST1	2.0753493	heparan sulfate 6-O-sulfotransferase 1
SLIT1	2.0757713	slit homolog 1 (Drosophila)
TMPRSS5	2.0762491	transmembrane protease, serine 5
BFSP1	2.076502	beaded filament structural protein 1, filensin
APC2	2.0772436	adenomatosis polyposis coli 2
LPCAT2	2.0785565	lysophosphatidylcholine acyltransferase 2
RAB36	2.0787408	RAB36, member RAS oncogene family
NTF4	2.0791702	neurotrophin 4
ICAM5	2.0796664	intercellular adhesion molecule 5, telencephalin
BRSK1	2.0805259	BR serine/threonine kinase 1
SLC10A4	2.0819528	solute carrier family 10, member 4
ZNF385B	2.0840237	zinc finger protein 385B
FLJ46284	2.0866656	uncharacterized LOC441369
FLJ46066	2.0885441	uncharacterized LOC401103
GOLT1A	2.089428	golgi transport 1A
CHRD1	2.0901766	chordin-like 1
CACHD1	2.0904849	cache domain containing 1
CHST1	2.0916307	carbohydrate (keratan sulfate Gal-6) sulfotransferase 1
IGHEP2	2.0918167	Immunoglobulin Heavy Constant Epsilon P2 (Pseudogene)
TREML4	2.0921683	triggering receptor expressed on myeloid cells-like 4

LINC00890	2.0921948	long intergenic non-protein coding RNA 890
MCHR2-AS1	2.0939481	MCHR2 antisense RNA 1
ADAMTS1	2.0942295	ADAM metalloproteinase with thrombospondin type 1 motif, 1
EPHA5-AS1	2.0943654	EPHA5 antisense RNA 1
IFT57	2.094667	intraflagellar transport 57
SDC1	2.0948222	syndecan 1
ANGPTL4	2.0953748	angiopoietin-like 4
LOC100996405	2.095698	uncharacterized LOC100996405
LINC01530	2.0968962	long intergenic non-protein coding RNA 1530
PKIA	2.0991318	protein kinase (cAMP-dependent, catalytic) inhibitor alpha
ANKRD26P3	2.1006157	ankyrin repeat domain 26 pseudogene 3
ELL3	2.1008282	elongation factor RNA polymerase II-like 3
ZNF804A	2.1010993	zinc finger protein 804A
C12orf57	2.1023064	chromosome 12 open reading frame 57
AC006328.2	2.1026585	Putative Ubiquitin-Conjugating Enzyme E2Q2-Like Protein
DNAAF3	2.103423	dynein, axonemal, assembly factor 3
RALGAPA2	2.1034465	Ral GTPase activating protein, alpha subunit 2 (catalytic)
lnc-KRT18-1	2.1048133	lnc-KRT18-1:1
LINC02599	2.1056645	long intergenic non-protein coding RNA 2599
LINC00589	2.1061552	long intergenic non-protein coding RNA 589
ZBTB42	2.10638	zinc finger and BTB domain containing 42
OR1M1	2.1068547	olfactory receptor, family 1, subfamily M, member 1
GFAP	2.107034	glial fibrillary acidic protein
VN1R1	2.1085403	vomer nasal 1 receptor 1
SHANK3	2.1090732	SH3 and multiple ankyrin repeat domains 3
HTRA1	2.109559	HtrA serine peptidase 1
lnc-CDYL2-6	2.1107216	lnc-CDYL2-6:5
HES7	2.1125567	hes family bHLH transcription factor 7
OCR1	2.1127522	ovarian cancer-related protein 1
KRT86	2.1136587	keratin 86, type II
SH2D1B	2.1146526	SH2 domain containing 1B
RAMP3	2.1150374	receptor (G protein-coupled) activity modifying protein 3
OR52M1	2.115443	olfactory receptor, family 52, subfamily M, member 1
FUT8-AS1	2.115573	FUT8 antisense RNA 1
KIAA0408	2.1157205	KIAA0408

LTBP3	2.115859	latent transforming growth factor beta binding protein 3
CFAP99	2.1174815	cilia and flagella associated protein 99
PSG2	2.1203585	pregnancy specific beta-1-glycoprotein 2
DHRS12	2.1215312	dehydrogenase/reductase (SDR family) member 12
FSTL3	2.1215804	follistatin-like 3 (secreted glycoprotein)
PROCR	2.1216884	protein C receptor, endothelial
MYL1	2.1246336	myosin, light chain 1, alkali; skeletal, fast
CLSTN2	2.1254666	calsyntenin 2
OR12D1	2.127134	Olfactory Receptor Family 12 Subfamily D Member 1 (Gene/Pseudogene)
HP09025	2.1291127	uncharacterized LOC100652929
HOXA13	2.1301398	homeobox A13
FAM66C	2.1306815	family with sequence similarity 66, member C
IGSF5	2.1313004	immunoglobulin superfamily, member 5
H2BFWT	2.132926	H2B histone family, member W, testis-specific
ATOH1	2.1342022	atonal homolog 1 (Drosophila)
LOC101929584	2.134238	uncharacterized LOC101929584
RADIL	2.1359892	Ras association and DIL domains
FBN1	2.1378574	fibrillin 1
LOC100128356	2.1381924	protein transactivated by hepatitis B virus E antigen
GPRIN2	2.138714	G protein regulated inducer of neurite outgrowth 2
SSPN	2.1388257	sarcospan
SMAD2	2.1389205	SMAD family member 2
LANCL2	2.1393778	LanC lantibiotic synthetase component C-like 2 (bacterial)
FLJ44674	2.1422474	FLJ44674 protein
UNC45B	2.1435318	unc-45 homolog B (C. elegans)
MYL12A	2.143573	myosin, light chain 12A, regulatory, non-sarcomeric
DLEC1	2.1443887	deleted in lung and esophageal cancer 1
ELF5	2.145925	E74-like factor 5 (ets domain transcription factor)
TGFBI	2.1460173	transforming growth factor, beta-induced, 68kDa
ZBED3	2.147618	zinc finger, BED-type containing 3
SOGA3	2.1484473	SOGA family member 3
LOC100129034	2.1486895	uncharacterized LOC100129034
GH2	2.1492825	growth hormone 2
TMPRSS12	2.1498268	transmembrane (C-terminal) protease, serine 12
TMEM215	2.1514616	transmembrane protein 215

SIRT4	2.1519587	sirtuin 4
GRXCR1	2.1531105	glutaredoxin, cysteine rich 1
KIAA1211L	2.1546578	KIAA1211-like
ROPN1	2.1558309	rhophilin associated tail protein 1
LINC00592	2.156181	long intergenic non-protein coding RNA 592
HMGB3P12	2.1589458	High Mobility Group Box 3 Pseudogene 12
MB	2.158999	myoglobin
LINC00304	2.1602378	long intergenic non-protein coding RNA 304
RAB26	2.1622617	RAB26, member RAS oncogene family
CRYGC	2.1649492	crystallin, gamma C
XLOC_12_006196	2.166712	XLOC_12_006196
lnc-C1orf86-1	2.16834	lnc-C1orf86-1:2
HLA-V	2.1696794	Major Histocompatibility Complex, Class I, V (Pseudogene)
BSPH1	2.1702805	binder of sperm protein homolog 1
FAM71D	2.171601	family with sequence similarity 71, member D
TMEM169	2.1721056	transmembrane protein 169
TM4SF18	2.1721485	transmembrane 4 L six family member 18
FABP4	2.174132	fatty acid binding protein 4, adipocyte
DNMT1	2.1751225	DNA (cytosine-5-)-methyltransferase 1
MOGS	2.178072	mannosyl-oligosaccharide glucosidase
LINC00667	2.1788504	long intergenic non-protein coding RNA 667
RPRML	2.1793041	reprimo-like
OR4P4	2.1829276	olfactory receptor, family 4, subfamily P, member 4
LOC100287704	2.1840348	uncharacterized LOC100287704
LINC01996	2.1849241	Long Intergenic Non-Protein Coding RNA 1996
DEFB123	2.185483	defensin, beta 123
GATA6-AS1	2.1855724	GATA6 antisense RNA 1 (head to head)
FFAR3	2.1861923	free fatty acid receptor 3
ALDOAP2	2.1864185	aldolase A, fructose-bisphosphate pseudogene 2
KDEL2	2.1867106	KDEL (Lys-Asp-Glu-Leu) containing 2
S1PR2	2.188958	sphingosine-1-phosphate receptor 2
CACNG7	2.189241	calcium channel, voltage-dependent, gamma subunit 7
LAMB2P1	2.1894429	laminin, beta 2 pseudogene 1
PIK3C2G	2.190364	phosphatidylinositol-4-phosphate 3-kinase, catalytic subunit type 2 gamma
NOP14-AS1	2.1918583	NOP14 antisense RNA 1

KCNJ2-AS1	2.192111	KCNJ2 antisense RNA 1 (head to head)
HIF1A	2.194977	hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor)
RASSF2	2.1949959	Ras association (RalGDS/AF-6) domain family member 2
ZFHX2	2.1955452	zinc finger homeobox 2
IGHV3-35	2.196696	Immunoglobulin Heavy Variable 3-35 (Non-Functional)
BEND4	2.1967275	BEN domain containing 4
FAM110D	2.1971648	family with sequence similarity 110, member D
RTN1	2.1978383	reticulon 1
LOC100506272	2.199099	uncharacterized LOC100506272
LYSMD3	2.199806	LysM, putative peptidoglycan-binding, domain containing 3
FUCA1	2.2004	fucosidase, alpha-L- 1, tissue
SECTM1	2.2007856	secreted and transmembrane 1
DBX1	2.2025626	developing brain homeobox 1
SHC4	2.2055464	SHC (Src homology 2 domain containing) family, member 4
DHX40	2.2059467	DEAH (Asp-Glu-Ala-His) box polypeptide 40
CCKAR	2.2064834	cholecystokinin A receptor
MEGF6	2.2065082	multiple EGF-like-domains 6
PTN	2.207771	pleiotrophin
ARHGAP15	2.208851	Rho GTPase activating protein 15
AC005524.1	2.20895	ENSG00000283269
SPATA3	2.2105055	spermatogenesis associated 3
ZNF469	2.2105076	zinc finger protein 469
PCDH1	2.211794	protocadherin 1
ERVW-1	2.2133923	endogenous retrovirus group W, member 1
OR2L3	2.2134562	olfactory receptor, family 2, subfamily L, member 3
XDH	2.2140663	xanthine dehydrogenase
OR2W3	2.2157586	olfactory receptor, family 2, subfamily W, member 3
MMP12	2.2170413	matrix metalloproteinase 12 (macrophage elastase)
ADIG	2.2173004	adipogenin
TNFRSF9	2.21779	tumor necrosis factor receptor superfamily, member 9
SPZ1	2.2180195	spermatogenic leucine zipper 1
IL21R	2.218276	interleukin 21 receptor
TMTC4	2.2185957	transmembrane and tetratricopeptide repeat containing 4
FGFR3	2.2223659	fibroblast growth factor receptor 3
DUSP10	2.2229218	dual specificity phosphatase 10

FTSJ1	2.2232232	FtsJ RNA methyltransferase homolog 1 (E. coli)
LOC285692	2.2246485	uncharacterized LOC285692
TANC2	2.2247462	tetratricopeptide repeat, ankyrin repeat and coiled-coil containing 2
CPEB1	2.2252486	cytoplasmic polyadenylation element binding protein 1
SURF4	2.2252638	surfeit 4
KDELC1	2.226539	KDEL (Lys-Asp-Glu-Leu) containing 1
PRAMEF16	2.2281935	PRAME family member 16
TAGLN	2.2285275	transgelin
CGGBP1	2.2302825	CGG triplet repeat binding protein 1
XPR1	2.231706	xenotropic and polytropic retrovirus receptor 1
KCNQ1DN	2.2319355	KCNQ1 downstream neighbor (non-protein coding)
DPRX	2.2321808	divergent-paired related homeobox
FIGLA	2.232566	folliculogenesis specific basic helix-loop-helix
CCDC7	2.2344182	coiled-coil domain containing 7
FAM182B	2.236157	family with sequence similarity 182, member B
OR8A1	2.2385237	olfactory receptor, family 8, subfamily A, member 1
SULT1C2P1	2.2402647	sulfotransferase family, cytosolic, 1C, member 2 pseudogene 1
HSF2BP	2.2407029	heat shock transcription factor 2 binding protein
ZNF286A	2.2439454	Zinc Finger Protein 286A
ZNF554	2.2448933	zinc finger protein 554
lnc-KIAA0087-2	2.245712	lnc-KIAA0087-2:2
PCDHGB3	2.2462869	protocadherin gamma subfamily B, 3
LOC101060498	2.2470083	uncharacterized LOC101060498
CYMP	2.2494032	chymosin pseudogene
TRBV4-1	2.2505684	T Cell Receptor Beta Variable 4-1
OFCC1	2.250591	orofacial cleft 1 candidate 1
PIP	2.2515426	prolactin-induced protein
FZD7	2.252108	frizzled class receptor 7
LOC100507634	2.2529786	uncharacterized LOC100507634
GCNT1	2.2530046	glucosaminyl (N-acetyl) transferase 1, core 2
RIMS3	2.2533023	regulating synaptic membrane exocytosis 3
KDELR3	2.254156	KDEL (Lys-Asp-Glu-Leu) endoplasmic reticulum protein retention receptor 3
EPHA5	2.254365	EPH receptor A5
ASIC1	2.2546375	acid sensing (proton gated) ion channel 1
FHL5	2.2558253	four and a half LIM domains 5

SSX7	2.2574456	synovial sarcoma, X breakpoint 7
VIPR1	2.2581925	vasoactive intestinal peptide receptor 1
WFDC10B	2.2585106	WAP four-disulfide core domain 10B
FAM87B	2.2589304	family with sequence similarity 87, member B
ARHGEF4	2.2617278	Rho guanine nucleotide exchange factor (GEF) 4
FLG	2.2629786	filaggrin
PAGE2B	2.2653651	P antigen family, member 2B
GDF10	2.2662213	growth differentiation factor 10
LINC01446	2.268163	long intergenic non-protein coding RNA 1446
CHN2	2.2683053	chimerin 2
lnc-CLEC14A-1	2.2707179	lnc-CLEC14A-1:2
NTNG2	2.2708797	netrin G2
NEUROG1	2.2711957	neurogenin 1
LOC100128006	2.2728343	uncharacterized LOC100128006
IPO9	2.272996	importin 9
PCDHAC2	2.2752047	protocadherin alpha subfamily C, 2
TRDC	2.2758055	T Cell Receptor Delta Constant
AF096876.1	2.276707	ENSG00000259981
ICAM4	2.278822	intercellular adhesion molecule 4 (Landsteiner-Wiener blood group)
EBF1	2.2798028	early B-cell factor 1
PSG8	2.2800682	pregnancy specific beta-1-glycoprotein 8
C10orf10	2.280221	chromosome 10 open reading frame 10
lnc-AC073343.1-1	2.2832913	lnc-AC073343.1-1:3
AL357140.3	2.2838044	ENSG00000280113
LRRC55	2.2843752	leucine rich repeat containing 55
MAP3K12	2.2848086	mitogen-activated protein kinase kinase kinase 12
OR14A16	2.2848668	olfactory receptor, family 14, subfamily A, member 16
CLIC3	2.2854455	chloride intracellular channel 3
ISM1	2.285938	isthmin 1, angiogenesis inhibitor
LRRC31	2.2881432	leucine rich repeat containing 31
AL390066.1	2.291662	ENSG00000224950
BMF	2.292679	Bcl2 modifying factor
AC092375.1	2.2943122	ATP Binding Cassette Subfamily A Member 3 Pseudogene
CES1P1	2.2946694	carboxylesterase 1 pseudogene 1
ANGPT1	2.2964566	angiopoietin 1

TP53INP2	2.2973402	tumor protein p53 inducible nuclear protein 2
GPR151	2.2974482	G protein-coupled receptor 151
DLG3	2.2977624	discs, large homolog 3 (Drosophila)
KBTBD3	2.3021472	kelch repeat and BTB (POZ) domain containing 3
GALNT16	2.3043456	polypeptide N-acetylgalactosaminyltransferase 16
AC009803.2	2.3048744	ENSG00000258108
TAS2R13	2.3074229	taste receptor, type 2, member 13
CD52	2.3089445	CD52 molecule
LOC654841	2.3089738	uncharacterized LOC654841
LOC100131023	2.3095477	uncharacterized LOC100131023
TPTE2P5	2.311309	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 5
LINC01139	2.313525	long intergenic non-protein coding RNA 1139
EVA1B	2.3147292	eva-1 homolog B (C. elegans)
LOC100132215	2.3188555	uncharacterized LOC100132215
NICN1	2.3193789	nicolin 1
LOC400940	2.319848	uncharacterized LOC400940
TFCP2L1	2.3223865	transcription factor CP2-like 1
VASN	2.3266537	vasorin
GRXCR2	2.3267224	glutaredoxin, cysteine rich 2
SCTR	2.3276777	secretin receptor
GRID1	2.3290772	glutamate receptor, ionotropic, delta 1
LOC102725300	2.3304455	uncharacterized LOC102725300
LINC00921	2.3314888	long intergenic non-protein coding RNA 921
LRIT2	2.3342557	leucine-rich repeat, immunoglobulin-like and transmembrane domains 2
IGHV5-78	2.3346837	Immunoglobulin Heavy Variable 5-78 (Pseudogene)
C4orf22	2.3371766	chromosome 4 open reading frame 22
KRTAP4-4	2.3385706	keratin associated protein 4-4
SMAD6	2.338835	SMAD family member 6
AC068020.1	2.3395486	ENSG00000268955
ENTPD8	2.3423417	ectonucleoside triphosphate diphosphohydrolase 8
IL36B	2.3435435	interleukin 36, beta
RBM15B	2.3435504	RNA binding motif protein 15B
MAGI2	2.3444233	membrane associated guanylate kinase, WW and PDZ domain containing 2
DKFZP434K028	2.3448431	uncharacterized LOC26070
PCDHB10	2.3449001	protocadherin beta 10



lnc-ENPP1-2	2.3479128	lnc-ENPP1-2:1
IL12RB2	2.348554	interleukin 12 receptor, beta 2
LZTS3	2.353542	leucine zipper, putative tumor suppressor family member 3
FIBIN	2.3536773	fin bud initiation factor homolog (zebrafish)
USP27X-AS1	2.3540635	USP27X antisense RNA 1 (head to head)
MYO15A	2.3548448	myosin XVA
LEFTY1	2.3561497	left-right determination factor 1
SS18L1	2.3569582	synovial sarcoma translocation gene on chromosome 18-like 1
FLJ32154	2.3574007	uncharacterized protein FLJ32154
UCK1	2.359747	uridine-cytidine kinase 1
PARM1	2.3597898	prostate androgen-regulated mucin-like protein 1
GPR137C	2.3616843	G protein-coupled receptor 137C
PWAR1	2.3625057	Prader Willi/Angelman region RNA 1
CCNYL2	2.3629968	cyclin Y-like 2, pseudogene
MMP8	2.3632257	matrix metalloproteinase 8 (neutrophil collagenase)
ST7-AS1	2.3634882	ST7 antisense RNA 1
lnc-ACOT12-2	2.364427	lnc-ACOT12-2:1
lnc-LBH-1	2.3645823	lnc-LBH-1:1
GRIK4	2.3654497	glutamate receptor, ionotropic, kainate 4
ADAMTS9	2.3657837	ADAM metalloproteinase with thrombospondin type 1 motif, 9
ADCY10P1	2.365901	adenylate cyclase 10 (soluble) pseudogene 1
HOMER1	2.3675451	homer homolog 1 (Drosophila)
PMCHL1	2.3688912	pro-melanin-concentrating hormone-like 1, pseudogene
RHBDL2	2.3692987	rhomboid, veinlet-like 2 (Drosophila)
LOC100147773	2.3701105	uncharacterized LOC100147773
AL161938.1	2.3730478	ENSG00000261411
LOC285000	2.37378	uncharacterized LOC285000
ROM1	2.373988	retinal outer segment membrane protein 1
PJA2	2.3758953	praja ring finger 2, E3 ubiquitin protein ligase
CCR4	2.3783832	chemokine (C-C motif) receptor 4
lnc-GFAP-1	2.3804991	lnc-GFAP-1:1
PTGS1	2.3809154	prostaglandin-endoperoxide synthase 1 (prostaglandin G/H synthase and cyclooxygenase)
HIVEP3	2.3811767	human immunodeficiency virus type I enhancer binding protein 3
PTPRS	2.382422	protein tyrosine phosphatase, receptor type, S
KRTDAP	2.383427	keratinocyte differentiation-associated protein

EQTN	2.3835669	equatorin, sperm acrosome associated
LOXL1	2.3842874	lysyl oxidase-like 1
lnc-PGR-1	2.3884203	lnc-PGR-1:1
SEPT5	2.3891006	septin 5
ARHGEF40	2.3893664	Rho guanine nucleotide exchange factor (GEF) 40
MBNL1-AS1	2.390983	MBNL1 antisense RNA 1
GDF11	2.3927298	growth differentiation factor 11
ADCY5	2.3949146	adenylate cyclase 5
HLA-DMB	2.39757	major histocompatibility complex, class II, DM beta
PTGFRN	2.397829	prostaglandin F2 receptor inhibitor
COL5A1	2.4022624	collagen, type V, alpha 1
DSEL	2.4026618	dermatan sulfate epimerase-like
CERCAM	2.4031773	cerebral endothelial cell adhesion molecule
C14orf180	2.403878	chromosome 14 open reading frame 180
EDNRB	2.4040966	endothelin receptor type B
DOPEY2	2.404877	dopey family member 2
KIAA1524	2.4079916	KIAA1524
GNGT2	2.4099236	guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 2
SERPINB10	2.4147778	serpin peptidase inhibitor, clade B (ovalbumin), member 10
DYNLRB2	2.416458	dynein, light chain, roadblock-type 2
CPE	2.416589	carboxypeptidase E
OR1L8	2.4171376	olfactory receptor, family 1, subfamily L, member 8
GP2	2.418341	glycoprotein 2 (zymogen granule membrane)
TGFB2	2.4197676	transforming growth factor, beta 2
MUC13	2.42288	mucin 13, cell surface associated
MIR181A1HG	2.42334	MIR181A1 host gene (non-protein coding)
HLA-DRB4	2.4235256	major histocompatibility complex, class II, DR beta 4
MIRLET7DHG	2.4235814	MIRLET7D host gene (non-protein coding)
CDRT15	2.4237752	CMT1A duplicated region transcript 15
CARF	2.4239893	calcium responsive transcription factor
lnc-ENDOD1-1	2.4244826	lnc-ENDOD1-1:7
SEMA6A	2.42618	sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6A
P3H4	2.4284647	prolyl 3-hydroxylase family member 4 (non-enzymatic)
SYCP2	2.4293094	synaptonemal complex protein 2
NPY2R	2.43185	neuropeptide Y receptor Y2

AC087491.1	2.432774	ENSG00000214546
GDAP1	2.433163	ganglioside induced differentiation associated protein 1
COLCA1	2.43432	colorectal cancer associated 1
NACAD	2.4364522	NAC alpha domain containing
TTPA	2.4369237	tocopherol (alpha) transfer protein
FSTL1	2.4389846	folliculin-like 1
RBPM5	2.4406312	RNA binding protein with multiple splicing 2
SLITRK5	2.4407063	SLIT and NTRK-like family, member 5
XLOC_12_011265	2.4470456	XLOC_12_011265
SEC1P	2.4483933	secretory blood group 1, pseudogene
GGTLC1	2.448972	gamma-glutamyltransferase light chain 1
DOK6	2.4494343	docking protein 6
ZNF425	2.4526942	zinc finger protein 425
LOC389895	2.4530973	chromosome 16 open reading frame 72-like
FAM210B	2.4549513	family with sequence similarity 210, member B
TMEM184B	2.456149	transmembrane protein 184B
MICALCL	2.4576266	MICAL C-terminal like
XLOC_12_013457	2.458091	XLOC_12_013457
CYB5R4	2.462049	cytochrome b5 reductase 4
TIMP3	2.4633458	TIMP metalloproteinase inhibitor 3
CRISPLD2	2.4649372	cysteine-rich secretory protein LCCL domain containing 2
BX284668.2	2.4651153	Transfer RNA-Glu (TTC) 3-1
FAM47A	2.465261	family with sequence similarity 47, member A
LINC00342	2.4666128	Long Intergenic Non-Protein Coding RNA 342
AC087203.1	2.4669728	Ubiquitin Specific Peptidase 17-Like Family Member 2 Pseudogene
HTN1	2.4684918	histatin 1
C19orf45	2.4725394	chromosome 19 open reading frame 45
LOC100129275	2.4747229	uncharacterized LOC100129275
FO393408.1	2.4748611	Uncharacterized LOC105373343
C8orf88	2.4754007	chromosome 8 open reading frame 88
DAPK2	2.4755101	death-associated protein kinase 2
IGKV1-16	2.475651	Immunoglobulin Kappa Variable 1-16
AASDH	2.4766886	aminoadipate-semialdehyde dehydrogenase
LOC283674	2.4770114	uncharacterized LOC283674
BMP3	2.4784222	bone morphogenetic protein 3

MORN5	2.4788024	MORN repeat containing 5
KLHL42	2.479367	kelch-like family member 42
LOC102725407	2.480939	uncharacterized LOC102725407
LIMCH1	2.4823024	LIM and calponin homology domains 1
SAMD15	2.4836905	sterile alpha motif domain containing 15
PDE4DIP	2.4838405	Phosphodiesterase 4D Interacting Protein
lnc-CHADL-3	2.4889925	lnc-CHADL-3:1
GPR68	2.4901862	G protein-coupled receptor 68
GGACT	2.4910038	gamma-glutamylamine cyclotransferase
IL10RB	2.491403	interleukin 10 receptor, beta
IGHV7-81	2.4926283	Immunoglobulin Heavy Variable 7-81 (Non-Functional)
TMEM200B	2.4928386	transmembrane protein 200B
TMEM191A	2.4955287	transmembrane protein 191A (pseudogene)
COL18A1	2.4958801	collagen, type XVIII, alpha 1
ANXA8L1	2.4959526	annexin A8-like 1
LINC01503	2.49697	long intergenic non-protein coding RNA 1503
PITX2	2.498814	paired-like homeodomain 2
LOC388242	2.5025651	coiled-coil domain containing 101 pseudogene
CSPG5	2.5097446	chondroitin sulfate proteoglycan 5 (neuroglycan C)
LINC00551	2.510108	long intergenic non-protein coding RNA 551
PLXDC2	2.512773	plexin domain containing 2
FZD10-AS1	2.513522	FZD10 antisense RNA 1 (head to head)
RAB27A	2.517787	RAB27A, member RAS oncogene family
GPR37	2.5179849	G protein-coupled receptor 37 (endothelin receptor type B-like)
TMEM30C	2.518716	transmembrane protein 30C
CDK14	2.519715	cyclin-dependent kinase 14
CCDC148	2.521116	coiled-coil domain containing 148
FAM170B	2.5211568	family with sequence similarity 170, member B
ADAMTS12	2.521478	ADAM metallopeptidase with thrombospondin type 1 motif, 12
C1orf195	2.5215478	chromosome 1 open reading frame 195
TP53INP1	2.5224965	tumor protein p53 inducible nuclear protein 1
SPANXN4	2.5243053	SPANX family, member N4
SMIM23	2.5246112	small integral membrane protein 23
CXCR4	2.528309	chemokine (C-X-C motif) receptor 4
SLC41A1	2.5283477	solute carrier family 41 (magnesium transporter), member 1

LPAR1	2.5293455	lysophosphatidic acid receptor 1
KRTAP4-9	2.5309665	keratin associated protein 4-9
FAM135A	2.531112	family with sequence similarity 135, member A
CSF1R	2.5365539	colony stimulating factor 1 receptor
CA12	2.5368717	carbonic anhydrase XII
LSM14A	2.5396929	LSM14A, SCD6 homolog A ( <i>S. cerevisiae</i> )
EPB41L1	2.5399096	erythrocyte membrane protein band 4.1-like 1
GNB1	2.5425284	guanine nucleotide binding protein (G protein), beta polypeptide 1
AFAPIL2	2.5448	actin filament associated protein 1-like 2
lnc-AC127496.3-7	2.5517685	lnc-AC127496.3-7:1
RFPL3	2.5533264	ret finger protein-like 3
DPY19L2	2.5542033	dpy-19-like 2 ( <i>C. elegans</i> )
PACS2	2.5545113	phosphofurin acidic cluster sorting protein 2
CHPF	2.5548964	chondroitin polymerizing factor
COL6A3	2.5556498	collagen, type VI, alpha 3
OR6K4P	2.5569472	Olfactory Receptor Family 6 Subfamily K Member 4 Pseudogene
GAB1	2.558024	GRB2-associated binding protein 1
LOC100133106	2.5600722	VCEW9374
PCDHB12	2.5644565	protocadherin beta 12
HOXC5	2.566734	homeobox C5
LOC100996634	2.573409	transmembrane protein FLJ37396
CAPN5	2.5772412	calpain 5
GPR162	2.5799105	G protein-coupled receptor 162
LOC730668	2.582304	dynein heavy chain -like pseudogene
LOC645188	2.5832753	uncharacterized LOC645188
LINC00242	2.58772	long intergenic non-protein coding RNA 242
TMEM87B	2.5901966	transmembrane protein 87B
RBPJ	2.5933568	recombination signal binding protein for immunoglobulin kappa J region
LBP	2.5968761	lipopolysaccharide binding protein
C1orf94	2.5973837	chromosome 1 open reading frame 94
SLC15A5	2.6032019	solute carrier family 15, member 5
PCDHB18P	2.6056912	protocadherin beta 18 pseudogene
ASPA	2.6071997	aspartoacylase
ZMAT4	2.6092558	zinc finger, matrin-type 4
TKTL2	2.6095424	transketolase-like 2

C3orf80	2.6097672	chromosome 3 open reading frame 80
CBX3P2	2.610207	chromobox homolog 3 pseudogene 2
LOC100509445	2.6113515	uncharacterized LOC100509445
OMG	2.6196084	oligodendrocyte myelin glycoprotein
LOC401463	2.6205504	uncharacterized LOC401463
MFSD6	2.6223938	major facilitator superfamily domain containing 6
HMG2P31	2.627972	High Mobility Group Nucleosomal Binding Domain 2 Pseudogene 31
LINC00314	2.6310096	long intergenic non-protein coding RNA 314
GOLGA4	2.633813	golgin A4
lnc-SLC6A12-6	2.6361055	lnc-SLC6A12-6:2
LCOR	2.6361296	ligand dependent nuclear receptor corepressor
RAPGEF4	2.6377084	Rap guanine nucleotide exchange factor (GEF) 4
AC108174.1	2.6378508	Uncharacterized LOC105379051
ZNF37BP	2.6427763	zinc finger protein 37B, pseudogene
SPINK4	2.6469758	serine peptidase inhibitor, Kazal type 4
LOC100134368	2.6499455	uncharacterized LOC100134368
CDH16	2.6516356	cadherin 16, KSP-cadherin
CLEC4C	2.6519253	C-type lectin domain family 4, member C
JPX	2.6526687	JPX transcript, XIST activator (non-protein coding)
TNNT1	2.6529222	troponin T type 1 (skeletal, slow)
ABCB4	2.6550307	ATP-binding cassette, sub-family B (MDR/TAP), member 4
AC092376.1	2.6574192	ENSG00000261722
PTBP3	2.6574311	polypyrimidine tract binding protein 3
SFT2D3	2.6574337	SFT2 domain containing 3
SLC2A12	2.6582172	solute carrier family 2 (facilitated glucose transporter), member 12
LOC146795	2.6616635	uncharacterized LOC146795
LINC00475	2.6696274	long intergenic non-protein coding RNA 475
HTR1E	2.6717803	5-hydroxytryptamine (serotonin) receptor 1E, G protein-coupled
PPAPDC3	2.674712	phosphatidic acid phosphatase type 2 domain containing 3
DLL1	2.674789	delta-like 1 (Drosophila)
SEMA3B	2.680552	sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B
LINC00701	2.6825755	long intergenic non-protein coding RNA 701
ELOVL7	2.6841197	ELOVL fatty acid elongase 7
FBXO32	2.6861875	F-box protein 32
FAM26E	2.687918	family with sequence similarity 26, member E

SUGT1P3	2.6929328	SUGT1 pseudogene 3
COLQ	2.6944149	collagen-like tail subunit (single strand of homotrimer) of asymmetric acetylcholinesterase
IGKV1D-13	2.695293	immunoglobulin kappa variable 1D-13
NKX2-1	2.7011387	NK2 homeobox 1
PNMA2	2.7050095	paraneoplastic Ma antigen 2
LTB	2.705326	lymphotoxin beta (TNF superfamily, member 3)
DNM1P17	2.7058616	Dynamin 1 Pseudogene 17
CLMP	2.710373	CXADR-like membrane protein
ADORA1	2.7124672	adenosine A1 receptor
TRGV8	2.732161	T Cell Receptor Gamma Variable 8
SLC26A4	2.7329931	solute carrier family 26 (anion exchanger), member 4
PRPH2	2.7336376	peripherin 2 (retinal degeneration, slow)
RAB14	2.7337313	RAB14, member RAS oncogene family
SPCS2	2.7347436	signal peptidase complex subunit 2 homolog ( <i>S. cerevisiae</i> )
PRAMEF22	2.7371013	PRAME family member 22
SPOCD1	2.7374566	SPOC domain containing 1
C3orf79	2.7405608	chromosome 3 open reading frame 79
PRDM5	2.7423851	PR/SET Domain 5
OR10A7	2.745613	olfactory receptor, family 10, subfamily A, member 7
HPVC1	2.746917	human papillomavirus (type 18) E5 central sequence-like 1
KIAA1456	2.7493675	KIAA1456
ALCAM	2.7526155	activated leukocyte cell adhesion molecule
PTGFR	2.7605264	prostaglandin F receptor (FP)
NDUFA5	2.765234	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 5
GALNT5	2.7657201	polypeptide N-acetylgalactosaminyltransferase 5
TRDN	2.765774	triadin
ERN2	2.7767472	endoplasmic reticulum to nucleus signaling 2
PLA2G2F	2.7775586	phospholipase A2, group IIF
LINC01552	2.7809253	long intergenic non-protein coding RNA 1552
FREM2	2.7891536	FRAS1 related extracellular matrix protein 2
TRAV19	2.7914386	T Cell Receptor Alpha Variable 19
LOC100132781	2.7942152	cyclin Y-like 1 pseudogene
LOC339988	2.7943356	uncharacterized LOC339988
AADAC	2.8002687	arylacetamide deacetylase
lnc-HERC1-1	2.8051546	lnc-HERC1-1:9

THG1L	2.8171828	tRNA-histidine guanylyltransferase 1-like ( <i>S. cerevisiae</i> )
PKIB	2.8180096	protein kinase (cAMP-dependent, catalytic) inhibitor beta
UBE2Q2P6	2.821349	Ubiquitin Conjugating Enzyme E2 Q2 Pseudogene 6
OR1J2	2.8232272	olfactory receptor, family 1, subfamily J, member 2
C17orf102	2.8259897	chromosome 17 open reading frame 102
LSP1P4	2.8271337	LSP1 pseudogene 4
LINC02279	2.828314	Long Intergenic Non-Protein Coding RNA 2279
NOTO	2.8343856	notochord homeobox
LYZL2	2.836221	lysozyme-like 2
ADRA1B	2.8374984	adrenoceptor alpha 1B
HSP90AB4P	2.8435802	heat shock protein 90kDa alpha (cytosolic), class B member 4, pseudogene
IL1B	2.8455312	interleukin 1, beta
PATE1	2.846714	prostate and testis expressed 1
DMP1	2.8498406	dentin matrix acidic phosphoprotein 1
NLRP8	2.8539493	NLR family, pyrin domain containing 8
DEFA4	2.8567367	defensin, alpha 4, corticostatin
MATN3	2.8589327	matrilin 3
IGSF6	2.86278	immunoglobulin superfamily, member 6
REEP2	2.8645918	receptor accessory protein 2
SELE	2.869732	selectin E
GAS6-AS1	2.8722892	GAS6 antisense RNA 1
TSPAN2	2.872598	tetraspanin 2
TRBV15	2.8733854	T Cell Receptor Beta Variable 15
LILRB2	2.8769782	leukocyte immunoglobulin-like receptor, subfamily B (with TM and ITIM domains), member 2
TOP1MT	2.8793468	topoisomerase (DNA) I, mitochondrial
TPTEP1	2.883613	TPTE Pseudogene 1
UST	2.886749	uronyl-2-sulfotransferase
C9orf170	2.8891284	chromosome 9 open reading frame 170
TBC1D3P6	2.8896213	TBC1 Domain Family Member 3 Pseudogene 6
ACKR2	2.8935637	atypical chemokine receptor 2
KDM5A	2.8939307	lysine (K)-specific demethylase 5A
TGM5	2.8983448	transglutaminase 5
RIPPLY2	2.9001846	rippy transcriptional repressor 2
CA5AP1	2.9067092	Carbonic Anhydrase 5A Pseudogene 1
TRIM31	2.9151268	tripartite motif containing 31



SPAM1	2.9156234	sperm adhesion molecule 1 (PH-20 hyaluronidase, zona pellucida binding)
TAS2R3	2.9156547	taste receptor, type 2, member 3
LUZP4	2.9159849	leucine zipper protein 4
MRC2	2.9165695	mannose receptor, C type 2
CHST15	2.9180293	carbohydrate (N-acetylgalactosamine 4-sulfate 6-O) sulfotransferase 15
PRR9	2.9201412	proline rich 9
IDI2	2.9218504	isopentenyl-diphosphate delta isomerase 2
CCDC146	2.922357	coiled-coil domain containing 146
NOG	2.9259195	noggin
ZBTB47	2.926235	zinc finger and BTB domain containing 47
C18orf63	2.9266639	chromosome 18 open reading frame 63
RYR3	2.933827	ryanodine receptor 3
GPR110	2.9361603	G protein-coupled receptor 110
HLX-AS1	2.937973	HLX antisense RNA 1
BHMT	2.9402304	betaine--homocysteine S-methyltransferase
MYL10	2.94085	myosin, light chain 10, regulatory
TSLP	2.9421382	thymic stromal lymphopoietin
CLEC4A	2.9422646	C-type lectin domain family 4, member A
SFTA3	2.9459448	surfactant associated 3
YPEL2	2.9529836	yippee-like 2 (Drosophila)
OR14A2	2.9546604	Olfactory Receptor Family 14 Subfamily A Member 2
YWHAH	2.9653962	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, eta
XLOC_I2_015641	2.9685156	XLOC_I2_015641
KCNS3	2.9738786	potassium voltage-gated channel, modifier subfamily S, member 3
PHLDA1	2.975853	pleckstrin homology-like domain, family A, member 1
ANK3	2.976545	Ankyrin 3
LPP-AS2	2.9788485	LPP antisense RNA 2
SMIM17	2.98036	small integral membrane protein 17
SNHG18	2.9953628	small nucleolar RNA host gene 18 (non-protein coding)
ASCL4	2.997351	achaete-scute family bHLH transcription factor 4
MATN2	2.999739	matrilin 2
AL121672.1	3.006664	ENSG00000231010
SAR1B	3.0070133	secretion associated, Ras related GTPase 1B
SPTBN1	3.0129585	spectrin, beta, non-erythrocytic 1
RUNX2	3.017161	runt-related transcription factor 2

PSAT1P2	3.0267015	Phosphoserine Aminotransferase 1 Pseudogene 2
ZFYVE26	3.0278652	Zinc Finger FYVE-Type Containing 26
PDGFRL	3.0310493	platelet-derived growth factor receptor-like
KCNJ15	3.0337784	potassium channel, inwardly rectifying subfamily J, member 15
C8ORF16	3.0366807	chromosome 8 open reading frame 16
OR56A1	3.0451727	olfactory receptor, family 56, subfamily A, member 1
IGFBP7	3.048245	insulin-like growth factor binding protein 7
ANO10	3.0507061	anoctamin 10
XKRY2	3.053792	XK, Kell blood group complex subunit-related, Y-linked 2
SHROOM2	3.0554147	shroom family member 2
AC084357.1	3.0604146	Ribosomal Protein S27 Pseudogene 24
CCDC175	3.0628254	coiled-coil domain containing 175
lnc-C1orf31-1	3.0629582	lnc-C1orf31-1:1
PLEKHA7	3.0635529	pleckstrin homology domain containing, family A member 7
FAM83B	3.07003	family with sequence similarity 83, member B
FABP7	3.0712824	fatty acid binding protein 7, brain
lnc-DPP4-2	3.080887	lnc-DPP4-2:3
MAEL	3.081873	maelstrom spermatogenic transposon silencer
AP003532.1	3.0844283	ENSG00000254591
MFSD6L	3.088498	major facilitator superfamily domain containing 6-like
C7orf33	3.0889208	chromosome 7 open reading frame 33
KLHL40	3.1023307	kelch-like family member 40
RLN1	3.103021	relaxin 1
LOC100129973	3.1062882	uncharacterized LOC100129973
HDX	3.1291654	highly divergent homeobox
HOXC10	3.1340876	homeobox C10
OR1B1	3.134997	olfactory receptor, family 1, subfamily B, member 1 (gene/pseudogene)
LINC01965	3.141224	Long Intergenic Non-Protein Coding RNA 1965
EFEMP1	3.1457443	EGF containing fibulin-like extracellular matrix protein 1
SKOR1	3.1519325	SKI family transcriptional corepressor 1
NKX2-6	3.1591618	NK2 homeobox 6
LOC100128320	3.159246	uncharacterized LOC100128320
PRRT2	3.1597252	proline-rich transmembrane protein 2
OLFML2A	3.161893	olfactomedin-like 2A
OR9Q2	3.1634395	olfactory receptor, family 9, subfamily Q, member 2

LOC100130857	3.1637626	uncharacterized LOC100130857
SDK2	3.166907	sidekick cell adhesion molecule 2
DCAF8L2	3.1697438	DDB1 and CUL4 associated factor 8-like 2
METTL21EP	3.1701472	methyltransferase like 21E, pseudogene
LINC00161	3.171417	long intergenic non-protein coding RNA 161
LY75	3.1759834	lymphocyte antigen 75
ARSI	3.1771636	arylsulfatase family, member I
C12orf77	3.1777537	chromosome 12 open reading frame 77
LOC154872	3.1801856	putative uncharacterized protein LOC154872
lnc-GUSB-5	3.203964	lnc-GUSB-5:1
EDN2	3.2060335	endothelin 2
CCDC180	3.2089424	coiled-coil domain containing 180
DEFA11P	3.2098315	defensin, alpha 11 pseudogene
F2R	3.2098463	coagulation factor II (thrombin) receptor
CTSO	3.2214363	cathepsin O
SCAMP1	3.2266674	secretory carrier membrane protein 1
DHH	3.2294707	desert hedgehog
PAPOLB	3.2360256	poly(A) polymerase beta (testis specific)
SH3PXD2A	3.2394187	SH3 and PX domains 2A
TBC1D32	3.2397742	TBC1 domain family, member 32
TNNI3K	3.2532063	TNNI3 Interacting Kinase
S1PR5	3.2562506	sphingosine-1-phosphate receptor 5
NUDT15	3.2718916	nudix (nucleoside diphosphate linked moiety X)-type motif 15
TMEM50B	3.2745335	transmembrane protein 50B
NAPIL5	3.280882	nucleosome assembly protein 1-like 5
KREMEN2	3.2891335	kringle containing transmembrane protein 2
CROCCP3	3.292982	ciliary rootlet coiled-coil, rootletin pseudogene 3
C10orf111	3.2933142	chromosome 10 open reading frame 111
TPTE2P3	3.295769	transmembrane phosphoinositide 3-phosphatase and tensin homolog 2 pseudogene 3
FRMPD4	3.3042116	FERM and PDZ domain containing 4
WNT5A	3.307835	wingless-type MMTV integration site family, member 5A
DEFB105B	3.307857	defensin, beta 105B
WISP1	3.3112485	WNT1 inducible signaling pathway protein 1
LINC02511	3.3131337	Long Intergenic Non-Protein Coding RNA 2511
KLHL1	3.3179526	kelch-like family member 1

TMOD2	3.3190176	tropomodulin 2 (neuronal)
FAM229B	3.3302648	family with sequence similarity 229, member B
OR52E4	3.336105	olfactory receptor, family 52, subfamily E, member 4
LINC00323	3.3413599	long intergenic non-protein coding RNA 323
LYPD1	3.3458593	LY6/PLAUR domain containing 1
EXOC5P1	3.3469677	Exocyst Complex Component 5 Pseudogene 1
UTS2B	3.356932	urotensin 2B
LOC100129940	3.3694568	uncharacterized LOC100129940
UNC93A	3.3830168	unc-93 homolog A (C. elegans)
ABCC6P1	3.3854003	ATP-binding cassette, sub-family C, member 6 pseudogene 1 (functional)
TRIM22	3.4092433	tripartite motif containing 22
SSU72P1	3.4449313	SSU72 Pseudogene 1
FLJ45079	3.4449866	FLJ45079 protein
MAMDC2	3.4476368	MAM domain containing 2
KRCC1	3.4484994	lysine-rich coiled-coil 1
LOC730100	3.449156	uncharacterized LOC730100
LOC100506351	3.4554586	uncharacterized LOC100506351
LURAP1L	3.4584923	leucine rich adaptor protein 1-like
WDR45	3.468709	WD Repeat Domain 45
ZFP2	3.484666	ZFP2 zinc finger protein
GPX7	3.4883242	glutathione peroxidase 7
PMEPA1	3.4958055	prostate transmembrane protein, androgen induced 1
NRK	3.505012	Nik related kinase
DEFT1P	3.5075002	defensin, theta 1 pseudogene
PQLC2L	3.5173633	PQ loop repeat containing 2-like
LOC101928561	3.524916	uncharacterized LOC101928561
XLOC_12_013461	3.5250218	XLOC_12_013461
AC245177.1	3.5275598	ENSG00000179028
BMPER	3.5313861	BMP binding endothelial regulator
lnc-ZNF395-2	3.5355256	lnc-ZNF395-2:1
RFX6	3.5445354	regulatory factor X, 6
ARNT2	3.5464509	aryl-hydrocarbon receptor nuclear translocator 2
FGF19	3.547985	fibroblast growth factor 19
SPOPL	3.556842	speckle-type POZ protein-like
MED28	3.5575898	mediator complex subunit 28

C14orf37	3.560733	chromosome 14 open reading frame 37
SLC44A3	3.5672863	solute carrier family 44, member 3
LINC01483	3.5699594	long intergenic non-protein coding RNA 1483
MC4R	3.5860322	melanocortin 4 receptor
AC051619.1	3.6216815	ENSG00000235390
MBD5	3.6246986	methyl-CpG binding domain protein 5
OR4C13	3.6257982	olfactory receptor, family 4, subfamily C, member 13
LOC286272	3.6367395	uncharacterized LOC286272
LOC100506393	3.651552	uncharacterized LOC100506393
AMY1C	3.6616604	amylase, alpha 1C (salivary)
C14orf28	3.663739	chromosome 14 open reading frame 28
FAM171B	3.6690044	family with sequence similarity 171, member B
FAM179A	3.6712594	family with sequence similarity 179, member A
SHISA2	3.6771104	shisa family member 2
PNPLA1	3.6834636	patatin-like phospholipase domain containing 1
HIPK3	3.68393	homeodomain interacting protein kinase 3
VSX2	3.6863794	visual system homeobox 2
CSNK1G2-AS1	3.70166	CSNK1G2 antisense RNA 1
CYP4X1	3.70192	cytochrome P450, family 4, subfamily X, polypeptide 1
LOC286059	3.714226	tumor necrosis factor receptor superfamily, member 10d, decoy with truncated death domain pseudogene
C8orf89	3.7142854	chromosome 8 open reading frame 89
STEAP3	3.7175224	STEAP family member 3, metalloredutase
SPOCK3	3.7189565	sparc/osteonectin, cwcv and kazal-like domains proteoglycan (testican) 3
CHSY3	3.7246103	chondroitin sulfate synthase 3
OR8J1	3.7411997	olfactory receptor, family 8, subfamily J, member 1
PGLYRP4	3.8024857	peptidoglycan recognition protein 4
KRT222	3.8197422	keratin 222, type II
LRRC15	3.8273253	leucine rich repeat containing 15
DISC1	3.8301725	disrupted in schizophrenia 1
VGLL1	3.8621068	vestigial-like family member 1
RAB15	3.8723693	RAB15, member RAS oncogene family
NMNAT2	3.8738275	nicotinamide nucleotide adenylyltransferase 2
ODF3L1	3.8789337	outer dense fiber of sperm tails 3-like 1
DCDC2B	3.8816023	doublecortin domain containing 2B
TK2	3.91665	thymidine kinase 2, mitochondrial

CELA1	3.917654	chymotrypsin-like elastase family, member 1
ANKRD34A	3.9283915	ankyrin repeat domain 34A
LOC100101478	3.9366157	H2B histone family, member M pseudogene
lnc-LGALS14-1	3.9443195	lnc-LGALS14-1:1
KCNRG	3.9563851	potassium channel regulator
LOC100128198	3.978158	uncharacterized LOC100128198
CMTM2	3.9867098	CKLF-like MARVEL transmembrane domain containing 2
QRFPR	3.990708	pyroglutamylated RFamide peptide receptor
HTRA4	4.0042653	HtrA serine peptidase 4
NCKAP1L	4.027715	NCK-associated protein 1-like
LINC00877	4.03304	long intergenic non-protein coding RNA 877
DEFB106B	4.033363	defensin, beta 106B
CYP26B1	4.036697	cytochrome P450, family 26, subfamily B, polypeptide 1
FLJ46120	4.0492373	uncharacterized LOC647008
LOC100128727	4.0951085	uncharacterized LOC100128727
COL7A1	4.097341	collagen, type VII, alpha 1
KEL	4.130324	Kell blood group, metallo-endopeptidase
MKRN9P	4.142573	makorin ring finger protein 9, pseudogene
TMEM86A	4.162809	transmembrane protein 86A
lnc-ADII-1	4.2209363	lnc-ADII-1:1
PPFIA2	4.2414436	protein tyrosine phosphatase, receptor type, f polypeptide (PTPRF), interacting protein (liprin), alpha 2
PLCE1-AS1	4.253688	PLCE1 antisense RNA 1
FER1L6	4.3142147	fer-1-like family member 6
UBTFL7	4.338758	UBTF Like 7 (Pseudogene)
CCKBR	4.370973	cholecystokinin B receptor
CNN1	4.385418	calponin 1, basic, smooth muscle
LINC00661	4.4008455	long intergenic non-protein coding RNA 661
PDGFRB	4.448225	platelet-derived growth factor receptor, beta polypeptide
lnc-C11orf39-3	4.448456	lnc-C11orf39-3:1
DYDC1	4.5047107	DPY30 domain containing 1
TGM3	4.517343	transglutaminase 3
RPL36AP33	4.532146	ribosomal protein L36a pseudogene 33
ANKRD30B	4.6248837	ankyrin repeat domain 30B
SAA2	4.6284084	serum amyloid A2
CAPN11	4.6287346	calpain 11

DMRTC1	4.6382923	DMRT-like family C1
RHOJ	4.6399584	ras homolog family member J
OPRK1	4.6880865	opioid receptor, kappa 1
PSG9	4.7019577	pregnancy specific beta-1-glycoprotein 9
TRBV17	4.7145433	T Cell Receptor Beta Variable 17 (Non-Functional)
C18orf65	4.8299627	chromosome 18 open reading frame 65
SPICE1	4.847913	spindle and centriole associated protein 1
MYH16	4.84978	myosin, heavy chain 16 pseudogene
OR2M5	4.8564897	olfactory receptor, family 2, subfamily M, member 5
EMP2	4.887733	epithelial membrane protein 2
SLC16A2	4.9565334	solute carrier family 16, member 2 (thyroid hormone transporter)
PRSS35	4.958851	protease, serine, 35
DMC1	4.982222	DNA meiotic recombinase 1
GPA33	5.007998	glycoprotein A33 (transmembrane)
AP003385.2	5.046489	ENSG00000254883
LINC00886	5.052836	long intergenic non-protein coding RNA 886
ZNF846	5.0721107	zinc finger protein 846
TEPP	5.080538	testis, prostate and placenta expressed
HSPA12B	5.095139	heat shock 70kD protein 12B
POU1F1	5.138877	POU class 1 homeobox 1
LOC100133669	5.3588076	uncharacterized LOC100133669
TLR8-AS1	5.556142	TLR8 antisense RNA 1
C5	5.557389	complement component 5
GBX2	5.6163473	gastrulation brain homeobox 2
MACC1	5.6470103	metastasis associated in colon cancer 1
UPK2	5.697173	uroplakin 2
C19orf71	5.7237444	chromosome 19 open reading frame 71
lnc-PIEZO2-7	5.7911763	lnc-PIEZO2-7:1
FAM138E	5.813194	family with sequence similarity 138, member E
TRBV12-4	6.0752926	T Cell Receptor Beta Variable 12-4
LOC399715	6.3741317	uncharacterized LOC399715
EDIL3	6.426505	EGF-like repeats and discoidin I-like domains 3
CCDC83	6.535713	coiled-coil domain containing 83
OXTR	6.880251	oxytocin receptor
KCNJ2	6.897635	potassium channel, inwardly rectifying subfamily J, member 2

AC091951.1	7.056184	Golgin A6 Family-Like 1 Pseudogene
LOC441601	7.0925455	septin 7 pseudogene
CYS1	7.133363	cystin 1
PRR31	7.4001007	proline rich 31
EFCAB8	7.4731894	EF-hand calcium binding domain 8
EPB41	7.9893103	erythrocyte membrane protein band 4.1
CCR1	8.17341	chemokine (C-C motif) receptor 1
SLC9A9	8.250493	solute carrier family 9, subfamily A (NHE9, cation proton antiporter 9), member 9
PSG5	8.345928	pregnancy specific beta-1-glycoprotein 5
RTP3	8.802638	receptor (chemosensory) transporter protein 3
LOC100422737	8.959873	uncharacterized LOC100422737
OR52L1	9.157012	olfactory receptor, family 52, subfamily L, member 1
OR2D2	9.507756	olfactory receptor, family 2, subfamily D, member 2
ITGB8	10.241307	integrin, beta 8
AL137100.1	10.747416	Zinc Finger CCHC-Type Containing 7 Pseudogene
FLI1	10.819079	Fli-1 proto-oncogene, ETS transcription factor
OR51Q1	14.845551	olfactory receptor, family 51, subfamily Q, member 1 (gene/pseudogene)
AC093689.1	15.492573	ENSG00000250597
CEMIP	20.366045	cell migration inducing protein, hyaluronan binding



Anexo C - Legenda dos símbolos apresentados nas análises *in silico* pelo software Metacore®

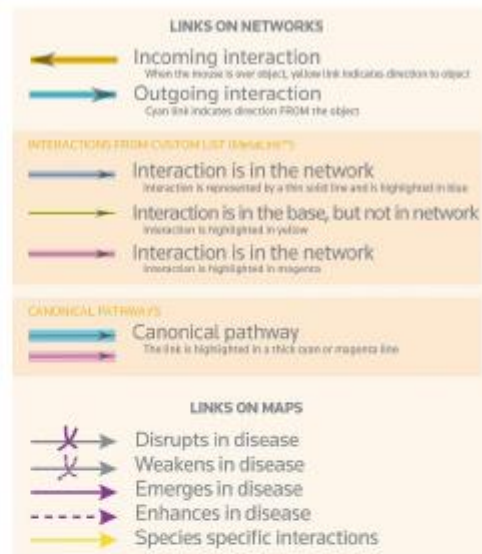
### User Data

NETWORKS	MAPS
Up-regulated (+) Object has user data with positive value	
Down-regulated (-) Object has user data with negative value	
Mixed-signal (+ /-) Object has user data with both positive and negative values	
Gene variants Object has user data with gene variants	
Mixed data Object has user data with both expression values and gene variants	

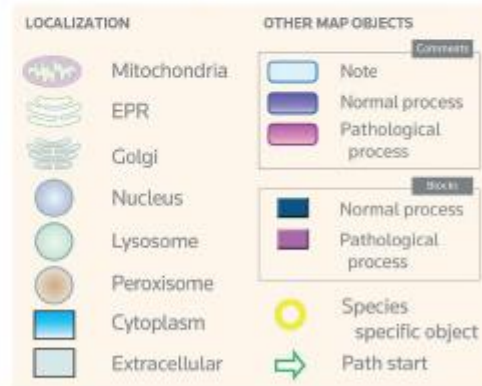
### Network Objects

ENZYMES	GENERIC CLASSES
<p><b>Generic enzyme</b></p> <p><b>KINASE</b></p> <ul style="list-style-type: none"> <li>Generic kinase</li> <li>Protein kinase</li> <li>Lipid kinase</li> </ul> <p><b>PHOSPHATASE</b></p> <ul style="list-style-type: none"> <li>Generic phosphatase</li> <li>Protein phosphatase</li> <li>Lipid phosphatase</li> </ul> <p><b>PHOSPHOLIPASE</b></p> <ul style="list-style-type: none"> <li>Generic phospholipase</li> </ul> <p><b>PROTEASE</b></p> <ul style="list-style-type: none"> <li>Generic protease</li> <li>Metalloprotease</li> </ul> <p><b>GTASE</b></p> <ul style="list-style-type: none"> <li>G-alpha</li> <li>RAS - superfamily</li> </ul>	<ul style="list-style-type: none"> <li>Receptor ligand</li> <li>Transcription factor</li> <li>Protein</li> <li>Compound</li> <li>Predicted metabolite or user's structure</li> <li>Inorganic ion</li> <li>Reaction</li> <li>DNA</li> <li>RNA</li> <li>Generic binding protein</li> <li>Cell membrane glycoprotein</li> </ul>
<p><b>CHANNELS/TRANSPORTERS</b></p> <ul style="list-style-type: none"> <li>Generic channel</li> <li>Ligand-gated ion channel</li> <li>Voltage-gated ion channel</li> <li>Transporter</li> </ul>	<p><b>RECEPTORS</b></p> <ul style="list-style-type: none"> <li>Generic Receptor</li> <li>GPCR</li> <li>Receptors with kinase activity</li> </ul>
<p><b>G PROTEIN ADAPTOR/REGULATORS</b></p> <ul style="list-style-type: none"> <li>G beta/gamma</li> <li>Regulators (GDI, GAP, GEF, etc.)</li> </ul>	
<p><b>GROUPS OF OBJECTS</b></p> <ul style="list-style-type: none"> <li><b>A complex or a group</b> Proteins physically connected into a complex or related as a family</li> <li><b>Logical association</b> Proteins linked by logical relations or physical interactions</li> <li><b>Custom association</b> Group of collapsed objects chosen by user</li> </ul>	

## Interactions between objects



## Objects on maps



For more information, visit [clarivate.com/metacore](http://clarivate.com/metacore)

**Anexo D - Tabela das proteínas identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MCF-7**

<b>Classe funcional</b>	<b>Símbolo</b>	<b>Nome da proteína</b>
Componentes do citoesqueleto	ABLIM1	Actin-binding LIM protein 1
	ANK3	Ankyrin-3
	CDC42EP1	Cdc42 effector protein 1
	CEP290	Centrosomal protein of 290 kDa
	CFAP53	Cilia- and flagella-associated protein 53
	CYFIP1	Cytoplasmic FMR1-interacting protein 1
	DCHS2	Protocadherin-23
	DNAH10	Dynein heavy chain 10, axonemal
	DNAH11	Dynein heavy chain 11, axonemal
	EHBP1	EH domain-binding protein 1
	EPB41L3	Band 4.1-like protein 3
	EZR	Ezrin
	GFAP	Glial fibrillary acidic protein
	JAKMIP1	Janus kinase and microtubule-interacting protein 1
	KRT4	Keratin, type II cytoskeletal 4
	KRT7	Keratin, type II cytoskeletal 7
	KRT80	Keratin, type II cytoskeletal 80
	MSN	Moesin
	MYO1E	Unconventional myosin-Ie
	NAV3	Neuron navigator 3
	PCDH15	Protocadherin-15
	PFN1	Profilin-1
	RDX	Radixin
	SHTN1	Shootin-1
	SPAG17	Sperm-associated antigen 17
	TNS1	Tensin-1
TOGARAM1	TOG array regulator of axonemal microtubules protein 1	
TPM2	Tropomyosin beta chain	
TUBA4B	Putative tubulin-like protein alpha-4B	
TUBB4A	Tubulin beta-4A chain	
TUBB4B	Tubulin beta-4B chain	
VILL	Villin-like protein	
Ciclo celular e divisão	CDK5RAP2	CDK5 regulatory subunit-associated protein 2
	CENPF	Centromere protein F
	CIT	Citron Rho-interacting kinase
	NEK2	Serine/threonine-protein kinase Nek2
	PCNT	Pericentrin
	SFN	Stratifin
	TEP1	Telomerase protein component 1
Transporte	ABCC10	ATP-binding cassette sub-family C member 10
	APOB	Apolipoprotein B-100
	CALHM3	Calcium homeostasis modulator protein 3
	FAM160A2	FTS and Hook-interacting protein
	GRIPAP1	GRIP1-associated protein 1
	KPNB1	Importin subunit beta-1
	RAB39A	Ras-related protein Rab-39A
SLC25A5	ADP/ATP translocase 2	

	SLC4A10	Sodium-driven chloride bicarbonate exchanger
	TF	Serotransferrin
	TFG	Protein TFG
	VDAC1	Voltage-dependent anion-selective channel protein 1
	ZFYVE27	Protrudin
Metabolismo	A2M	Alpha-2-macroglobulin
	ADAM9	Disintegrin and metalloproteinase domain-containing protein 9
	ADK	Adenosine kinase
	AK9	Adenylate kinase 9
	ARMC5	Armadillo repeat-containing protein 5
	B4GALT1	Beta-1,4-galactosyltransferase 1
	BMP1	Bone morphogenetic protein 1
	CHIT1	Chitotriosidase-1
	COG7	Conserved oligomeric Golgi complex subunit 7
	CP	Ceruloplasmin
	DUS2	Dihydrouridine synthase 2
	EIF2AK1	Eukaryotic translation initiation factor 2-alpha kinase 1
	FAR2	Fatty acyl-CoA reductase 2
	FBP1	Fructose-1,6-bisphosphatase 1
	G6PD	Glucose-6-phosphate 1-dehydrogenase
	GANAB	Neutral alpha-glucosidase AB
	GUCY1B1	Guanylate cyclase soluble subunit beta-1
	GZMH	Granzyme H
	IRAK1	Interleukin-1 receptor-associated kinase 1
	LARGE1	LARGE xylosyl- and glucuronyltransferase 1
	LDHA	L-lactate dehydrogenase A chain
	MAT2A	S-adenosylmethionine synthase isoform type-2
	MINK1	Misshapen-like kinase 1
	MOGS	Mannosyl-oligosaccharide glucosidase
	NME2P1	Putative nucleoside diphosphate kinase
	P4HB	Protein disulfide-isomerase
	PAM	Peptidyl-glycine alpha-amidating monooxygenase
	PEAK1	Inactive tyrosine-protein kinase PEAK1
	PGA3	Pepsin A-3
	PGA4	Pepsin A-4
	PGA5	Pepsin A-5
	PHKB	Phosphorylase b kinase regulatory subunit beta
	PPIA	Peptidyl-prolyl cis-trans isomerase A
	PPP1CA	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit
	PPP1CB	Serine/threonine-protein phosphatase PP1-beta catalytic subunit
	PPP1CC	Serine/threonine-protein phosphatase PP1-gamma catalytic subunit
RANBP2	E3 SUMO-protein ligase RanBP2	
RNF220	E3 ubiquitin-protein ligase RNF220	
STK25	Serine/threonine-protein kinase 25	
TKFC	Triokinase/FMN cyclase	
UBR4	E3 ubiquitin-protein ligase UBR4	
USP47	Ubiquitin carboxyl-terminal hydrolase 47	
Sinalização celular	AGER	Advanced glycosylation end product-specific receptor
	BANK1	B-cell scaffold protein with ankyrin repeats
	CARD16	Caspase recruitment domain-containing protein 16
	CD2	T-cell surface antigen CD2
	CELSR1	Cadherin EGF LAG seven-pass G-type receptor 1

	DOCK7	Dedicator of cytokinesis protein 7
	EPHA6	Ephrin type-A receptor 6
	FGF18	Fibroblast growth factor 18
	FLOT1	Flotillin-1
	IL31RA	Interleukin-31 receptor subunit alpha
	KIF27	Kinesin-like protein KIF27
	LRBA	Lipopolysaccharide-responsive and beige-like anchor protein
	MAP4K4	Mitogen-activated protein kinase kinase kinase 4
	NCDN	Neurochondrin
	NCSTN	Nicastrin
	NLRC3	NLR family CARD domain-containing protein 3
	OR4A47	Olfactory receptor 4A47
	RASGEF1B	Ras-GEF domain-containing family member 1B
	RWDD3	RWD domain-containing protein 3
	SIPA1L1	Signal-induced proliferation-associated 1-like protein 1
	TFRC	Transferrin receptor protein 1
	TNIK	TRAF2 and NCK-interacting protein kinase
	TREM1	Triggering receptor expressed on myeloid cells 1
	VSIR	V-type immunoglobulin domain-containing suppressor of T-cell activation
	YWHAB	14-3-3 protein beta/alpha
	YWHAE	14-3-3 protein epsilon
	YWHAG	14-3-3 protein gamma
	YWHAH	14-3-3 protein eta
	YWHAQ	14-3-3 protein theta
	YWHAZ	14-3-3 protein zeta/delta
Proteínas asociadas a DNA e RNA	ANKRD11	Ankyrin repeat domain-containing protein 11
	APLF	Aprataxin and PNK-like factor
	BANF1	Barrier-to-autointegration factor 1
	BAZ2B	Bromodomain adjacent to zinc finger domain protein 2B
	BCOR	BCL-6 corepressor
	BRCA2	Breast cancer type 2 susceptibility protein
	CBX8	Chromobox protein homolog 8
	DHX9	ATP-dependent RNA helicase A
	EIF3A	Eukaryotic translation initiation factor 3 subunit A
	GEMIN5	Gem-associated protein 5
	GTF3C4	General transcription factor 3C polypeptide 4
	HFM1	Probable ATP-dependent DNA helicase HFM1
	HNRNPK	Heterogeneous nuclear ribonucleoprotein K
	JADE 1 (PHF17)	Protein Jade-1
	JADE3 (PHF16)	Protein Jade-3
	KAT6A	Histone acetyltransferase KAT6A
	KAT6B	Histone acetyltransferase KAT6B
	KDM5A	Lysine-specific demethylase 5A
	KMT2A	Histone-lysine N-methyltransferase 2A
	LIG3	DNA ligase 3
	LTV1	Protein LTV1 homolog
	MLH1	DNA mismatch repair protein Mlh1
	MTHFSD	Methenyltetrahydrofolate synthase domain-containing protein
	NCL	Nucleolin
	NOL11	Nucleolar protein 11

	NOP56	Nucleolar protein 56
	NPM1	Nucleophosmin
	NR1H3	Oxysterols receptor LXR-alpha
	PARPBP	PCNA-interacting partner
	RBM25	RNA-binding protein 25
	RBM28	RNA-binding protein 28
	SAMD4B	Protein Smaug homolog 2
	SMARCAD1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1
	SREK1	Splicing regulatory glutamine/lysine-rich protein 1
	TEAD2	Transcriptional enhancer factor TEF-4
	THAP9	DNA transposase THAP9
	THOC2	THO complex subunit 2
	TICRR	Treslin
	TLE2	Transducin-like enhancer protein 2
	TSN	Translin
	TWNK	Twinkle protein, mitochondrial
	YTHDC2	3'-5' RNA helicase YTHDC2
	ZC3H13	Zinc finger CCCH domain-containing protein 13
	ZMYND11	Zinc finger MYND domain-containing protein 11
	ZNF432	Zinc finger protein 432
Outros	AHI1	Abelson helper integration site 1 protein homolog
	ARMH3	Armadillo-like helical domain-containing protein 3
	C3orf49	Chromosome 3 Open Reading Frame 49
	CCDC105	Coiled-coil domain-containing protein 105
	CCDC112	Coiled-coil domain-containing protein 112
	CCDC150	Coiled-coil domain-containing protein 150
	CCDC18	Coiled-coil domain-containing protein 18
	CCDC30	Coiled-coil domain-containing protein 30
	CCT2	T-complex protein 1 subunit beta
	CCT4	T-complex protein 1 subunit delta
	COL4A5	Collagen alpha-5(IV) chain
	CRP	C-reactive protein
	F7	Coagulation factor VII
	FAM186B	Protein FAM186B
	FAM208B	Protein FAM208B
	FHAD1	Forkhead-associated domain-containing protein 1
	HCG22	Protein PBMUCL2
	HLAC	HLA class I histocompatibility antigen, Cw-3 alpha chain
	HSPA9	Stress-70 protein, mitochondrial
	IGHA1	Immunoglobulin heavy constant alpha 1
	IGHA2	Immunoglobulin heavy constant alpha 2
	KIAA1109	Transmembrane protein KIAA1109
	LINC00271	Putative uncharacterized protein encoded by LINC00271
	LSMEM2	Leucine-rich single-pass membrane protein 2
	MIDN	Midnolin
	MMACHC	Methylmalonic aciduria and homocystinuria type C protein
	PACS1	Phosphofurin acidic cluster sorting protein 1
	PMS2P2	Putative postmeiotic segregation increased 2-like protein 2
	RPL6	60S ribosomal protein L6
	RPLP1	60S acidic ribosomal protein P1
	RPLP2	60S acidic ribosomal protein P2
	RTKN2	Rhotekin-2

	SH3D21	SH3 domain-containing protein 21
	THEM6	Protein THEM6
	TNXB	Tenascin-X

**Anexo E - Tabela das proteínas identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MDA-MB-231**

<b>Classe funcional</b>	<b>Símbolo</b>	<b>Nome da proteína</b>
Componentes do citoesqueleto	CAP1	Adenylyl cyclase-associated protein 1
	CCDC39	Coiled-coil domain-containing protein 39
	CLIP1	CAP-Gly domain-containing linker protein 1
	COBLL1	Cordon-bleu protein-like 1
	DNAH5	Dynein heavy chain 5, axonemal
	ELMO2	Engulfment and cell motility protein 2
	FAM110A	Protein FAM110A
	FN1	Fibronectin
	FRYL	Protein furry homolog-like
	KLHL1	Kelch-like protein 1
	KRT1	Keratin, type II cytoskeletal 1
	KRT10	Keratin, type II cytoskeletal 10
	KRT13	Keratin, type II cytoskeletal 13
	KRT2	Keratin, type II cytoskeletal 2 epidermal
	KRT24	Keratin, type I cytoskeletal 24
	KRT6B	Keratin, type II cytoskeletal 6B
	KRT77	Keratin, type II cytoskeletal 1b
	MAP7D1	MAP7 domain-containing protein 1
	MYH1	Myosin-1
	MYH14	Myosin-14
	MYH4	Myosin-4
	MYH6	Myosin-6
	MYH7	Myosin-7
	MYH8	Myosin-8
	MYO16	Unconventional myosin-XVI
	MYO18B	Unconventional myosin-XVIIIb
	NEGR1	Neuronal growth regulator 1
	ODF2	Outer dense fiber protein 2
	PATJ	InaD-like protein
	PCDHA3	Protocadherin alpha-3
	PHACTR2	Phosphatase and actin regulator 2
	SNTG1	Gamma-1-syntrophin
	SPEF2	Sperm flagellar protein 2
SPTB	Spectrin beta chain, erythrocytic	
TEKT4	Tektin-4	
TUBB8	Tubulin beta-8 chain	
VCL	Vinculin	
Ciclo celular e divisão	ATR	Serine/threonine-protein kinase ATR
	HAUS4	HAUS augmin-like complex subunit 4
Transporte	CLTC	Clathrin heavy chain 1
	CTTNBP2NL	CTTNBP2 N-terminal-like protein
	EEA1	Early endosome antigen 1
	GOLGA4	Golgin subfamily A member 4
	PLEKHA8P1	Putative protein PLEKHA9
	SNX17	Sorting nexin-17



	TFG	Protein TFG
	TIMM9	Mitochondrial import inner membrane translocase subunit Tim9
Metabolismo	ATP5F1A	ATP synthase subunit alpha, mitochondrial
	CAMK1D	Calcium/calmodulin-dependent protein kinase type 1D
	CBLB	E3 ubiquitin-protein ligase CBL-B
	COP1	E3 ubiquitin-protein ligase COP1
	ECH1	Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial
	GLB1L2	Beta-galactosidase-1-like protein 2
	GSTA1	Glutathione S-transferase A1
	GSTA2	Glutathione S-transferase A2
	GSTA3	Glutathione S-transferase A3
	GSTA4	Glutathione S-transferase A4
	GSTA5	Glutathione S-transferase A5
	IMPDH2	Inosine-5'-monophosphate dehydrogenase 2
	MACROD2	O-acetyl-ADP-ribose deacetylase MACROD2
	NAGK	N-acetyl-D-glucosamine kinase
	NDUFAF7	Protein arginine methyltransferase NDUFAF7, mitochondrial
	NEU4	Sialidase-4
	PGK1	Phosphoglycerate kinase 1
	PLD2	Phospholipase D2
	PRODH2	Hydroxyproline dehydrogenase
	PXDN	Peroxidasin homolog
	SURF1	Surfeit locus protein 1
	UBE2H	Ubiquitin-conjugating enzyme E2 H
USP48	Ubiquitin carboxyl-terminal hydrolase 48	
VCP	Transitional endoplasmic reticulum ATPase	
Sinalização celular	BRAF	Serine/threonine-protein kinase B-raf
	BTBD7	BTB/POZ domain-containing protein 7
	CCDC88A	Girdin
	CTNND2	Catenin delta-2
	DENND2A	DENN domain-containing protein 2A
	FER	Tyrosine-protein kinase Fer
	GMP	GEM-interacting protein
	GUCY1B1	Guanylate cyclase soluble subunit beta-1
	KIF27	Kinesin-like protein KIF27
	LRP1B	Low-density lipoprotein receptor-related protein 1B
	LYN	Tyrosine-protein kinase Lyn
	MAP4K3	Mitogen-activated protein kinase kinase kinase kinase 3
	OTOF	Otoferlin
	PRPSAP2	Phosphoribosyl pyrophosphate synthase-associated protein 2
	PSD3	PH and SEC7 domain-containing protein 3
	STK25	Serine/threonine-protein kinase 25
	TENM2	Teneurin-2
UACA	Uveal autoantigen with coiled-coil domains and ankyrin repeats	
Proteínas associadas a DNA e RNA	ANKRD11	Ankyrin repeat domain-containing protein 11
	ANKRD12	Ankyrin repeat domain-containing protein 12
	BOD1L1	Biorientation of chromosomes in cell division protein 1-like 1
	BPTF	Nucleosome-remodeling factor subunit BPTF
	CCDC171	Coiled-coil domain-containing protein 171
	EIF5A	Eukaryotic translation initiation factor 5A-1

	FAM208A	Protein TASOR
	GRHL1	Grainyhead-like protein 1 homolog
	GTF3C4	General transcription factor 3C polypeptide 4
	HNRNPK	Heterogeneous nuclear ribonucleoprotein K
	ISL2	Insulin gene enhancer protein ISL-2
	JADE 1 (PHF17)	Protein Jade-1
	NKRF	NF-kappa-B-repressing factor
	NOBOX	Homeobox protein NOBOX
	NOL11	Nucleolar protein 11
	PGBD5	PiggyBac transposable element-derived protein 5
	PTBP1	Polypyrimidine tract-binding protein 1
	RSBN1L	Round spermatid basic protein 1-like protein
	SLFN11	Schlafen family member 11
	SMARCD3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3
	SNRNP35	U11/U12 small nuclear ribonucleoprotein 35 kDa protein
	SPOCD1	SPOC domain-containing protein 1
	STAU2	Double-stranded RNA-binding protein Staufen homolog 2
	THOC2	THO complex subunit 2
	YTHDC2	3'-5' RNA helicase YTHDC2
	ZC3H13	Zinc finger CCCH domain-containing protein 13
	ZNF107	Zinc finger protein 107
	ZNF117	Zinc finger protein 117
	ZNF273	Zinc finger protein 273
	ZNF506	Zinc finger protein 506
	ZNF626	Zinc finger protein 626
	ZNF66	Putative zinc finger protein 66
	ZNF681	Zinc finger protein 681
	ZNF737	Zinc finger protein 737
	ZNF85	Zinc finger protein 85
	ZNF92	Zinc finger protein 92
	ZNHIT2	Zinc finger HIT domain-containing protein 2
Outros	A1BG	Alpha-1B-glycoprotein
	ASB17	Ankyrin repeat and SOCS box protein 17
	C11orf40	Putative uncharacterized protein C11orf40
	CCDC18	Coiled-coil domain-containing protein 18
	CCER2	Coiled-coil domain-containing glutamate-rich protein 2
	CCSER1	Serine-rich coiled-coil domain-containing protein 1
	EFHB	EF-hand domain-containing family member B
	EFHC2	EF-hand domain-containing family member C2
	F5	Coagulation factor V
	FAM216B	Protein FAM216B
	FHAD1	Forkhead-associated domain-containing protein 1
	GOLGA8S	Golgin subfamily A member 8S
	HSPA4	Heat shock 70 kDa protein 4
	HYOU1	Hypoxia up-regulated protein 1
	IGFBPL1	Insulin-like growth factor-binding protein-like 1
	IGSF9	Protein turtle homolog A
	LSMEM2	Leucine-rich single-pass membrane protein 2
	MRPL9	39S ribosomal protein L9, mitochondrial
	PID1	PTB-containing, cubilin and LRP1-interacting protein

	PRRC2C	Protein PRRC2C
	PSMD2	26S proteasome non-ATPase regulatory subunit 2
	PSME1	Proteasome activator complex subunit 1
	RANBP2	E3 SUMO-protein ligase RanBP2

**Anexo F - Tabela das proteínas associadas a DNA e RNA identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MCF-7**

<b>Subclasse funcional</b>	<b>Símbolo</b>	<b>Nome da proteína</b>	
Proteínas basais e estruturais do núcleo	BANF1	Barrier-to-autointegration factor 1	
	HFM1	Probable ATP-dependent DNA helicase HFM1	
	LIG3	DNA ligase 3	
	NPM1	Nucleophosmin	
	THAP9	DNA transposase THAP9	
	TSN	Translin	
Regulação da replicação	TICRR	Treslin	
	TWINK	Twinkle protein, mitochondrial	
Regulação da transcrição	DHX9	ATP-dependent RNA helicase A	
	GTF3C4	General transcription factor 3C polypeptide 4	
	NCL	Nucleolin	
	NR1H3	Oxysterols receptor LXR-alpha	
	SAMD4B	Protein Smaug homolog 2	
	TEAD2	Transcriptional enhancer factor TEF-4	
	TLE2	Transducin-like enhancer protein 2	
	ZNF432	Zinc finger protein 432	
	Regulação da tradução	EIF3A	Eukaryotic translation initiation factor 3 subunit A
		YTHDC2	3'-5' RNA helicase YTHDC2
Splicing e processamento de RNA	GEMIN5	Gem-associated protein 5	
	HNRNPK	Heterogeneous nuclear ribonucleoprotein K	
	LTV1	Protein LTV1 homolog	
	NOL11	Nucleolar protein 11	
	NOP56	Nucleolar protein 56	
	RBM25	RNA-binding protein 25	
	RBM28	RNA-binding protein 28	
	SREK1	Splicing regulatory glutamine/lysine-rich protein 1	
	THOC2	THO complex subunit 2	
	ZC3H13	Zinc finger CCCH domain-containing protein 13	
	Resposta a dano de DNA	APLF	Aprataxin and PNK-like factor
BRCA2		Breast cancer type 2 susceptibility protein	
MLH1		DNA mismatch repair protein Mlh1	
PARPBP		PCNA-interacting partner	
Modificações de histona e remodelamento de cromatina	ANKRD11	Ankyrin repeat domain-containing protein 11	
	BAZ2B	Bromodomain adjacent to zinc finger domain protein 2B	
	BCOR	BCL-6 corepressor	
	CBX8	Chromobox protein homolog 8	
	JADE 1 (PHF17)	Protein Jade-1	
	JADE3 (PHF16)	Protein Jade-3	
	KAT6A	Histone acetyltransferase KAT6A	
	KAT6B	Histone acetyltransferase KAT6B	
	KDM5A (RBBP2)	Lysine-specific demethylase 5A	

	KMT2A (MLL1)	Histone-lysine N-methyltransferase 2A
	SMARCAD1	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1
	ZMYND11	Zinc finger MYND domain-containing protein 11
Funções pouco conhecidas	MTHFSD	Methenyltetrahydrofolate synthase domain-containing protein

**Anexo G - Tabela das proteínas associadas a DNA e RNA identificadas como parceiras de ASXL2 por abordagem proteômica na linhagem MDA-MB-231**

<b>Subclasse funcional</b>	<b>Símbolo</b>	<b>Nome da proteína</b>
Proteínas basais e estruturais do núcleo	PGBD5	PiggyBac transposable element-derived protein 5
Regulação da replicação	BOD1L1	Biorientation of chromosomes in cell division protein 1-like 1
Regulação da transcrição	GRHL1	Grainyhead-like protein 1 homolog
	ISL2	Insulin gene enhancer protein ISL-2
	NKRF	NF-kappa-B-repressing factor
	NOBOX	Homeobox protein NOBOX
	SPOCD1	SPOC domain-containing protein 1
	ZNF107	Zinc finger protein 107
	ZNF117	Zinc finger protein 117
	ZNF273	Zinc finger protein 273
	ZNF506	Zinc finger protein 506
	ZNF626	Zinc finger protein 626
	ZNF66	Putative zinc finger protein 66
	ZNF681	Zinc finger protein 681
	ZNF737	Zinc finger protein 737
	ZNF85	Zinc finger protein 85
ZNF92	Zinc finger protein 92	
Regulação da tradução	EIF5A	Eukaryotic translation initiation factor 5A-1
	YTHDC2	3'-5' RNA helicase YTHDC2
Splicing e processamento de RNA	HNRNPK	Heterogeneous nuclear ribonucleoprotein K
	NOL11	Nucleolar protein 11
	PTBP1	Polypyrimidine tract-binding protein 1
	SNRNP35	U11/U12 small nuclear ribonucleoprotein 35 kDa protein
	THOC2	THO complex subunit 2
	ZC3H13	Zinc finger CCCH domain-containing protein 13
Resposta a dano de DNA	SLFN11	Schlafen family member 11
Modificações de histona e remodelamento de cromatina	ANKRD11	Ankyrin repeat domain-containing protein 11
	ANKRD12	Ankyrin repeat domain-containing protein 12
	BPTF	Nucleosome-remodeling factor subunit BPTF
	FAM208A	Protein TASOR
	GTF3C4	General transcription factor 3C polypeptide 4
	JADE 1 (PHF17)	Protein Jade-1
	RSBN1L	Round spermatid basic protein 1-like protein
	SMARCD3	SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily D member 3
Funções pouco conhecidas	CCDC171	Coiled-coil domain-containing protein 171
	STAU2	Double-stranded RNA-binding protein Staufen homolog 2
	ZNHIT2	Zinc finger HIT domain-containing protein 2