

**UNIVERSIDADE DE SÃO PAULO**

**INSTITUTO DE QUÍMICA**

**Programa de Pós-Graduação em Ciências Biológicas (Bioquímica)**

**HENRIQUE CÉSAR DE JESUS FERREIRA**

**Caracterização funcional do gene codificador de  
proteína *CD14* e do gene não codificador de proteína  
*LINC01133* em linhagens celulares humanas de  
câncer de mama triplo negativo**

**Versão corrigida da Tese defendida**

São Paulo

Data do Depósito na SPG:

**08/09/2021**

HENRIQUE CÉSAR DE JESUS FERREIRA

**Caracterização funcional do gene codificador de  
proteína *CD14* e do gene não codificador de proteína  
*LINC01133* em linhagens celulares humanas de  
câncer de mama triplo negativo**

*Tese apresentada ao Instituto de Química da  
Universidade de São Paulo para obtenção do  
Título de Doutor em Ciências Biológicas  
(Bioquímica).*

*Orientadora: Profa. Dra. Mari Cleide Sogayar*

São Paulo

2021

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha Catalográfica elaborada eletronicamente pelo autor, utilizando o programa desenvolvido pela Seção Técnica de Informática do ICMC/USP e adaptado para a Divisão de Biblioteca e Documentação do Conjunto das Químicas da USP

Bibliotecária responsável pela orientação de catalogação da publicação:  
Marlene Aparecida Vieira - CRB - 8/5562

J58c	Jesus-Ferreira, Henrique César de Caracterização funcional do gene codificador de proteína CD14 e do gene não codificador de proteína LINC01133 em linhagens celulares humanas de câncer de mama triplo negativo / Henrique César de Jesus-Ferreira. - São Paulo, 2021. 201 p.
	Tese (doutorado) - Instituto de Química da Universidade de São Paulo. Departamento de Bioquímica. Orientador: Sogayar, Mari Cleide Coorientador: Carreira, Ana Claudia Oliveira
	1. Câncer de mama triplo negativo (TNBC). 2. Biomarcadores. 3. CD14. 4. lncRNAs. 5. LINC01133. I. T. II. Sogayar, Mari Cleide, orientador. III. Carreira, Ana Claudia Oliveira, coorientador.



Dedicado aos meus amados pais, pela educação e dedicação incondicional.

À minha amada irmã, pela força e pelo exemplo de ser humano.

## **AGRADECIMENTOS**

À Deus, por me dar a força necessária para enfrentar os obstáculos que passei ao longo desses anos, que foram muitos, e por me ajudar a atingir mais um objetivo. Nas minhas inúmeras quedas você me ajudou a levantar e seguir em frente.

À minha amiga e orientadora, Profa. Dra. Mari Cleide Sogayar, que aceitou meus pedidos e me acolheu como aluno de doutorado, por toda confiança, por todos os ensinamentos, pela paciência ao longo desses anos, pela amizade, pelo exemplo de liderança, por me tratar como um filho em um dos momentos mais críticos da minha vida e por ser o pilar do nosso grupo de pesquisa NUCEL. Sua experiência foi fundamental para o meu crescimento pessoal e para minha carreira científica.

À minha amiga e co-orientadora, Dra. Ana Claudia Oliveira Carreira, pelos inúmeros ensinamentos científicos, pelo auxílio com os experimentos, por me tranquilizar em momentos de angústia, por sempre acreditar no meu potencial, por nunca desistir de mim e principalmente pela amizade. Você foi um dos principais pilares para que este objetivo fosse alcançado.

Às minhas grandes amigas, Dra. Miriam Lemos, Dra. Marluce Mantovani e Zizi de Mendonça. Sem vocês eu não teria chegado até o fim. Agradeço pelo apoio técnico, pela amizade, pelos conselhos, pelos inúmeros momentos de diversão e pela ajuda em um dos momentos mais críticos da minha vida. Vocês estarão para sempre em meu coração.

À Dra. Raquel Arminda Carvalho Machado, pela disponibilidade, pelos conselhos científicos e pelo auxílio em alguns experimentos.

Ao Dr. Marcos Demasi, pelo exemplo de profissional e pelas discussões científicas.

Ao Dr. Fernando Lojudice, por todo conhecimento técnico e científico compartilhado, pelos conselhos profissionais e pelo auxílio em alguns experimentos.

Aos amigos e funcionários do grupo NUCEL, Alan, Érika, Renatinha e Michele, pela amizade, pelo trabalho essencial para o funcionamento do grupo e pelos inúmeros momentos de descontração durante à pausa dos experimentos. Vocês fazem parte desta conquista.

À aluna de Iniciação Científica que orientei, Isabeli Yumi, pelo auxílio nos experimentos e por contribuir para o meu crescimento científico.

A todos os outros amigos e colegas da família NUCEL que estiveram presentes ao longo desta jornada: Mariele Moraes, Bruna Aguiar, Mariana Araújo, Andrew Douglas, Camila

Azzi, Camila Lopes, Marina Trombetta, Gabriela Manini, Renato Astorino, Ana Paula Barreto, Luiz Pomatay, Patrícia Kossugue, Aline Loba, Gabriel Levin, Alice Sampaio, Lauren Camargo, Mariel Acorte, Emília Zenteno, Darshna Pryia, Raja, Isaura Beatriz, Vitor Colantoni, Thamara Reis, Aduino Spindola, Thais Ribas e Roberta Facioli.

Aos meus amigos e eternos “*roomies*”, Vinícius de Moraes, Felipe Gerônimo, Sandro Filho e João Paulo, pela amizade, pelas discussões científicas, pelos inúmeros momentos de diversão e descontração, pelas reuniões nas sextas-feiras à noite e por todo suporte em um dos momentos mais críticos da minha vida. Nem só de pesquisa vive um Doutorando.

Aos meus amigos de São Paulo - SP, Marcos Devaner, Thiago Augusto, Vere Rohr, Viviane Sardilli, Lesly Gomes, Alex Ribeiro, Nero Wanderley, Fernando Borges, Vinícius Marques, André Damasceno, Deusa Medeiros e Tajiuro Himeno.

Aos meus amigos de Natal - RN, que sempre me acompanharam e torceram por mim, independente da distância: Isabelle Arruda, Pollyana Gomes, William Danilo, Felipe Ferreira, Aristécio Júnior, Luciene Araújo, Taty Silva, Dhyego Melo, Maycon Ranniere, Mario Canuto, Ritinha Araújo e Joaquim Aristelo.

À Marcela Giancoli, pelo auxílio na confecção das imagens de conclusão dos Capítulos da Tese.

À Coordenação do Curso de Pós-Graduação em Bioquímica.

Ao CNPq e a CAPES pelo apoio financeiro para o desenvolvimento deste trabalho.

À minha afilhada e sobrinha, Cecília. Você foi um dos maiores presentes dados por Deus para a nossa família. Cuidarei de você eternamente.

À minha amiga e irmã, Amanda, por me ouvir e aconselhar nos momentos que pensei em desistir, por acreditar sempre em mim, pelas inúmeras palavras de força, por ser um exemplo de ser humano e por toda cumplicidade ao longo desses anos. Você é e sempre será um exemplo para mim. Te amo.

Finalmente, aos meus amados pais, Oswaldo e Inácia, pelo exemplo de luta, por sempre terem acreditado em mim, por estarem ao meu lado nos momentos que mais precisei, por compreenderem a minha ausência ao longo destes anos, por me acolherem nos meus momentos de tristeza e cansaço excessivo e pelo suporte financeiro. Vocês foram os principais alicerces para que eu pudesse suportar todos os desafios e medos que apareceram neste longo caminho. Palavras não são suficientes para expressar o tamanho do meu amor por vocês. Obrigado por tudo.

“Só se pode alcançar um grande êxito quando nos mantemos fieis a nós mesmos”

**Friedrich Nietzsche**

## RESUMO

JESUS-FERREIRA, H. C. **Caracterização funcional do gene codificador de proteína *CD14* e do gene não codificador de proteína *LINC01133* em linhagens celulares humanas de câncer de mama triplo negativo.** 2021. 201p. Tese de Doutorado. Programa de Pós-Graduação em Ciências Biológicas (Bioquímica). Instituto de Química, Universidade de São Paulo, São Paulo.

Dentre os subtipos de câncer de mama, o triplo negativo (TNBC) é o que apresenta as maiores taxas de mortalidade, sendo, portanto, considerado um enorme desafio para a clínica. O uso de moléculas como marcadores tumorais vem auxiliando o clínico no diagnóstico, no prognóstico e, até mesmo, no tratamento do TNBC, sendo essenciais na redução de suas altas taxa de mortalidade. No entanto, um pequeno grupo de marcadores tumorais são validados na prática clínica, estimulando à busca por novos alvos, e sua caracterização funcional, como forma de se entender a Biologia desta doença. Assim, o objetivo deste trabalho é caracterizar funcionalmente o gene codificador de proteína *CD14* e o gene não codificador de proteína *LINC01133* em linhagens celulares humanas de TNBC, no intuito de descobrir o papel destas moléculas na progressão tumoral.

Na primeira parte deste trabalho, analisou-se a expressão do *CD14* frente à um painel de linhagens celulares que representam os diferentes subtipos dos tumores mamários. O *CD14* exibiu elevados níveis de expressão nas linhagens não-tumorigênicas MCF10A e MCF12A e baixos níveis na linhagem triplo negativa Hs578T. A partir destes resultados, o *CD14* foi superexpresso na linhagem Hs578T. Ensaios de caracterização funcional mostraram que a superexpressão do *CD14* reduziu a capacidade migratória e invasiva das células, efeito que foi hipoteticamente relacionado ao aumento da expressão da E-caderina. No entanto, observou-se aumento no potencial tumorigênico, levando-nos a sugerir seu envolvimento num possível mecanismo utilizado pelas células para compensar a significativa redução do potencial migratório e invasivo. Os resultados obtidos indicam que o nível basal de expressão do *CD14* observado na linhagem Hs578T é importante, podendo contribuir para a desenvolvimento primário do tumor, atuando como um oncogene.

Na segunda parte deste trabalho, analisou-se a expressão de 10 RNAs longos não codificadores (lncRNAs), frente ao mesmo painel de linhagens descrito

anteriormente. Dentre estes, o lncRNA LINC01133 exibiu baixos níveis de expressão nas linhagens não-tumorigênicas MCF10A e MCF12A e elevados níveis na linhagem triplo negativa Hs578T, sendo, então, escolhido como alvo de estudo. A partir destes resultados, decidimos superexpressar, de forma indutível, o *LINC01133* na linhagem MCF10A e nocautear este gene, via sistema CRISPR/Cas9, na linhagem Hs578T. Ensaios de caracterização funcional mostraram que a superexpressão do *LINC01133* na linhagem MCF10A reduziu a proliferação celular e inibiu o crescimento de colônias dependente de ancoragem, mas, em contrapartida, aumentou o crescimento de colônias independente de ancoragem e a capacidade migratória e invasiva destas células. No entanto, sugerimos que isto não seja suficiente para tornar estas células tumorigênicas e metastáticas. Por outro lado, o nocauteamento do LINC01133 na linhagem triplo negativa Hs578T aumentou de forma considerável todos os parâmetros de malignidade analisados. Baseado nos dados obtidos, sugerimos que o elevado nível de expressão do *LINC01133* na linhagem Hs578T é importante na regulação negativa de processos relacionados com a progressão tumoral, atuando com um supressor tumoral. Os dados obtidos em nosso estudo contribuem para o enriquecimento de informações relacionadas à Biologia do TNBC, auxiliando, desta forma, no desenvolvimento de potenciais protocolos clínicos e terapêuticos utilizando estes biomarcadores.

Palavras-chave: Câncer de mama triplo negativo (TNBC), Biomarcadores, *CD14*, LncRNAs, *LINC01133*.

## ABSTRACT

JESUS-FERREIRA, H. C. **Functional characterization of the *CD14* protein-coding gene and the non-protein-coding *LINC01133* gene in triple negative human breast cancer cell lines.** 2021. 201p. PhD Thesis. Graduate Program in Biological Sciences (Biochemistry). Instituto de Química, Universidade de São Paulo, São Paulo.

Among the breast cancer subtypes, the triple negative (TNBC) displays the highest mortality rates, being, therefore, considered a major challenge for the clinic. The use of molecules as tumor markers has helped clinicians in the diagnosis, prognosis and even in treatment of TNBC, being essential in reducing its high mortality rate. However, a small group of tumor markers is validated in clinical practice, stimulating the search for new targets, and their functional characterization, as a way to understand the biology of this disease. Thus, the aim of this work is to functionally characterize the *CD14* protein-coding gene and the non-protein-coding *LINC01133* gene in human TNBC cell lines, in order to probe into the role of these molecules in tumor progression.

In the first part of this work, the expression of *CD14* was analyzed in a panel of cell lines that represent the different subtypes of breast tumors. High expression levels of *CD14* were observed in the non-tumorigenic MCF10A and MCF12A lineages and low levels in the triple negative Hs578T lineage. Based on these results, *CD14* was overexpressed in the Hs578T lineage. Functional characterization assays showed that *CD14* overexpression reduced the migratory and invasive capacity of cells, an effect that was hypothetically related to increased E-cadherin expression. However, increased tumorigenic potential was observed, leading us to suggest its involvement in a possible mechanism used by cells to compensate for the significant reduction in the migratory and invasive potential. The results obtained indicate that *CD14* expression basal level observed in the Hs578T lineage may be important to contribute to the primary development of tumor, thus acting as an oncogene.

In the second part of this work, the expression of 10 long non-coding RNAs (lncRNAs) was analyzed against the same lineage panel described above. Among these, the *LINC01133* lncRNA exhibited low expression levels in the non-tumorigenic MCF10A and MCF12A lineages and high levels in the triple negative Hs578T lineage, being, then, chosen as a target for this study. Based on these results, we decided to

inducibly overexpress *LINC01133* in the MCF10A lineage and knockout this gene, via the CRISPR/Cas9 system, in the Hs578T lineage. Functional characterization assays showed that overexpression of *LINC01133* in the MCF10A lineage reduced cell proliferation and inhibited anchorage-dependent colony growth, but, on the other hand, increased anchorage-independent colony growth and the migratory and invasive capacity of these cells. However, we suggest that this is not sufficient to render these cells tumorigenic and metastatic. On the other hand, the knockout of *LINC01133* in the triple negative Hs578T lineage considerably increased all the analyzed malignancy parameters. Based on the results obtained, we suggest that the high expression level of *LINC01133* in the Hs578T lineage is important for down-regulation of processes related to tumor progression, acting as a tumor suppressor. The data obtained in our study contribute to the enrichment of information related to TNBC Biology, thus assisting in the development of potential clinical and therapeutic protocols using these biomarkers.

Keywords: Triple-negative breast cancer (TNBC), biomarkers, *CD14*, lncRNAs, *LINC01133*.

## LISTA DE ABREVIATURAS E SIGLAS

<b>µg</b>	Micrograma
<b>µL</b>	Microlitro
<b>µM</b>	Micromolar
<b>µm</b>	Micrometro
<b>ASO</b>	<i>Antisense Oligonucleotide</i>
<b>ATCC</b>	<i>American Type Culture Collection</i>
<b>BCSC</b>	<i>Breast Cancer Stem Cell</i>
<b>BL1</b>	<i>Basal-like 1</i>
<b>BL2</b>	<i>Basal-like 2</i>
<b>BLIS</b>	<i>Basal-like and imune supressed</i>
<b>BRCA1</b>	<i>Breast Cancer gene 1</i>
<b>BRCA2</b>	<i>Breast Cancer gene 2</i>
<b>Cas9</b>	Proteína 9 associada ao CRISPR
<b>cDNA</b>	DNA Complementar
<b>CD14</b>	<i>Cluster of differentiation 14</i>
<b>COL3A1</b>	<i>Colágeno Tipo III Cadeia Alpha 1</i>
<b>CRISPR</b>	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i>
<b>Ct</b>	<i>Threshold cycle</i>
<b>CXCL12</b>	C-X-C Motif Chemokine Ligand 12
<b>DAMPs</b>	Padrões moleculares associados ao dano
<b>DNA</b>	Ácido Desoxirribonucleico
<b>DMEM</b>	<i>Dulbecco's Modified Eagle Medium</i>
<b>DMSO</b>	Dimetilsulfóxido
<b>E-CAD</b>	E-caderina
<b>EGFR</b>	Receptor do fator de crescimento epidérmico
<b>EMT</b>	<i>Epithelial–Mesenchymal Transition</i>
<b>ER</b>	Receptor de estrogênio
<b>FACS</b>	<i>Fluorescence-activated cell sorting</i>
<b>GAPDH</b>	<i>Glyceraldehyde 3-phosphate dehydrogenase</i>
<b>GPI</b>	Glicosilfosfatidilinositol
<b>gRNA</b>	RNA guia
<b>HER2</b>	Receptor do fator de crescimento epidérmico humano do tipo 2

<b>HMBS</b>	<i>Hydroxymethylbilane Synthase</i>
<b>HPRT1</b>	<i>Hypoxanthine Phosphoribosyltransferase 1</i>
<b>IARC</b>	<i>International Agency for Research on Cancer</i>
<b>IDC-NST</b>	<i>Invasive Ductal Breast Cancer of No Special Type</i>
<b>IHC</b>	Imunohistoquímica
<b>IM</b>	Imunomodulador
<b>Ki-67</b>	<i>Marker of proliferation Ki-67</i>
<b>Kpb</b>	<i>Kilo pares de base</i>
<b>KO</b>	<i>Knockout</i>
<b>LAR</b>	<i>Luminal androgen receptor</i>
<b>LINC01133</b>	<i>Long Intergenic Non-Protein Coding RNA 1133</i>
<b>LincRNA</b>	RNA longo não codificador intergênico
<b>LncRNA</b>	RNA longo não codificador
<b>LPS</b>	Lipopolissacarídeo
<b>MAPK</b>	<i>Mitogen-activated protein kinase</i>
<b>mCD14</b>	CD14 ancorado à membrana
<b>MES</b>	<i>Mesenchymal-like</i>
<b>miRNA</b>	microRNA
<b>mL</b>	Mililitro
<b>MOI</b>	<i>Multiplicity of Infection</i>
<b>mRNA</b>	RNA mensageiro
<b>MSL</b>	<i>Mesenchymal stem-like</i>
<b>ncRNA</b>	RNA não codificador
<b>N-CAD</b>	N-caderina
<b>nt</b>	Nucleotídeo
<b>PAMPs</b>	Padrões moleculares associados a patógenos
<b>PARP1</b>	<i>Poly (ADP-ribose) Polymerase 1</i>
<b>pb</b>	Pares de base
<b>PBS</b>	<i>Phosphate-buffered saline</i>
<b>PR</b>	Receptor de progesterona
<b>PRP</b>	Receptor de reconhecimento padrão
<b>qRT-PCR</b>	PCR quantitativo em tempo real
<b>RNA</b>	Ácido Ribonucleico
<b>SATB1</b>	<i>Special AT-rich Sequence Binding-Protein 1</i>
<b>sCD14</b>	CD14 como proteína solúvel

<b>SFB</b>	Soro Fetal Bovino
<b>SFE</b>	Soro Fetal Equino
<b>SPARC</b>	<i>Secreted Protein Acidic And Cysteine Rich</i>
<b>TGF-<math>\beta</math></b>	<i>Transforming Growth Factor beta</i>
<b>TLRs</b>	<i>Toll-like Receptors</i>
<b>TNBC</b>	<i>Triple-negative breast cancer</i>
<b>VEGF-A</b>	<i>Vascular Endothelial Growth Factor A</i>
<b>XIAP</b>	<i>X-Linked Inhibitor Of Apoptosis</i>
<b>WHO</b>	<i>World Health Organization</i>
<b>WT</b>	<i>Wild Type</i>

# SUMÁRIO

<b>1. INTRODUÇÃO GERAL</b>	<b>20</b>
1.1. Câncer	20
1.2. Câncer de mama	23
1.2.1. Classificação em Tipos Histológicos	28
1.2.2. Classificação em Grau Histológico	31
1.2.3. Classificação por Estadiamento	31
1.2.4. Classificação Molecular	32
1.3. Câncer de mama triplo negativo (TNBC)	37
1.4. Biomarcadores no câncer de mama triplo negativo	45
<b>2. CAPÍTULO I: Caracterização funcional (superexpressão) do gene codificador de proteína CD14 em linhagens celulares humanas de câncer de mama triplo negativo</b>	<b>50</b>
<b>2.1. INTRODUÇÃO</b>	<b>51</b>
<b>2.2. OBJETIVOS</b>	<b>57</b>
2.2.1. Objetivo Geral	57
2.2.2. Objetivos Específicos	57
<b>2.3. MATERIAIS E MÉTODOS</b>	<b>58</b>
2.3.1. Linhagens Celulares e Condições de Cultivo	58
2.3.2. Ensaio de PCR quantitativo em tempo real (qRT-PCR)	60
2.3.2.1. Extração do RNA total	60
2.3.2.2. Desenho dos <i>primers</i>	60
2.3.2.3. Transcrição reversa e síntese de cDNA	61
2.3.2.4. Determinação da concentração final de <i>primers</i>	62
2.3.2.5. Determinação da eficiência dos <i>primers</i>	62
2.3.2.6. Confirmação da expressão diferencial	63
2.3.3. Construção do vetor lentiviral pLV-eGFP-CD14 para ensaio de superexpressão	63
2.3.3.1. Desenho dos primers para amplificação da região codificadora do <i>CD14</i>	64
2.3.3.2. Amplificação da região codificadora do gene <i>CD14</i>	64
2.3.3.3. Clonagem da região codificadora do gene <i>CD14</i>	65
2.3.3.4. Transformação por eletroporação de bactérias <i>Escherichia coli</i> eletrocompetentes com DNA plasmideal	65
2.3.3.5. Preparação de DNA plasmideal em pequena escala	66
2.3.3.6. Análise da sequência de nucleotídeos	67
2.3.3.7. Sub-clonagem da construção no vetor lentiviral de terceira geração	67
2.3.4. Geração da linhagem Hs578T superexpressando o gene <i>CD14</i>	67

2.3.4.1. Produção de partículas lentivirais em células 293T	67
2.3.4.2. Titulação das partículas lentivirais eGFP positivas (eGFP <sup>+</sup> )	68
2.3.4.3. Transdução viral da linhagem Hs578T com partículas lentivirais	69
2.3.5. Citometria de Fluxo	69
2.3.6. Confirmação da superexpressão do <i>CD14</i> nas linhagens transduzidas	70
2.3.7. Caracterização das linhagens celulares geradas	70
2.3.7.1. Curva de crescimento celular	70
2.3.7.2. Ensaio de tumorigênese <i>in vitro</i>	71
2.3.7.2.1. Ensaio de crescimento de colônias em substrato sólido	71
2.3.7.2.2. Ensaio de crescimento de colônias em suspensão de agarose	72
2.3.7.3. Ensaio de migração celular <i>in vitro</i>	72
2.3.7.3.1. Ensaio de migração celular por <i>scratching</i> ( <i>wound healing</i> )	72
2.3.7.3.2. Ensaio de migração celular por <i>Transwell</i> <sup>TM</sup>	73
2.3.7.4. Ensaio de invasão celular <i>in vitro</i>	74
2.3.7.5. Análise da expressão de genes envolvidos na malignidade tumoral	74
2.3.8. Análises Estatísticas	75
<b>2.4. RESULTADOS</b>	<b>76</b>
2.4.1. Análise da expressão do <i>CD14</i> no modelo de progressão tumoral do carcinoma mamário	76
2.4.2. Construção do vetor lentiviral pLV-EGFP/ <i>CD14</i> para ensaio de superexpressão do <i>CD14</i>	77
2.4.3. Geração da linhagem celular Hs578T superexpressando o gene <i>CD14</i>	78
2.4.4. Confirmação da superexpressão do <i>CD14</i> na linhagem Hs578T transduzida	82
2.4.5. Caracterização das linhagens celulares geradas	82
2.4.5.1. Curva de crescimento celular	82
2.4.5.2. Ensaio de tumorigênese <i>in vitro</i>	83
2.4.5.2.1. Ensaio de crescimento de colônias em substrato sólido	83
2.4.5.2.2. Ensaio de crescimento de colônias em suspensão de agarose	84
2.4.5.3. Ensaio de migração celular <i>in vitro</i>	85
2.4.5.3.1. Ensaio de migração celular por <i>scratching</i> ( <i>wound healing</i> )	86
2.4.5.3.2. Ensaio de migração celular por <i>Transwell</i> <sup>TM</sup>	86
2.4.5.4. Ensaio de invasão celular <i>in vitro</i>	87
2.4.5.5. Análise da expressão de genes envolvidos na malignidade tumoral	88
<b>2.5. DISCUSSÃO</b>	<b>90</b>
<b>2.6. CONCLUSÃO</b>	<b>99</b>
<b>2.7. PRÓXIMOS PASSOS</b>	<b>100</b>

<b>3. CAPÍTULO II: Identificação e caracterização funcional (superexpressão indutível e nocauteamento) do gene não codificador de proteína <i>LINC01133</i> em linhagens celulares humanas de câncer de mama triplo negativo</b>	<b>101</b>
<b>3.1. INTRODUÇÃO</b>	<b>102</b>
<b>3.2. OBJETIVOS</b>	<b>108</b>
3.2.1. Objetivo Geral	108
3.2.2. Objetivos Específicos	108
<b>3.3. MATERIAIS E MÉTODOS</b>	<b>110</b>
3.3.1. Linhagens Celulares e Condições de Cultivo	110
3.3.2. Identificação in silico de lncRNAs associados ao câncer de mama	110
3.3.3. Ensaio de PCR quantitativo em tempo real (qRT-PCR)	110
3.3.3.1. Extração do RNA total	110
3.3.3.2. Desenho dos <i>primers</i>	110
3.3.3.3. Transcrição reversa e síntese de cDNA	111
3.3.3.4. Determinação da concentração final de <i>primers</i>	112
3.3.3.5. Determinação da eficiência dos <i>primers</i>	112
3.3.3.6. Confirmação da expressão diferencial	112
3.3.4. Seleção de RNA longo não codificador para análise funcional	112
3.3.5. Construção do vetor lentiviral pLVX-TetOne-Puro-LINC01133 para ensaio de superexpressão do <i>LINC01133</i>	112
3.3.5.1. Síntese da região codificadora do <i>LINC01133</i>	113
3.3.5.2. Transformação por eletroporação de bactérias <i>Escherichia coli</i> eletrocompetentes com DNA plasmideal	113
3.3.5.3. Preparação de DNA plasmideal em pequena escala	113
3.3.5.4. Análise da sequência de nucleotídeos	113
3.3.5.5. Sub-clonagem da construção no vetor lentiviral de quarta geração	114
3.3.6. Geração da linhagem MCF10A superexpressando o gene <i>LINC01133</i>	114
3.3.6.1. Produção de partículas lentivirais em células 293T	114
3.3.6.2. Titulação das partículas lentivirais resistentes a puomicina	114
3.3.6.3. Transdução viral da linhagem MCF10A com partículas lentivirais	115
3.3.7. Ensaio de metabolismo celular através da redução do sal tetrazólico (MTT)	115
3.3.8. Confirmação da superexpressão indutível do <i>LINC01133</i> nas linhagens transduzidas	116
3.3.9. Construção dos vetores lentivirais do sistema CRISPR/Cas9 para ensaio de nocauteamento do <i>LINC01133</i>	117
3.3.9.1. Seleção da região alvo e desenho dos <i>guides</i> RNAs	117
3.3.9.2. Fosforilação, anelamento e diluição dos <i>guides</i> RNAs	118
3.3.9.3. Digestão e desfosforilação dos plasmídeos lentivirais do sistema CRISPR/Cas9	118
3.3.9.4. Reação de ligação dos gRNAs aos vetores lentivirais	119
3.3.9.5. Transformação por eletroporação de bactérias <i>Escherichia coli</i> eletrocompetentes com DNA plasmideal	119

3.3.9.6. Preparação de DNA plasmideal em pequena escala	119
3.3.9.7. Análise da sequência de nucleotídeos	120
3.3.10. Geração da linhagem Hs578T nocauteada do gene <i>LINC01133</i>	120
3.3.10.1. Produção de partículas lentivirais em células 293T	120
3.3.10.2. Titulação das partículas lentivirais GFP positivas	120
3.3.10.3. Titulação das partículas lentivirais resistentes a puromicina	121
3.3.10.4. Transdução viral da linhagem Hs578T com partículas lentivirais	122
3.3.11. Confirmação do nocauteamento do <i>LINC01133</i> nas linhagens transduzidas	122
3.3.12. Caracterização das linhagens celulares geradas	123
3.3.12.1. Curva de crescimento celular	123
3.3.12.2. Ensaio de tumorigênese <i>in vitro</i>	123
3.3.12.2.1. Ensaio de crescimento de colônias em substrato sólido	123
3.3.12.2.2. Ensaio de crescimento de colônias em suspensão de agarose	123
3.3.12.3. Ensaio de migração celular <i>in vitro</i>	124
3.3.12.3.1. Ensaio de migração celular por <i>scratching</i> ( <i>wound healing</i> )	124
3.3.12.3.2. Ensaio de migração celular por <i>Transwell</i> <sup>TM</sup>	124
3.3.12.4. Ensaio de invasão celular <i>in vitro</i>	124
3.3.12.5. Imunofluorescência	125
3.3.12.6. Análise da expressão de genes envolvidos na malignidade tumoral	126
3.3.13. Análises Estatísticas	126
<b>3.4. RESULTADOS</b>	<b>127</b>
3.4.1. Identificação <i>in silico</i> de lncRNAs associados ao câncer de mama	127
3.4.2. Análise da expressão gênica dos lncRNAs selecionados no modelo de progressão tumoral do carcinoma mamário	127
3.4.3. Seleção de RNA longo não codificador (lncRNA) para análise funcional	130
3.4.4. Construção do vetor lentiviral pLVX-TetOne-Puro-LINC01133 para ensaio de superexpressão do <i>LINC01133</i>	130
3.4.5. Geração da linhagem MCF10A superexpressando o gene <i>LINC01133</i>	132
3.4.6. Ensaio de metabolismo celular através da redução do sal tetrazólico (MTT)	133
3.4.7. Confirmação da superexpressão indutível do <i>LINC01133</i> nas linhagens transduzidas	134
3.4.8. Construção dos vetores lentivirais do sistema CRISPR/Cas9 para ensaio de nocauteamento do <i>LINC01133</i>	135
3.4.8.1. Seleção da região alvo e desenho dos <i>guides</i> RNAs (gRNAs)	135
3.4.8.2. Digestão dos plasmídeos lentivirais e ligação dos gRNAs	137
3.4.9. Geração da linhagem Hs578T nocauteada para o gene <i>LINC01133</i>	138
3.4.10. Confirmação do nocauteamento do <i>LINC01133</i> nas linhagens transduzidas	140
3.4.11. Caracterização das linhagens celulares geradas	143
3.4.11.1. Curva de crescimento celular	143
3.4.11.2. Ensaio de tumorigênese <i>in vitro</i>	144
3.4.11.2.1. Ensaio de crescimento de colônias em substrato sólido	144

3.4.11.2.2. Ensaio de crescimento de colônias em suspensão de agarose	146
3.4.11.3. Ensaio de migração celular <i>in vitro</i>	148
3.4.11.3.1 Ensaio de migração celular por <i>scratching</i> ( <i>wound healing</i> )	148
3.4.11.3.2. Ensaio de migração celular por <i>Transwell™</i>	150
3.4.11.4. Ensaio de invasão celular <i>in vitro</i>	152
3.4.11.5. Imunofluorescência	154
3.4.11.6. Análise da expressão de genes envolvidos na malignidade tumoral	158
<b>3.5. DISCUSSÃO</b>	<b>160</b>
<b>3.6. CONCLUSÃO</b>	<b>180</b>
<b>3.7. PRÓXIMOS PASSOS</b>	<b>181</b>
<b>4. CONCLUSÕES GERAIS</b>	<b>183</b>
<b>5. REFERÊNCIAS BIBLIOGRÁFICAS</b>	<b>185</b>
<b>ANEXO – Súmula Curricular</b>	<b>202</b>

## 1. INTRODUÇÃO GERAL

### 1.1. Câncer

As células animais são formadas basicamente por três componentes importantes: a membrana celular, a qual funciona como uma barreira seletiva para a passagem de diversas moléculas e norteia os limites celulares, o citoplasma, região onde se encontra todas as organelas citoplasmáticas e onde ocorrem diversos processos celulares, e o núcleo, região responsável pelo armazenamento das informações genéticas, na forma de cromossomos. Os cromossomos abrigam os genes responsáveis por originar todas as moléculas funcionais presentes na célula, sejam elas proteínas ou RNAs. Estes genes podem sofrer mutações, que caso não sejam reparadas, podem interferir na função ou na estrutura do seu produto final e, conseqüentemente, interferir no funcionamento normal da célula. Algumas destas mutações podem ocorrer em genes específicos conhecidos como proto-oncogenes, culminando em sua ativação. Quando ativados, esses genes se transformam em oncogenes, responsáveis por dar origem ao câncer [Adamson, 1987; Weinstein & Joe, 2006].

Câncer é um termo utilizado para definir um vasto grupo de doenças que podem afetar qualquer região do corpo, tendo como características principais: sinalização proliferativa autossustentável, evasão de sinais de parada do crescimento celular, resistência à morte celular, imortalidade replicativa, atividade angiogênica, instabilidade genômica, potencial celular invasivo e metastático, entre outras [Hanahan *et al.*, 2011]. Os diferentes tipos de câncer são classificados de acordo com sua localização e os tecidos de origem. Se o câncer tiver origem em tecidos epiteliais, como pele ou mucosa, é denominado de carcinoma. Caso sua origem seja em tecidos conjuntivos, como osso, músculo ou cartilagem, é denominado de sarcoma. Por fim,

se o início for a partir do tecido glandular, é denominado de adenocarcinoma. Outras características que diferenciam os diferentes tipos de câncer entre si são a velocidade de multiplicação celular e a capacidade de invadir tecidos e órgãos vizinhos ou distantes (metástases) [Nall, 2015].

De acordo com estimativas da Organização Mundial da Saúde (WHO – *World Health Organization*) em 2019, o câncer foi a primeira ou segunda principal causa de mortes antes dos 70 anos em 112 de 183 países, e ocupa a terceira e quarta posições em outros 23 países, tornando-o um dos principais problemas de Saúde Pública Mundial [Bray *et al.*, 2021]. O envelhecimento e o crescimento populacional, além da mudança na distribuição e na prevalência dos fatores de risco, especialmente aqueles associados ao desenvolvimento socioeconômico, são fatores que podem ser citados como sendo responsáveis pelo aumento dos números de incidência e mortalidade ocasionados pela doença. É observada uma importante transição dos principais tipos de câncer detectados em países em desenvolvimento, com uma expressiva diminuição dos tipos de câncer associados a infecções e um aumento dos tipos relacionados à melhoria das condições socioeconômicas, como a incorporação de hábitos e atitudes associados à urbanização (alimentação inadequada, sedentarismo, entre outros) [Omran, 1971; Bray *et al.*, 2018].

Segundo dados da mais recente estimativa mundial, realizada no ano de 2020 pelo projeto GLOBOCAN da Agência Internacional de Pesquisa Sobre o Câncer (IARC - *International Agency for Research on Cancer*), foram registrados 19,3 milhões de novos casos de câncer (18,1 milhões excluindo os casos de câncer de pele não melanoma) e 10 milhões de óbitos (9,9 milhões excluindo os casos de câncer de pele não melanoma). Levando em consideração ambos os sexos, o câncer de mama é o de maior incidência no Mundo (11,7% do total de casos), seguido pelo câncer de

pulmão (11,4%), colorretal (10%), próstata (7,3%) e estômago (5,6%). Por outro lado, o câncer de pulmão é a principal causa de óbitos (18% do total de óbitos), seguido pelo colorretal (9,4%), fígado (8,3%), estômago (7,7%) e mama feminino (6,9%). Em homens, o câncer de maior incidência e com maior número de óbitos é o de pulmão (14,3% e 21,5%, respectivamente), seguido por próstata (14,1%) e colorretal (10,6%) para incidência, e fígado (10,5%) e colorretal (9,3%) para mortalidade. Em mulheres, o câncer de maior incidência e com maior número de óbitos é o câncer de mama (24,5% e 15,5%, respectivamente), seguido pelo colorretal (9,4%) e pulmão (8,4%) para incidência, e pulmão (13,7%) e colorretal (9,5%) para mortalidade [Sung *et al.*, 2021].

No Brasil, segundo estimativas para cada ano do triênio 2020-2022, realizada pelo Ministério da Saúde (MS) e pelo Instituto Nacional do Câncer José Alencar Gomes da Silva (INCA), ocorrerão 625 mil novos casos de câncer (450 mil excluindo os casos de câncer de pele não melanoma). Levando em consideração ambos os sexos, os cânceres de mama e de próstata serão os mais incidentes (66 mil novos casos), seguido por colorretal (41 mil), pulmão (30 mil) e estômago (21 mil), excluindo os casos de câncer de pele não melanoma (177 mil) [INCA, 2019]. Em relação ao número de óbitos, segundo estimativas do ano de 2019, o câncer de pulmão ocupa a primeira colocação (12,6%), seguido por colorretal (8,8%), mama feminino (7,7%), próstata (6,8%) e estômago (6,5%). Em homens, o câncer mais incidente é o de próstata (29,2%), seguido pelo colorretal (9,1%) e pulmão (7,9%). Para mortalidade, o câncer de pulmão ocupa a primeira colocação (13,8%), seguido pelo de próstata (13,1%) e colorretal (8,4%). Em mulheres, o câncer mais incidente e com maior número de óbitos é o câncer de mama (29,7% e 16,4%, respectivamente), seguido pelo colorretal (9,2%) e colo do útero (7,5%) para incidência, e pulmão (11,4%) e

colorretal (9,4%) para mortalidade [MS – Sistema de Informação sobre Mortalidade, 2021].

Os números descritos anteriormente revelam uma informação importante: o câncer de mama ultrapassou o câncer de pulmão como o de maior incidência no Mundo, quando comparamos a estimativa realizada no ano de 2020 com a estimativa realizada no ano de 2018 pelo projeto GLOBOCAN da IARC [Bray *et al.*, 2018; Sung *et al.*, 2021]. Além disto, o câncer de mama feminino ocupa posição de destaque em relação ao número de óbitos, seja no Mundo (5º colocação) ou no Brasil (3º colocação). Quando se leva em consideração apenas as mulheres, esses números se tornam ainda mais evidentes, com o câncer de mama ocupando as primeiras colocações, tanto para incidência quanto para mortalidade, seja no Mundo ou no Brasil. Esses dados indicam a necessidade de desenvolvimento de mais estudos relacionados a este tipo de câncer, visando compreender as bases moleculares da doença, contribuindo, desta forma, para a diminuição de suas altas taxas de incidência e mortalidade.

## 1.2. Câncer de mama

O câncer de mama é uma doença heterogênea caracterizada pela multiplicação desordenada das células mamárias que adquiriram características anormais causadas por mutações no material genético da célula. Esta doença pode ter características patológicas variáveis, diferentes respostas à terapias e diferenças substanciais na sobrevida do paciente a longo prazo. A heterogeneidade observada entre os cânceres de mama reflete a ideia de que não é apenas uma doença com alguns subtipos variantes, mas que o câncer de mama representa um conjunto de doenças neoplásicas distintas da mama e de suas células constituintes. Clinicamente,

o câncer se desenvolve por um longo período de tempo, requer diversas mudanças moleculares e envolve a evolução de populações de células com características fenotípicas cada vez mais agressivas. Embora o tempo necessário para o processo de carcinogênese não esteja bem estabelecido, as estimativas sugerem que esse processo leva muitos anos e possivelmente várias décadas [Riverbank *et al.*,2013]. Pesquisas indicam que as células do câncer de mama sobrevivem na forma latente por um longo tempo, no qual param de se dividir, mas sobrevivem em um estado quiescente enquanto esperam por condições adequadas para iniciar a proliferação [Ono *et al.*, 2014].

Em 2020, 2,3 milhões de mulheres foram diagnosticadas com câncer de mama, ocorrendo em torno de 685.000 mortes em todo o Mundo. No final de 2020, tornou-se o câncer mais prevalente do Mundo, visto que 7,8 milhões de mulheres vivas foram diagnosticadas nos últimos cinco anos. É uma doença presente em todos os países do mundo, ocorrendo em mulheres de qualquer idade após a puberdade, mas com taxas mais elevadas em mulheres adultas [WHO, 2021]. Raramente, o câncer de mama atinge homens, representando apenas 1% de todos os casos da doença [Korde *et al.*,2010]. O sintoma mais comum do câncer de mama é o aparecimento de nódulos, geralmente indolores, duros e irregulares, embora existam tumores moles, globulares e bem definidos. Outros sintomas incluem dor, hiperemia, inversão mamilar, alterações da pele da mama (abaulamento ou retração dando à pele o aspecto de casca de laranja), descamação ou ulceração mamilar, secreção papilar (podendo ser rosa ou vermelha devido à presença de glóbulos vermelhos) e presença de linfonodos palpáveis na axila [Adami *et al.*, 2008].

A doença não apresenta uma causa única e diversos fatores de risco podem contribuir para seu desenvolvimento, sendo eles: a) Idade: mulheres mais velhas, a

partir dos 50 anos de idade, exibem maior risco devido ao envelhecimento e ao acúmulo de mutações ao longo da vida [Silva & Silva, 2005]; b) Fatores genéticos/hereditários: relacionados a mulheres com mutações nos genes *BRCAs* (*BRCA1* e *BRCA2* - mais frequente), *CHEK2*, *ATM*, *RAD51*, *TP53*, entre outros [Breast Cancer Association Consortium *et al.*, 2021] e a mulheres com casos de câncer de mama e/ou câncer de ovário detectados em parentes co-sanguíneos [Adami *et al.*, 2008]; c) Fatores endócrinos/vida reprodutiva: relacionados a mulheres com histórico de menarca precoce (menstruação antes dos 12 anos), menopausa tardia (após os 55 anos), nuliparidade, primeira gravidez após os 30 anos, uso de contraceptivos orais e terapia de reposição hormonal pós-menopausa [Silva & Silva, 2005]; d) Fatores comportamentais/ambientais: relacionados à ingestão de bebidas alcoólicas, ao tabagismo, à obesidade e à exposição à radiação ionizante [Inumaru *et al.*, 2011; Nothaisintawee *et al.*, 2013]; e e) Densidade do tecido mamário: relacionado à maior proporção de tecido fibroglandular na mama quando comparado com o tecido adiposo [Boyd *et al.*, 1998].

Aumentos na incidência, sobrevida e mortalidade são observadas em todo o Mundo, variando consideravelmente entre grupos raciais e étnicos, assim como, também, entre países desenvolvidos e em desenvolvimento. Alguns fatores que contribuem para essa variação são: a) o sítio de origem do câncer; b) a prevalência dos fatores de risco; c) o acesso a cuidados de saúde de alta qualidade, incluindo a prevenção, o diagnóstico precoce e a escolha do tratamento ideal, sendo a principal fonte dessa desigualdade o alto índice de pobreza observado na comunidade negra. Mulheres negras apresentam maior risco de morte por câncer de mama devido à maior prevalência de comorbidades, ao maior índice de massa corporal, a maior prevalência de características tumorais mais agressivas, e por fim, a opções de tratamento de

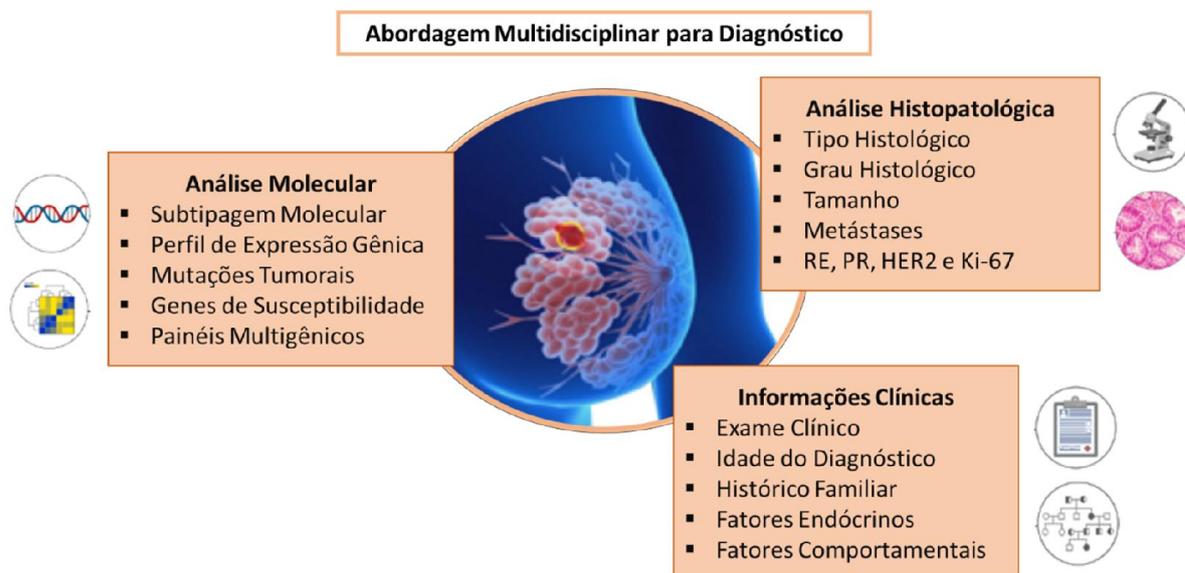
baixa qualidade [Siegel *et al.*, 2016]. Em relação à mortalidade, um outro fator que pode contribuir para seu aumento é o atraso no diagnóstico devido às limitações da mamografia, recomendada para mulheres com mais de 50 anos, mas que em países em desenvolvimento ou em áreas rurais, o acesso a clínicas é difícil, além do fato de mulheres mais jovens não terem o hábito de realizar auto-exame [Dündar *et al.*, 2006]. Nos últimos 40 anos, estratégias de diagnóstico precoce tem elevado a sobrevida, em países desenvolvidos, exibindo taxas de 85% em cinco anos. Nos países em desenvolvimento, estas taxas permanecem na faixa de 50 à 60%, possivelmente devido ao diagnóstico da doença ser feito em estágios mais avançados [Da Costa Vieira *et al.*, 2017].

O principal fator que contribui para a mortalidade por câncer de mama é a presença de metástases, responsáveis por cerca 400.000 mortes anualmente em todo o Mundo [Kamangar *et al.*, 2006]. Cerca de 30% dos casos de câncer de mama evoluem para metástases, com 90% delas resultando em morte. Os principais sítios de metástase são os ossos, os pulmões, o cérebro e o fígado, mas o câncer de mama também é capaz de se metastizar para diversos outros sítios. A possibilidade de ocorrer metástase depende fundamentalmente de dois fatores: a extensão/volume da doença no diagnóstico (tamanho do tumor e grau de envolvimento linfonodal) e também do subtipo tumoral, visto que existem tumores mais agressivos do que outros [Weigelt *et al.*, 2005; Foulkes *et al.*, 2010].

Estudos mostram que a detecção precoce e a escolha das melhores opções de tratamento pode elevar a probabilidade de sobrevida em até 90% e reduzir a mortalidade por câncer de mama em até 30% [Kamangar *et al.*, 2006]. As opções de tratamento geralmente são a cirurgia, a radioterapia, o transplante de medula óssea e a terapia sistêmica (medicamentos anticâncer administrados por via oral ou

intravenosa no intuito de tratar e/ou reduzir o risco de metástases). Estes medicamentos incluem a terapia endócrina (hormonal), a quimioterapia e, em alguns casos, terapia direcionadas (por exemplo, terapia com anticorpos monoclonais). Em muitos casos, é necessário a combinação de mais de uma modalidade [Velazquez *et al.*, 2018; Caparica *et al.*, 2019; Mutebi *et al.*, 2020].

No entanto, é extremamente difícil diferenciar morfológicamente os diferentes tipos de tumores mamários e diagnosticá-los corretamente é de extrema importância. Inicialmente, o diagnóstico era baseado apenas em informações clínicas. Atualmente, o diagnóstico é realizado através de abordagens multidisciplinares, envolvendo, além das informações clínicas, classificações histopatológicas e moleculares dos tumores (Figura 1). Estas classificações são importantes por diversas razões: a) na gestão clínica são utilizadas para decidir ou padronizar o atendimento e o tratamento do paciente, contribuindo, desta forma, na tomada de decisão correta pelo clínico; b) em ensaios clínicos são utilizadas para avaliar a resposta à terapia; e c) em ensaios epidemiológicos e funcionais são importantes por fornecer mais informações sobre os mecanismos de desenvolvimento e evolução tumoral durante o tratamento. Diversos fatores são avaliados para se ter uma classificação precisa do câncer de mama, tais como a origem, o tamanho, o potencial de avanço, a expressão de receptores hormonais (estrogênio e progesterona) e do receptor do fator de crescimento epidermal humano tipo 2 (HER2), a análise de mutações tumorais, a análise de genes de susceptibilidade ao câncer, entre outros. Baseado nessas características, o câncer de mama pode ser classificado em tipo ou grau histológico, por estadiamento e, também, molecularmente [Sorlie *et al.*, 2001; Rusness *et al.*, 2017].



**Figura 1: Diagnóstico Multidisciplinar.** O diagnóstico de câncer de mama é composto por vários componentes (informações clínicas, análises histopatológicas e análises moleculares), importantes por fornecer a base para a decisão do tratamento padrão, planejar o acompanhamento e fortalecer pesquisas translacionais [adaptado de Szymiczek *et al.*, 2021]. ER: receptor de estrogênio; PR: receptor de progesterona; HER2: receptor do fator de crescimento epidermal humano tipo 2.

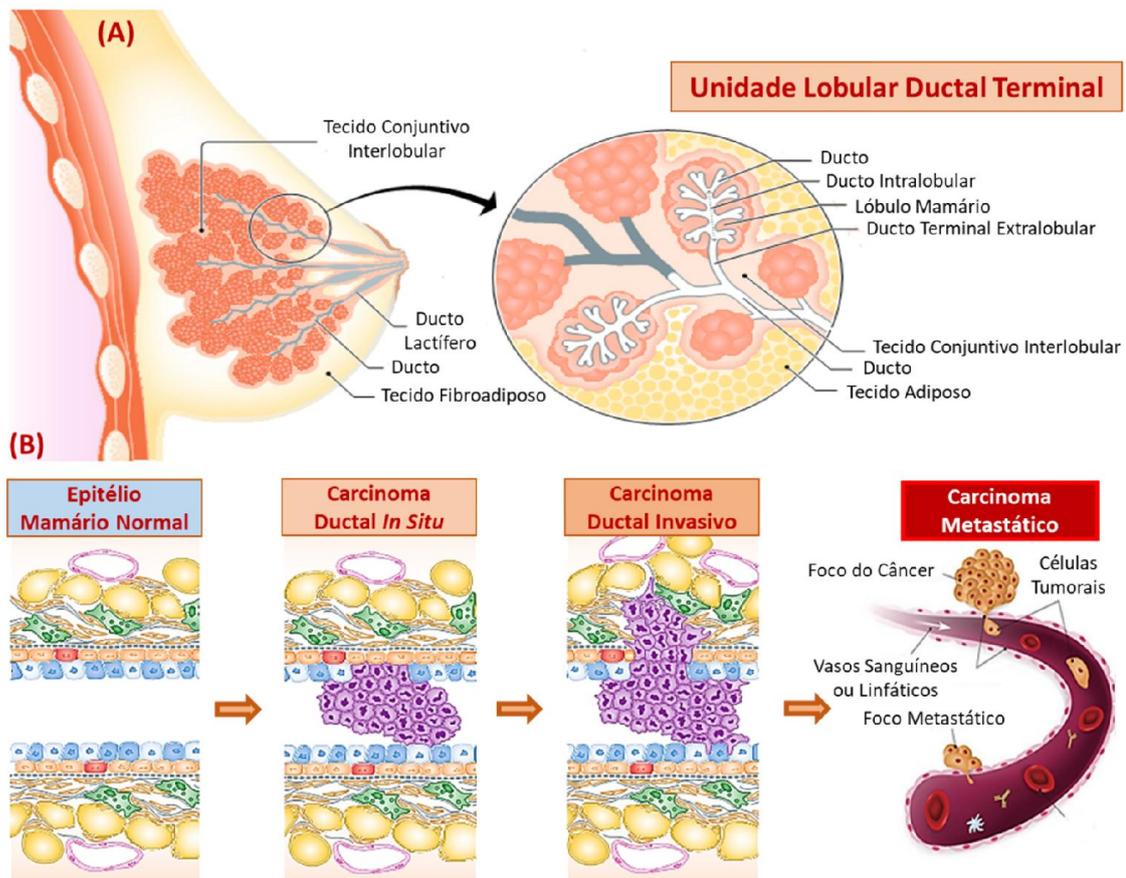
### 1.2.1. Classificação em Tipos Histológicos

As glândulas mamárias são glândulas sudoríparas apócrinas modificadas. Elas estão situadas na parede torácica anterior e são compostas por tecido epitelial glandular, tecido conjuntivo fibroso e tecido adiposo. O tecido glandular é constituído por 8-15 lobos secretores, os quais são formados por numerosos lóbulos e estes por ácinos. Diversos ductos principais e secundários conectam as unidades lobulares secretoras de leite aos mamilos. Pequenos ductos percorrem toda a mama, convergindo em ductos coletores largos (ductos lactíferos) que se abrem nos seios lactíferos localizados na base do mamilo, o qual é responsável por auxiliar na função da mama. Os mamilos são cercados por uma região circular de pele pigmentada, chamada de aréola, que contém pequenas elevações pontuais em sua superfície, as quais são produzidas pelas diversas glândulas areolares [Pandya & Moore, 2011; Hassiotou & Geddes, 2013].

Na região posterior das aréolas encontra-se o plexo linfático sub-areolar, responsável por drenar a linfa proveniente dos lobos, dos mamilos e das auréolas. Partindo deste plexo, a drenagem linfática da mama pode seguir três diferentes caminhos: cerca de 75% da linfa é drenada para os linfonodos axilares, enquanto o restante drena ou para os linfonodos paraesternais ou para os linfonodos intercostais [Edition *et al.*, 2017]. A maior parte dos cânceres originam-se inicialmente nas unidades lobulares ductais terminais da mama, mas podem se disseminar para qualquer tecido ou órgão do corpo através do sistema vascular sanguíneo ou linfático, ocasionando as metástases. Baseado nisto, o câncer de mama pode se apresentar de três formas distintas: carcinoma não-invasivo ou *in situ*, no qual as células tumorais estão limitadas aos componentes epiteliais da mama (ductos ou lóbulos mamários); carcinoma invasivo, no qual as células tumorais migram dos ductos ou lóbulos para o tecido mamário circundante; e carcinoma metastático, no qual as células tumorais migram para os linfonodos próximos (metástases regionais) ou para a corrente sanguínea e atingem outros órgãos do corpo (metástases distantes) (Figura 2) [Allred *et al.*, 2001].

Os carcinomas invasivos são subdivididos morfológicamente de acordo com seus padrões de crescimento, tipo de célula tumoral, secreção extracelular, características arquitetônicas e perfil imunohistoquímico. Essas características irão definir se o tumor é ductal ou lobular, além de definir também suas subclassificações [Lakhani *et al.*, 2012; Makki, 2015]. Cerca de 50 a 80% dos tumores de mama são classificados como carcinoma ductal invasivo “sem tipo específico” (IDC-NST – *invasive ductal carcinoma no specific type*), visto que não possuem características morfológicas suficientes para serem classificados como um tipo histológico característico. Por outro lado, 25% dos carcinomas invasivos são classificados como

de “tipo específico”, ou seja, possuem características morfológicas e comportamentos celulares/moleculares específicos [Henry & Cannon-Allbright, 2019]. São relativamente incomuns, estando inclusos nesse grupo o carcinoma lobular invasivo, carcinoma medular, carcinoma metaplásico, carcinoma apócrino, carcinoma cribriforme, carcinoma mucinoso, carcinoma micropapilar, carcinoma tubular, carcinoma adenóide cístico, carcinoma secretor, carcinoma neuroendócrino, entre outros [Masood, 2016]. Como consequência, a função da tipagem histológica na tomada de decisão de manejo clínico é atualmente limitada, sendo necessário avaliar parâmetros adicionais [Pereira *et al.*, 1995].



**Figura 2: A mama: estrutura e estágios do desenvolvimento tumoral.** (A) Esquema da mama, destacando as unidades lobulares ductais terminais. (B) Esquema de um corte transversal ductal representando a progressão do câncer de mama, referente a classificação em tipos histológicos [adaptado de Tharmapalan *et al.*, 2019].

### 1.2.2. Classificação em Grau Histológico

O grau histológico é um dos fatores prognósticos mais bem estabelecidos no câncer de mama. Ele representa o grau de diferenciação do tecido tumoral, medido através da avaliação de características morfológicas, e tem se mostrado capaz de gerar informações importantes relacionadas ao comportamento clínico dos cânceres de mama [Rakha *et al.*, 2010]. Diferentes sistemas de pontuação estão disponíveis para avaliação do grau histológico, sendo o Sistema de Graduação de Nottingham o mais amplamente utilizado. Este sistema é baseado na avaliação de três características morfológicas: a) grau de formação de túbulos ou glândulas; b) pleomorfismo nuclear; e c) contagem mitótica [Elston & Ellis, 1991; Elston & Ellis, 1993]. Cada característica é avaliada e recebe uma pontuação, as quais são somadas para se obter um número entre 3 e 9. Este número é utilizado para classificar os tumores em três diferentes graus histológicos: a) Grau I ou bem diferenciado (pontuação de 3, 4 ou 5), no qual as células cancerosas tem crescimento mais lento e são semelhantes ao tecido mamário normal; b) Grau II ou moderadamente diferenciado (pontuação de 6 ou 7), no qual as células cancerosas crescem numa velocidade intermediária e são em parte semelhantes ao tecido mamário normal e na outra parte semelhante às células cancerosas; e c) Grau III ou pouco diferenciado (pontuação de 8 ou 9), no qual as células cancerosas são bastante diferentes do tecido mamário normal e, provavelmente, crescem e se espalham rapidamente [Edition *et al.*, 2017].

### 1.2.3. Classificação por Estadiamento

O estadiamento é uma forma de descrever a extensão do câncer de mama, sendo utilizado para determinar o prognóstico. A ferramenta mais aceita

internacionalmente para descrever o estadiamento é o sistema TNM, desenvolvido pelo *American Joint Committee on Cancer* (AJCC). Este sistema, proposto inicialmente no ano de 1959, tem como objetivo uniformizar a descrição do câncer de mama, contribuindo para que os resultados do tratamento de todos os pacientes possam ser comparados e compreendidos. No passado, três características clínicas e patológicas foram utilizadas para descrever o estágio do câncer de mama: a) o tamanho do tumor (T); b) a disseminação para gânglios linfáticos próximos (N); e c) a disseminação para órgãos distantes – metástase (M) [Edition *et al.*, 2017].

Ao longo das décadas o sistema foi aperfeiçoado, devido ao avanço nas ciências laboratoriais e nas pesquisas translacionais, tornando a determinação do estadiamento mais complexa, porém mais precisa. Na última atualização, no ano de 2018, novas características foram incorporadas para descrever o estágio do câncer de mama: o grau histológico e a classificação molecular, a qual baseia-se no *status* dos receptores hormonais de estrogênio (ER) e progesterona (PR) e do receptor do fator de crescimento epidermal humano tipo 2 (HER2). A combinação de todos estes fatores estratifica a doença em 5 estágios (0, I, II, III e IV), sendo o estágio 0 referente aos carcinomas *in situ* e os estágios I, II, III e IV referentes aos carcinomas invasivos. Quanto maior o estágio, mais avançada se encontra a doença [Weiss *et al.*, 2018].

#### 1.2.4. Classificação Molecular

Na era das tecnologias de alto rendimento, as classificações descritas anteriormente, baseadas em parâmetros morfológicos e patológicos clínicos, tornaram-se insuficientes para prever o real comportamento da fisiopatologia dos tumores mamários, visto que muitos deles apresentam características morfológicas semelhantes, mas com diferentes prognósticos e respostas à terapias [Van't Veer *et*

*al.*, 2002; Sorlie *et al.*, 2003; Masood, 2016; Fragomeni *et al.*, 2018]. Buscando contornar este problema e com o objetivo de agrupar esses tumores em classes/grupos específicos, auxiliando, desta forma, no manejo clínico e na elaboração de estudos funcionais/epidemiológicos, diversos estudos baseados na análise de padrões moleculares no câncer de mama foram propostos [Perou *et al.*, 2000; Prat & Perou, 2011; Sorlie *et al.*, 2001; Holliday & Speirs, 2011; Lehmann *et al.*, 2011; Burstein *et al.*, 2015].

Baseado na similaridade do perfil de expressão gênica, utilizando a técnica de cDNA microarray, Perou e colaboradores foram os pioneiros, publicando um estudo classificando os tumores de mama em cinco subtipos intrínsecos: Luminal-A, Luminal-B, HER2-positivo, *Basal-like* (Triplo Negativo) e *Normal Breast-like*. A classificação nestes subtipos foi baseada, inicialmente, na expressão de três tipos de receptores: os receptores hormonais de estrogênio (ER) e de progesterona (PR) e o receptor do fator de crescimento epidermal humano tipo 2 (HER2) [Perou *et al.*, 2000]. Estes diferentes subtipos mostraram diferenças na incidência, nas características clínicas e patológicas, e em grande parte, se sobrepõem com as classificações clínicas e histopatológicas já estabelecidas [Tsang & Tse, 2020]. No entanto, a diferenciação dos subtipos Luminal A do Luminal B resultou em importantes implicações terapêuticas e, no ano de 2013, buscando contornar este problema, as diretrizes da conferência de *Saint Gallen* recomendaram a avaliação adicional do índice de proliferação celular através da análise do marcador Ki-67 [Coates *et al.*, 2015]. Além disto, no ano de 2008, Peppercorn e colaboradores classificaram o subtipo *normal breast-like*, caracterizado pela expressão de genes semelhantes ao epitélio mamário normal, como controverso e ele passou a ser considerado um artefato, levando à sua exclusão de parte das classificações atuais [Peppercorn *et al.*, 2008].

A busca pelo entendimento preciso da biologia do tumor, além da crescente necessidade de aperfeiçoamento da estratificação de risco e da determinação precisa do prognóstico, levou ao desenvolvimento de diversos painéis multigênicos, tais como o *Mammaprint*®, o *Oncotype Dx*® e o *PAM50 (ProSigma)*®. Estes painéis baseiam-se na assinatura de diversos genes e estão sendo utilizados, por exemplo, para avaliar o risco de recorrência da doença e a eficácia da terapia adjuvante [Van De Vijver *et al.*, 2002; Paik *et al.*, 2004; Vieira & Schmitt, 2018].

Embora a análise dos perfis de expressão gênica tenha se tornado uma técnica de laboratório corriqueira, os custos e a complexidade da técnica limitam sua aplicação na prática clínica diária. Visando solucionar este problema, uma classificação molecular baseada em imunohistoquímica (IHC) vem sendo proposta e utilizada. Ela baseia-se em um painel com os quatro biomarcadores descritos anteriormente (ER, PR, HER2 e Ki-67) e foi considerada eficiente e significativa na identificação de diferentes subtipos moleculares [Goldhirsch *et al.*, 2013; Vuong *et al.*, 2014; Tsang & Tse, 2020]. Ambas as classificações moleculares, seja a baseada no perfil de expressão gênica ou aquela baseada em IHC, dividem atualmente os tumores de mama em quatro subtipos intrínsecos: Luminal A, Luminal B, HER2+, e *Basal-Like* (Triplo Negativo) (Tabela 1). Estas classificações, seja a baseada no perfil de expressão gênica ou a baseada em imunohistoquímica, apresentam elevada sobreposição, mas algumas discrepâncias são observadas e serão descritas posteriormente (relacionadas ao subtipo HER2+ e ao *Basal-Like*) [Prat *et al.*, 2014].

O subtipo Luminal A é o mais comum e representa cerca de 40% a 50% dos casos de câncer de mama invasivo [Voduc *et al.*, 2010]. São considerados de grau histológico I ou II (bem ou moderadamente diferenciado), possuindo características de células epiteliais luminais da mama, tais como a alta expressão das citoqueratinas

7, 8, 18 e 19 [Gao & Swain, 2018]. Apresentam como perfil imunohistoquímico: ER+(≥1%), alta expressão de PR+(≥20%), HER2-(≤10%) e baixa expressão do Ki-67(<14%) [Goldhirsch *et al.*, 2013]. Além disto, possuem o melhor prognóstico entre todos os subtipos, e, por apresentarem status positivo para os receptores hormonais, os pacientes podem ser tratados com terapias endócrinas, tais como o tamoxifeno, considerado um modulador do receptor de estrogênio, ou com o anastrozol, um inibidor de aromatase que bloqueia a produção do estrogênio [Rocca *et al.*, 2015].

O subtipo Luminal B é responsável por cerca de 20% a 30% dos casos de câncer de mama invasivo. São considerados de grau histológico II (moderadamente diferenciado) e possuem baixa expressão das citoqueratinas 7, 8, 18 e 19 [Feng *et al.*, 2018]. Podem ser categorizados em dois tipos de perfil imunohistoquímico: Luminal B (HER-), o qual apresenta ER+(≥1%), PR-(<20%), HER2-(≤10%) e alta expressão do Ki-67(<14%); ou Luminal B (HER2+), o qual apresenta ER+(≥1%), HER2+(>10%) e níveis de expressão variáveis para PR e Ki-67 [Goldhirsch *et al.*, 2013; Cheang *et al.*, 2009]. São a forma mais agressiva dos casos de câncer de mama dependente de hormônios e demonstram ter um pior prognóstico quando comparados com o subtipo Luminal A. Os pacientes deste subtipo requerem tratamentos adicionais em relação à terapia endócrina, tais como a quimioterapia (para HER2+/HER2-) ou terapia alvo (para HER2+) [Fragomeni *et al.*, 2018].

O subtipo HER2+ representa cerca de 15% dos casos de câncer de mama invasivo [Perou *et al.*, 2000; Cho, 2016]. São considerados de grau histológico III (pouco diferenciado) e apresentam como perfil imunohistoquímico: alta expressão de HER2 (quase sempre atribuída à amplificação do gene *HER2*) (>10%), ausência da expressão do ER-(<1%) e do PR-(<20%) e alta expressão do Ki-67(>20%) [Goldhirsch *et al.*, 2013; Cho, 2016]. Uma discrepância observada é que apenas 65% dos tumores

HER2+ pertencem a este subtipo [Prat *et al.*, 2014]. As características listadas acima tornam estes tumores propensos a metástases, levando a uma sobrevida livre da doença curta e ao segundo pior prognóstico entre os quatro subtipos. Apesar disto, apresentam resposta terapêutica positiva a medicamentos que bloqueiam a atividade do HER2, como inibidores de receptores moleculares de tirosina quinase (Lapatinib) e anticorpos monoclonais humanizados (Trastuzumab). A combinação destas drogas com a quimioterapia citotóxica tem melhorado drasticamente os resultados a longo prazo para os pacientes deste subtipo de câncer de mama [Mukai, 2010; Llombart-Cussac *et al.*, 2017; Tsang & Tse, 2020].

O subtipo *Basal-Like* (BLBC – *Basal-like Breast Cancer*) representa aproximadamente cerca de 15% dos casos de câncer de mama invasivo [Eliyatkin *et al.*, 2015]. Apresentam grau histológico III (pouco diferenciado) e elevada expressão de genes característicos das células basais/mioepiteliais da mama, incluindo citoqueratinas de alto peso molecular (CK 5/6, 14 e 17), ciclina E, P-caderina, receptor do fator de crescimento epidermal (EGFR) e receptor de tirosina quinase (c-Kit) [Provenzano *et al.*, 2018; Tsang & Tse, 2020]. Altas taxas proliferativas são observadas neste subtipo devido à elevada expressão do Ki-67 (>30%). São também denominados de Triplo Negativos (TNBC – *Triple-negative Breast Cancer*) devido à ausência da expressão dos receptores de estrogênio (ER<1%), de progesterona (PR<20%) e do HER2 (≤10%) [Badve *et al.*, 2011]. A elevada heterogeneidade, seja ela morfológica, molecular ou clínica, aliada à ausência da expressão dos receptores citados anteriormente, tornam este subtipo um desafio para o manejo clínico, principalmente no que diz respeito às opções de tratamento, as quais são limitadas, e à ausência de estratégias terapêuticas direcionadas. Devido a isto, apresentam o

pior prognóstico entre todos os subtipos, refletido na elevada taxa de mortalidade observada nos pacientes [Rusness et al., 2017].

**Tabela 1: Visão geral dos diferentes subtipos moleculares do câncer de mama.**

SUBTIPOS MOLECULARES	LUMINAL A	LUMINAL B		HER2+	BLBC (TNBC)
		(HER2-)	(HER2+)		
FREQUÊNCIA	40 a 50%	20 a 30%		15 a 20%	15%
BIOMARCADORES	ER+ PR+ HER2- Ki-67-	ER+ PR- HER2- Ki-67+	ER+ PR+/- HER2+ Ki-67+/-	ER- PR- HER2+ Ki-67+	ER- PR- HER2- Ki-67+
GRAU HISTOLÓGICO	Grau I ou II	Grau II		Grau III	Grau III
TIPO HISTOLÓGICO	CT CLlc IDC-NSTb	CMic CLlp IDC-NST		CLlp IDC-NSTa	CMet CMed IDC-NSTa
PROGNÓSTICO	Bom	Intermediário		Ruim	Ruim
LINHAGENS CELULARES	MCF7 SUM185	ZR-75-1 BT474		SK-BR-3 MDA-MB-453	Hs578T MDA-MB-231

BLBC: basal-like breast cancer; TN: triple-negative breast cancer; ER: receptor de estrogênio; PR: receptor de progesterona; HER2: receptor do fator de crescimento epidermal humano tipo 2; CT: carcinoma tubular; CLlc: carcinoma lobular invasivo clássico; CLlp: carcinoma lobular invasivo pleomórfico; CMic: carcinoma micropapilar; CMet: carcinoma metaplástico; CMed: carcinoma medular; IDC-NSTa: carcinoma ductal infiltrante sem tipo específico de alto grau; IDC-NSTb: carcinoma ductal infiltrante sem tipo específico de baixo grau [adaptado de Holliday & Speirs, 2011; Tsang & Tse, 2020].

### 1.3. Câncer de mama triplo negativo (TNBC)

Os cânceres de mama triplo negativos (TNBCs) são definidos como tumores que não apresentam expressão significativa dos receptores hormonais de estrogênio (ER) e progesterona (PR) e do receptor do fator de crescimento epidermal humano tipo 2 (HER2). Esses tumores são caracterizados por apresentar um considerável aumento da expressão de citoqueratinas de alto peso molecular e desregulação em genes supressores tumorais importantes, tais como *p53*, *pRb*, *p16* e *BRCA1* (Bianchini *et al.*, 2016). Além disto, são enriquecidos para marcadores de transição epitélio-mesenquimal e de células-tronco tumorais, exibindo ainda altos índices mitóticos [Cheung *et al.*, 2015]. Os termos BLBC e TNBC tem sido usados como

sinônimos devido à semelhança de seus fenótipos, entretanto, nem todos os TNBC são BLBC. Análises recentes frequentemente classificam o TNBC como um subtipo do BLBC, devido à taxa de sobreposição de seus perfis de expressão gênica que é de aproximadamente 56%, podendo chegar até 90%, em relação a apenas 11,5% quando comparamos tumores não-TNBC e BLBC. Bertucci e colaboradores mostraram que cerca de 71% dos TNBC eram BLBC, enquanto que 77% dos BLBC eram TNBC [Bertucci *et al.*, 2008; Foulkes *et al.*, 2010; Prat *et al.*, 2015]. Histologicamente, cerca de 95% dos casos de TNBC são classificados como carcinoma ductal invasivo “sem tipo específico” (IDC-NST). Os outros 5% dos casos são referentes ao carcinoma lobular invasivo (1-2% dos casos) e aos outros carcinomas de “tipo específico” (1% ou menos dos casos) [Weigelt & Reis-Filho, 2009].

Dados epidemiológicos mostram que estes tumores ocorrem com maior frequência em mulheres mais jovens (<50 anos de idade) e descendentes de afro-americanos ou hispânicos do que de outros grupos raciais ou étnicos, em mulheres múltiparas e em mulheres que apresentam mutações no gene *BRCA1* [Foulkes *et al.*, 2010; Badve *et al.*, 2011]. Os genes supressores tumorais *BRCA1* e *BRCA2* são considerados de extrema importância para a integridade do genoma, desempenhando um papel importante no reparo por recombinação homóloga. O gene *BRCA1* apresenta um papel mais amplo e controla diversos processos celulares em resposta ao dano no DNA, tais como o reconhecimento do dano, as vias de transdução de sinais envolvidas no reparo, o recrutamento de proteínas de reparo, entre outros. Enquanto isto, o gene *BRCA2* tem como papel principal recrutar a DNA recombinase RAD51 e direcioná-la para a região danificada do DNA [Filipponi & Bulavin, 2013; Savage & Harkin, 2015; D’Alessandro *et al.*, 2018]. Mutações que ocasionam a perda de função dos genes *BRCA* levam à sua inativação, ocasionando instabilidade

genômica, devido à parada do ciclo celular e perda do reparo das quebras da dupla fita do DNA levando, assim, à predisposição a diversos tipos de câncer, como os de ovário, próstata e mama [Foulkes *et al.*, 2003]. Estudos anteriores mostraram que mais de 75% dos cânceres de mama em mulheres que possuem mutações BRCA1 na linhagem germinativa são TNBC [Quinn *et al.*, 2003; Rottenberg *et al.*, 2008; Andres *et al.*, 2014], enquanto que 10 à 31% dos TNBC tem mutações BRCA1/2 na linhagem germinativa [Greenup *et al.*, 2013; Sharma *et al.*, 2014]. A fim de manter a viabilidade celular, quando estes genes estão defeituosos/inativos, o reparo é realizado através do mecanismo de união de extremidade não-homóloga (NHEJ, da sigla em inglês), o qual é dependente da enzima poli-ADP ribose polimerase 1 (PARP1). Esta enzima nuclear é expressa constitutivamente e ativa vias de sinalização intracelulares que modulam o reparo do DNA nestas células [Prakash *et al.*, 2015; Ronson *et al.*, 2018].

O TNBC apresenta comportamento clínico altamente agressivo levando ao pior prognóstico entre os subtipos. O tempo médio de sobrevivência de pacientes com TNBC é menor quando comparado com os outros subtipos de câncer de mama, e a taxa de mortalidade é de aproximadamente 40% nos primeiros cinco anos após o diagnóstico [Dent *et al.*, 2007]. A taxa de recidiva após a cirurgia é maior do que 25% [Lin *et al.*, 2008]. O tempo médio de recidiva em pacientes com TNBC é de 19 à 40 meses, enquanto que para os outros subtipos é de 35 à 67 meses. Após a recidiva, o tempo médio de sobrevivência é de apenas nove meses e a taxa de mortalidade dentro dos três primeiros meses é de alarmantes 75% [Haffty *et al.*, 2006; Gluz *et al.*, 2009; Zhang *et al.*, 2015]. Estes tumores são extremamente invasivos com alta tendência a metástases precoces e distantes, geralmente ocorrendo nos primeiros três anos após o diagnóstico. Cerca de 45% dos pacientes com TNBC terão metástases distantes, geralmente envolvendo o cérebro (30%) e órgãos viscerais

como o pulmão (40%) e o fígado (20%). Apenas 10% das metástase distantes atingem os ossos, enquanto que para os outros subtipos este valor chega a 40% [Foulkes *et al.*, 2010]. O tempo médio de sobrevivência após a metástase é de aproximadamente um ano [Lin *et al.*, 2008].

Uma outra característica interessante observada neste subtipo é a sua elevada heterogeneidade, a maior entre todos os subtipos, com duas hipóteses tendo surgido para tentar explicá-la. A primeira hipótese baseia-se na evolução clonal e infere que mutações aleatórias e a seleção clonal dão origem à heterogeneidade observadas nestes tumores [Greaves & Maley, 2012]. A segunda hipótese (mais aceita) baseia-se nas células-tronco tumorais e postula que a hierarquia celular e a diversidade tumoral são oriundas das células-tronco do câncer de mama (BCSCs – *Breast Cancer Stem Cells*), pequeno grupo de células com capacidade de auto-renovação e habilidade de originar diferentes tipos celulares [Vermeulen *et al.*, 2008]. Este grupo de células, importantes por revelar a origem celular, a manutenção tumoral e a progressão do câncer de mama, apresentam um perfil de marcadores de superfície celular caracterizado pelo fenótipo CD44+/CD24- (elevados níveis de expressão do CD44 e baixos/ausentes níveis de expressão do CD24), além da expressão da enzima aldeído desidrogenase 1 (ALDH1+) [Morrison *et al.*, 2008]. Os diferentes subtipos do câncer de mama exibem diferentes abundâncias das BCSCs, e os TNBCs abrigam a maior proporção quando comparado com os outros subtipos, contribuindo, desta forma, para o seu pior prognóstico. São estas células responsáveis pelas elevadas taxas de recidiva e pela resistência ao tratamento associados ao TNBC, visto que exibem resistência à radiação, quimioterapia e terapias moleculares direcionadas [Yu *et al.*, 2013].

Por outro lado, diversos estudos de perfil de expressão gênica tem sido realizados para auxiliar na compreensão da heterogeneidade dos TNBCs, caracterizando-os em diferentes subgrupos a partir de impressões digitais moleculares. A subclassificação mais abrangente, que é amplamente utilizada, foi proposta por Lehmann e colaboradores no ano de 2011. Eles foram os pioneiros na subtipagem dos TNBCs e, baseando-se em 21 conjuntos de dados de *microarray*, analisaram cerca de 857 TNBCs dividindo-os em seis subgrupos distintos: imunomodulador (IM), basal-like 1 (BL1), basal-like 2 (BL2), mesenquimal (M), mesenquimal semelhantes a células-tronco (MSL – *mesenchymal stem-like*) e receptor de androgênio luminal (LAR – *luminal androgen receptor*). Além disto, eles também analisaram e compararam o perfil de expressão gênica de diversas linhagens celulares de TNBC existentes, classificando-as nestes subtipos e fornecendo modelos celulares precisos, os quais podem auxiliar no tratamento destes tumores [Lehmann *et al.*, 2011].

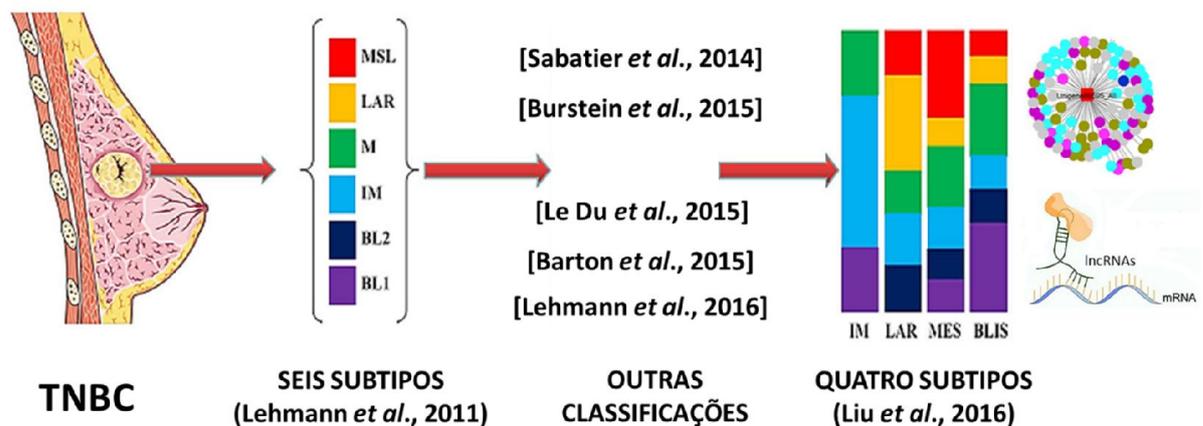
O grupo dos IM é caracterizado pelo aumento na expressão de genes envolvidos em processos celulares imunológicos e vias de transdução de sinais, tais como as vias de citocinas (interleucinas 12 e 7), de processamento de antígenos (sinalização dos receptores de linfócitos T e B), entre outros. As opções de tratamento para este subtipo envolve o uso de inibidores de *checkpoints* imunológicos como o PD1 e o PDL1 [Bertucci *et al.*, 2006]. O grupo BL1 apresenta elevada amplificação de genes relacionados a regulação do ciclo celular (*KRAS*, *CDK6*, *IGF1R*, *FGFR1*, *MYC* e *PIK3CA*) e alta frequência de deleção de genes relacionados ao reparo do DNA (*RB1*, *BRCA1/2*, *PTEN* e *TP53*), sendo utilizados como opção de tratamento inibidores da PARP1 e a quimioterapia baseada em compostos genotóxicos como a cisplatina. Por outro lado, o grupo BL2 é enriquecido para genes envolvidos na glicólise e na

gliconeogênese, além de exibir ativação anormal de vias de sinalização de fatores de crescimento (vias do EGFR, Wnt/ $\beta$ -catenin e IGF-1R). Inibidores de mTOR e de fatores de crescimento (lapatinib e cetuximab) podem ser utilizados como potenciais estratégias terapêuticas para este grupo [Lehmann *et al.*, 2014].

Os grupos M e MSL são caracterizados pelo enriquecimento da expressão de genes associados a vias de diferenciação (via Wnt e sinalização por TGF- $\beta$ ), de migração celular e de interação receptor-matriz extracelular. Entretanto, o grupo MSL apresenta um perfil de expressão gênica semelhante àquele de células-tronco mesenquimais, exibindo elevados níveis de expressão dos genes *HOX* e dos marcadores *BMP2*, *PDGFR*, *THY1* e *VCAM1*, e baixos níveis de expressão de genes relacionados à proliferação celular. O grupo M pode ser tratado com drogas que atuem na transição epitélio-mesênquima ou com inibidores de mTOR [Gibson *et al.*, 2005], enquanto as opções de tratamento para o grupo MSL envolvem o uso de drogas anti-angiogênicas e inibidores de PI3K. Por fim, o grupo LAR é enriquecido com genes envolvidos no metabolismo do androgênio/estrogênio, no metabolismo de porfirinas e na síntese de esteróides. Este grupo apresenta altos níveis de expressão do receptor de androgênio (RA) e a opção de tratamento envolve a terapia anti-RA [Lehmann *et al.*, 2014].

Com o passar dos anos novos estudos foram realizados e novas subtipagens passaram a ser propostas (Figura 3) [Sabatier *et al.*, 2014; Burstein *et al.*, 2015; Le Du *et al.*, 2015; Barton *et al.*, 2015; Lehmann *et al.*, 2016; Prado-Vázquez *et al.*, 2019]. Uma delas, proposta por Liu e colaboradores, no ano de 2016, baseou-se na integração de perfis transcriptômicos de RNAs mensageiros (mRNAs) e de RNAs longos não codificadores (lncRNAs), devido ao papel emergente destas moléculas na regulação de diversos processos celulares. A partir da análise de 165 amostras de

TNBC, foi identificada uma interação entre estas moléculas e uma nova subtipagem foi proposta, refinando os TNBC em apenas quatro subgrupos distintos: imunomoduladores (IM), basal-like/imunosuprimidos (BLIS – *basal-like and immune suppressed*), semelhante ao mesenquimal (MES – *mesenchymal-like*) e receptor de androgênio luminal (LAR - *luminal androgen receptor*). O grupo IM é enriquecido em vias que regulam as células do sistema imunológico, exibindo elevada expressão de genes associados com a resposta imune (inata e adquirida) e co-estimulação de linfócitos T. O grupo BLIS é enriquecido para vias que regulam o ciclo celular e o reparo do DNA, apresentando elevada expressão de genes associado à proliferação. Em contrapartida exibem baixa expressão de genes envolvidos em processos imunológicos (ativação do complemento, sinalização por linfócitos T e B, entre outros). O grupo MES é enriquecido para vias de sinalização, tais como a via do TGF- $\beta$  e de fatores de crescimento, e para vias que regulam a interação receptor-matriz extracelular e as junções comunicantes. Por fim, o grupo LAR é enriquecido em vias de regulação hormonal, tais como vias do metabolismo de androgênio e estrogênio e de biossíntese de esteróides e porfirinas [Liu *et al.*, 2016].



**Figura 3: Progresso da subtipagem molecular do câncer de mama triplo negativo (TNBC).** A primeira subtipagem proposta era baseada apenas no perfil de expressão gênica de tumores triplo negativos. Com o passar dos anos, interações entre moléculas (como, por exemplo, entre RNAs mensageiros e RNAs não codificadores de proteínas) passaram a ser analisadas [adaptado de Yin *et al.*, 2020].

Embora estas subtipagens utilizem um grande número de amostras de paciente portadores de TNBC e sejam importantes ferramentas que podem ser utilizadas no diagnóstico, prognóstico e na escolha de terapias direcionadas, elas baseiam-se principalmente na expressão de mRNA, o qual não reflete, com precisão, o nível de expressão proteica. No processo de tradução das proteínas existem diversas modificações e processos regulatórios que interferem na previsão do prognóstico e na escolha das estratégias terapêuticas direcionadas, fazendo com que apenas uma pequena parte se mostre útil em ensaios clínicos [Yin *et al.*, 2020]. Isto faz com que a quimioterapia citotóxica seja ainda a principal abordagem para tratamento dos TNBCs [Rusness *et al.*, 2017], além de justificar o aumento exponencial, nos últimos 10 anos, de estudos que buscam entender melhor a biologia deste subtipo tumoral. Em uma rápida busca no Pubmed utilizando o termo “*triple negative breast cancer*”, pode-se observar que, no ano de 2010, o número de estudos/publicações relacionados a este subtipo era de apenas 393, enquanto que no ano de 2020 este número chega a expressivos 2.860 [Pubmed, 2021].

Mesmo com o avanço dos estudos e das metodologias de classificação, a taxa de mortalidade observada nestes tumores permanece elevada, principalmente devido à sua agressividade (recidivas e metástases) associada à escassez de tratamentos direcionados eficazes. Este grupo de tumores tem sido um desafio para a clínica, portanto, buscar novos marcadores que auxiliem na identificação de pacientes com tumores potencialmente agressivos é essencial para diminuição da alta taxa de mortalidade observada neste grupo de tumores de mama [Rusness *et al.*, 2017].

#### 1.4. Biomarcadores no câncer de mama triplo negativo

Apresentando mecanismos de surgimento e evolução ainda não esclarecidos, o câncer de mama é considerado um dos mais graves problemas de Saúde Pública Mundial. Dentre os subtipos existentes, o câncer de mama triplo negativo se destaca por sua elevada taxa de mortalidade, devido à alta tendência de metástases e recidivas [WHO, 2021]. Quando diagnosticado em estágios iniciais, a possibilidade de cura é maior, mas devido ao acesso ainda limitado ao atendimento médico, para uma parcela significativa da população feminina, esse diagnóstico muitas vezes acontece tardiamente, impedindo o tratamento. Outro problema importante relacionado aos TNBCs refere-se à previsão correta do tipo de tratamento, uma vez que as atuais classificações moleculares destes tumores não são consideradas eficientes. Diante disso, buscar novos biomarcadores que possam estar relacionados a este subtipo e à sua progressão é de extrema importância, tanto na clínica médica, para o diagnóstico e prognóstico correto, quanto para a eleição de estratégias terapêuticas direcionadas [Phillips *et al.*, 1999].

O primeiro biomarcador tumoral foi descoberto no ano de 1847, antes mesmo da descoberta da molécula de DNA. Henry Bence Jones, químico e físico, detectou a presença de uma molécula específica na urina de pacientes com mieloma múltiplo, batizando-a de proteína de Bence Jones, a qual é utilizada até os dias de hoje e encontrada em 70% dos pacientes com esse tipo de câncer [Jones, 1851]. Biomarcadores são moléculas biológicas quantificáveis, incluindo proteínas (como, por exemplo, peptídeos, enzimas, receptores, anticorpos, marcadores de células-tronco, vias metabólicas, entre outros) ou ácidos nucleicos (como, por exemplo, genes e RNAs não codificadores), os quais são utilizados como assinaturas moleculares de um processo biológico normal, de um processo patogênico ou em resposta a uma

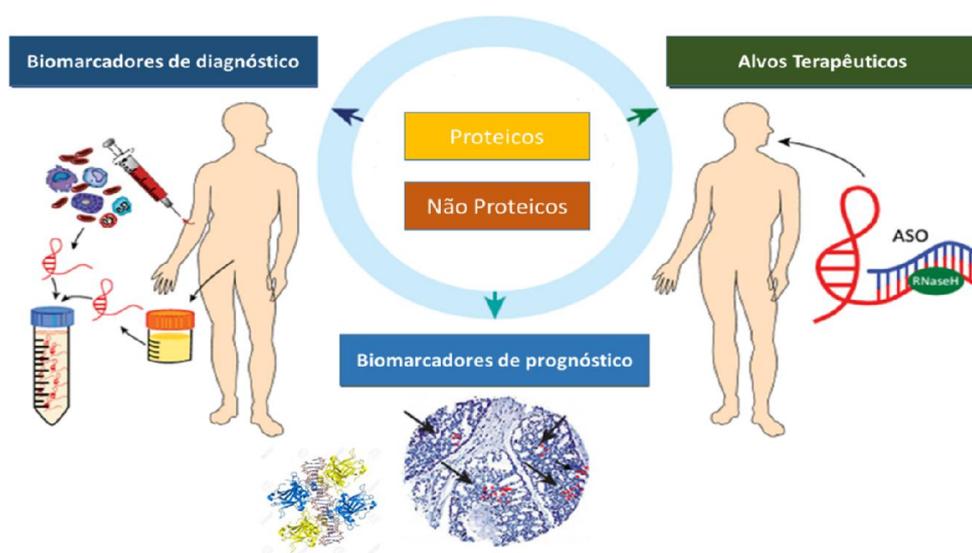
exposição ou intervenção, incluindo intervenções terapêuticas. Estes biomarcadores podem ser encontrados no sangue, nos fluidos corporais, em tecidos saudáveis, assim como, também, em tecidos tumorais [Pepe *et al.*, 2001]. Geralmente, são produzidos por indivíduos saudáveis, mas mudanças em sua quantidade, forma ou sequência gênica, devido à mutações, podem levar ao surgimento do câncer. Úteis no manejo clínico de pacientes com câncer, esses marcadores auxiliam no diagnóstico, prognóstico, detecção de recidiva, avaliação da resposta terapêutica, além de auxiliar no desenvolvimento de novas modalidades de tratamentos [Alonzo, 2005; De Almeida *et al.*, 2007].

Os biomarcadores podem ser classificados em três diferentes tipos: diagnóstico, prognóstico e preditivo. Biomarcadores de diagnóstico são utilizados para identificar indivíduos que tem uma doença específica, auxiliando na tomada de decisões sobre a escolha do tratamento, enquanto os de prognóstico servem para prever os resultados de sobrevivência dos pacientes, descrevendo a provável progressão da doença com ou sem intervenção terapêutica. Por fim, os biomarcadores preditivos são utilizados para prever a sensibilidade e a resposta dos pacientes a tratamentos específicos. Biomarcadores de diagnóstico e prognóstico também são considerados marcadores potencialmente preditivos e esta ligação continua a ser examinada em alguns estudos [Tanaka *et al.*, 2013; Barras & Savas, 2015]. Outra classificação refere-se ao seu local de detecção, podendo ser considerados séricos ou não-séricos. Os marcadores séricos são aqueles detectados na circulação (sangue, soro ou plasma), nas excreções e secreções (fezes, urina, expectoração ou secreção mamilar) sendo, portanto, facilmente medidos de forma não invasiva e em série. Os não séricos não são detectados na circulação e envolvem receptores celulares, fatores de transcrição, entre outros [Henry *et al.*, 2012].

A identificação de novos biomarcadores associados ao câncer pode ser realizada por meio de múltiplas abordagens, baseando-se na biologia tumoral, no ambiente circundante ou no metabolismo de agentes farmacêuticos específicos. Com a explosão de novos conhecimentos sobre a biologia tumoral e o advento de diversas técnicas de biologia molecular, como experimentos de ganho e perda de função através da superexpressão e nocauteamento de genes, sequenciamento em larga escala, análise de perfis de expressão gênica e espectrometria de massas, uma comparação entre células/tecidos saudáveis e tumorais pode ser feita e, portanto, marcadores que diferem entre as amostras podem ser identificados [Henry *et al.*, 2012]. Porém, para se pensar em um novo estudo de biomarcadores, é importante ter uma boa justificativa, visando atender uma necessidade clínica não atendida, além do mesmo necessitar ser preciso, reprodutível, simples e de baixo custo. O marcador ideal deve ser funcional para o diagnóstico precoce de neoplasias e suas origens, na determinação da extensão da doença, na detecção precoce de recidivas/metástases e no monitoramento da resposta terapêutica, além de ser órgão-sítio específico e ter meia-vida curta, permitindo acompanhar temporariamente as alterações no desenvolvimento tumoral [Reis, 2005; De Almeida *et al.*, 2007; Wang & Jiang, 2013].

Proteínas e RNAs não codificadores (ncRNAs) são reguladores finos da expressão gênica, tendo uma ampla variedade de funções que permitem sua interação com diversos processos que modulam o câncer, como: surgimento, desenvolvimento, progressão de metástases e recidivas. Diversos estudos demonstraram que essas moléculas atuam como possíveis oncogenes, sendo regulados positivamente no câncer de mama e facilitando o crescimento e a progressão tumoral, ou como possíveis supressores tumorais, sendo regulados negativamente no câncer de mama e inibindo a tumorigênese. Seu papel como

agentes terapêuticos está sendo explorado, tornando-os marcadores atraentes para o diagnóstico e prognóstico precisos de diversas doenças, incluindo o TNBC (Figura 4) (Yadav *et al.*, 2015; Cocco *et al.*, 2020; Da Silva *et al.*, 2020; Volovat *et al.*, 2020; Zhang *et al.*, 2021).



**Figura 4: A importância de proteínas e RNAs não codificadores na regulação da expressão gênica.** Proteínas e RNAs não codificadores podem ser utilizados como biomarcadores de diagnóstico, prognóstico ou como possíveis alvos terapêuticos [adaptado de Sahu *et al.*, 2015].

Recentemente, o marcador proteico de células-tronco CD14 foi descrito por nosso grupo de pesquisa, como sendo diferencialmente expresso em linhagens celulares humanas de câncer de mama, podendo estar relacionados com seu grau de malignidade [Lobba *et al.*, 2012], constituindo, portanto, interessante alvo que necessita ser melhor estudado do ponto de vista funcional, visando entender seu papel no câncer de mama humano. Até o momento ainda não existe informação disponível sobre o potencial papel do CD14 em TNBC. Por outro lado, nos últimos anos, tem aumentado significativamente o número de publicações que relacionam o papel de marcadores não proteicos, como RNAs não codificadores, no desenvolvimento e progressão do TNBC [Gibb *et al.*, 2011; Vikram *et al.*, 2014; Wang *et al.*, 2017].

No capítulo I, intitulado “Caracterização funcional (superexpressão) do gene codificador de proteína *CD14* em linhagens celulares humanas de câncer de mama triplo negativo”, eu descrevo nossos resultados de experimentos de ganho de função, via superexpressão gênica, do *CD14*, *in vitro*, utilizando linhagens celulares humanas de câncer de mama triplo negativo.

No capítulo II, intitulado “Identificação e caracterização funcional (superexpressão indutível e nocauteamento) do gene não codificador de proteína *LINC01133* em linhagens celulares humanas de câncer de mama triplo negativo”, eu descrevo nossos resultados referente à identificação de novos RNAs longos não codificadores (lncRNAs) diferencialmente expressos em um painel de linhagens celulares humanas de câncer de mama, além dos resultados de experimentos de ganho e perda de função (superexpressão e nocauteamento gênico), *in vitro*, de um dos lncRNAs escolhidos como alvo de estudo, o *LINC01133*, utilizando o mesmo painel de linhagens celulares.

## 2. CAPÍTULO I:

**Caracterização funcional (superexpressão) do gene codificador de proteína  
*CD14* em linhagens celulares humanas de câncer de mama triplo negativo**

## 2.1. INTRODUÇÃO

Dados recentes sugerem que um dos principais mecanismos envolvidos na progressão tumoral é a existência de subpopulações de células, as quais apresentam propriedades de células-tronco. Estas células são conhecidas como células-tronco tumorais ou células iniciadoras tumorais. Partindo desse pressuposto, a teoria mais recente de desenvolvimento do câncer propõe que a malignidade se origina no compartimento imortal de células-tronco e resulta da maturação de células-tronco mutantes ao invés da desdiferenciação de células somáticas [Cournoyer *et al.*, 2012; Vangipuram *et al.*, 2012]. Pesquisas promissoras envolvendo células-tronco do câncer de mama (BCSCs) vêm sendo desenvolvidas para estudar a progressão tumoral e descrever novos marcadores que possam ser utilizados no desenvolvimento de protocolos clínicos [Brown *et al.*, 2020; Du *et al.*, 2020; Elden *et al.*, 2021].

Um painel de marcadores moleculares vem sendo utilizado para identificar as BCSCs. Dentre eles, os mais utilizados são o CD44+/CD24- e o ALDH1+, os quais representam duas subpopulações distintas de BCSCs [Ginestier *et al.*, 2007]. Estudos baseados em perfis de expressão gênica mostraram que o fenótipo CD44+/CD24- é referente a um tipo mesenquimal e quiescente de BCSCs, enquanto que o fenótipo ALDH1+ é referente a um tipo epitelial e proliferativo [Mani *et al.*, 2008; Brabletz, 2012]. Outros marcadores destas células foram identificados em diferentes modelos de camundongos transgênicos. Um estudo, desenvolvido por Vassilopoulos e colaboradores, mostrou que camundongos BRCA1-mutantes com elevado potencial metastático eram enriquecidos para populações de BCSCs com o fenótipo CD24+/CD29+/CD49f+ [Vassilopoulos *et al.*, 2014]. Wright e colaboradores descreveram que tumores mamários de camundongos deficientes em BRCA1 exibiam subpopulações heterogêneas de BCSCs, incluindo células-tronco CD44+/CD24- e

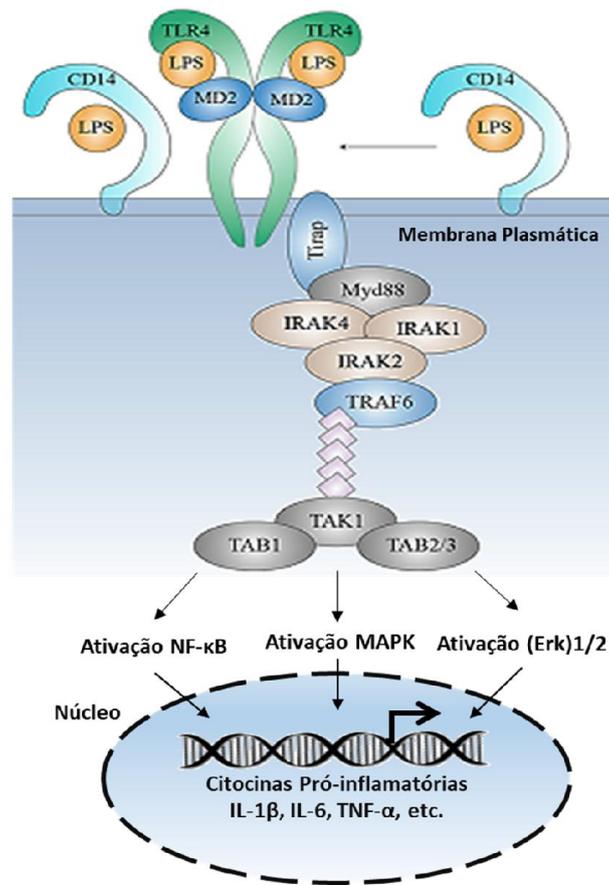
CD133+ [Wright *et al.*, 2008]. Grange e colaboradores relataram que camundongos BALB-neuT apresentaram uma subpopulação de BCSCs com o fenótipo Sca1+ [Grange *et al.*, 2008]. Marcadores também foram identificados em diferentes linhagens celulares de câncer de mama. Na linhagem MCF7 foram identificados o MUC1+ e o Lgr5+ [Engelmann *et al.*, 2008; Yang *et al.*, 2015]. Na linhagem HCC1937, o ABCG2+ [Lecia *et al.*, 2014]. Já na linhagem MDA-MB-231, os marcadores identificados foram o Procr+/ESA+, o GD2+, a Nectina-4+ e também o Lgr5+ [Hwang-Verslues *et al.*, 2009; Battula *et al.*, 2012; Yang *et al.*, 2015]. Por fim, na linhagem Hs578T, nosso grupo identificou o marcador CD90+ [Lobba *et al.*, 2018].

Atualmente, existe a necessidade de se identificar e caracterizar novos marcadores que estejam relacionados com tumores de mamas agressivos, como os TNBCs, complementando, de forma mais precisa, o diagnóstico, o protocolo clínico de prognóstico e a estratégia terapêutica a ser utilizada nas pacientes. Como citado anteriormente, o gene codificador de proteína *CD14* foi descrito por nosso grupo como sendo diferencialmente expresso em linhagens celulares humanas de câncer de mama, tornando-se, assim, um alvo promissor que deve ser melhor estudado e caracterizado [Lobba *et al.*, 2012].

O gene *CD14* consiste de 1.974 pb que codifica um único pré-mRNA de 1.400 pb, o qual sofre *splicing* alternativo podendo originar até quatro transcritos diferentes. Apesar disto, todos eles codificam a mesma proteína por apresentarem a mesma região codificadora. A proteína CD14 apresenta 375 aminoácidos, sendo composta por 11 repetições ricas em leucina. Sua estrutura cristalina revela que a mesma possui formato de fechadura, o que é importante para sua ligação a moléculas específicas [Kim *et al.*, 2005]. Ela é processada no retículo endoplasmático e se expressa como uma glicoproteína de 55 kDa na membrana celular via âncora de

glicosilfosfatidilinositol (GPI) (mCD14) [Wright *et al.*, 1990]. Da mesma forma que outras proteínas ancoradas ao GPI, a mCD14 acumula-se na superfície celular em microdomínios conhecidos como *lipid rafts*, que são ricos em colesterol e acumulam diversas quinases no seu sítio intracelular. Essa molécula é altamente expressa na superfície de células mielóides, tais como monócitos, macrófagos e neutrófilos e, com menor nível de expressão, na superfície de células hematopoiéticas, em células epiteliais, células endoteliais e fibroblastos [Anas *et al.*, 2010].

O CD14 também pode ser secretado como proteína solúvel (sCD14), sendo gerado por, pelo menos, três mecanismos distintos, os quais incluem o contorno da adição de glicosilfosfatidilinositol, a clivagem da âncora de GPI por fosfolipase D, ou a clivagem proteolítica direta a partir da superfície celular [Wright *et al.*, 1990]. Estudos relatam que a forma solúvel já foi encontrada no soro [(Ziegler-Heitbrock & Ulevitch, 1993; Klaassen *et al.*, 2013], no líquido, no leite materno [Snijders *et al.*, 2006] e em outros fluidos corporais, conferindo responsividade ao LPS para células que não expressam CD14 [He *et al.*, 2013; Lee *et al.*, 2015]. Primeiro receptor de reconhecimento padrão (PRR) identificado que se liga aos lipopolissacarídeos bacterianos (LPS), o antígeno de diferenciação de monócitos humanos CD14 tem a função de aumentar as respostas imunes inatas celulares. Tanto o mCD14 quanto o sCD14 são considerados co-receptores para vários receptores Toll-like (TLRs), transferindo produtos da parede celular bacteriana para estes receptores e, conseqüentemente, ativando cascatas de sinalização, tanto na superfície celular quanto em compartimento endossomais, que fazem o sistema imunológico inato reagir prontamente a infecções ou lesões que resultam em danos celulares (Figura 5) [Wright *et al.*, 1990; Granucci & Zanone, 2013].



**Figura 5: Ativação da via de sinalização dos TLRs.** A ligação do lipopolissacarídeos bacteriano (LPS) ao CD14 ativa a via de sinalização dos TLRs, a qual estimula a transcrição de citocinas pró-inflamatórias [adaptado de Wu *et al.*, 2019].

Estudos mostram que o CD14 é indispensável em processos de inflamação e infecção. Por exemplo, o CD14 promoveu significativamente a inflamação induzida por LPS em órgãos como o intestino [He *et al.*, 2016], fígado [Ogawa *et al.*, 2013], pulmão [Dagvadorj *et al.*, 2015], entre outros. O CD14 também pode atuar como um receptor para moléculas endógenas tais como a molécula de adesão intercelular 3 (ICAM-3) na superfície de células apoptóticas, peptídeos amilóides, proteínas de choque térmico (HSP70), ceramida e cristais de urato. A ligação de CD14 a estes ligantes, com exceção das células apoptóticas, medeia a ativação de respostas inflamatórias [Anas *et al.*, 2010]. O CD14 também é descrito por exibir papéis

importantes em doenças, tais como na aterosclerose, em doenças metabólicas e na regulação do câncer [Patino *et al.*, 2000; Waldo *et al.*, 2008; Kanczkowski *et al.*, 2010].

Apesar de ser classicamente descrito como um componente do sistema imune, existem alguns relatos da expressão de *CD14* em células do tecido mamário. Stein e colaboradores relataram a presença de *CD14*, associado ao *STAT3*, em células epiteliais luminais de camundongos associadas ao processo de involução da glândula mamária [Stein *et al.*, 2004]. Asselin-Labat e colaboradores identificaram o *CD14*, em conjunto com o *CD61*, como novos marcadores de células progenitoras luminais mamárias [Asselin-Labat *et al.*, 2007], as quais foram caracterizadas por Shehata e colaboradores, os quais mostraram a expressão do *CD14* em 86% das células progenitoras examinadas, confirmando, assim, sua relevância no tecido mamário [Shehata *et al.*, 2012]. A forma solúvel do *CD14* tem sido amplamente encontrada no leite materno, e a alteração dos seus níveis foi associada com uma série de doenças atópicas, tais como eczema e asma [Snijders *et al.*, 2006]. Foi relatada uma associação entre a redução dos níveis de *sCD14* no leite materno e o diagnóstico de eczema em crianças de até seis meses de idade [Jones *et al.*, 2002].

Na literatura, ainda é observada uma escassez de dados que relacionem o marcador *CD14*, seja na forma ancorada à membrana ou solúvel, com o câncer, mostrando a necessidade de mais estudos envolvendo esta proteína. Alguns trabalhos relatam que o nível sérico de *sCD14* foi maior em pacientes com câncer, tais como o glioma, do que em pacientes com doenças benignas ou pessoas normais, indicando seu possível papel no desenvolvimento do câncer [Wu *et al.*, 2010; Zhou *et al.*, 2010]. Recentemente, foi relatado que a alta expressão do *CD14* em uma subpopulação de células de câncer de bexiga está associada ao desenvolvimento do mesmo através do aumento da produção de citocinas pró-inflamatórias [Cheah *et al.*,

2015]. Em outro estudo, He e colaboradores mostraram a presença de sCD14 em pacientes com câncer de mama HER2+. Neste trabalho, foi observado que pacientes com menor nível de sCD14 no soro apresentaram um maior risco de recorrência da doença, quando comparados com aqueles que apresentaram um maior nível desta proteína no soro, indicando que o CD14 poderia ser um potencial biomarcador no prognóstico de recidiva neste subtipo de câncer de mama [He *et al.*, 2013]. Até o momento, são escassas as informações disponíveis na literatura sobre o potencial papel do *CD14* em TNBC. Lobba e colaboradores relataram uma alta expressão de *CD14* na linhagem celular não-tumorigênica de mama MCF10A (maior que 60%) e baixa expressão na linhagem de câncer de mama triplo negativa Hs578T (menor que 15%) [Lobba *et al.*, 2012]. Com base nestes dados, os autores sugerem que o CD14 poderia ser um marcador de bom prognóstico no carcinoma mamário, mas que para isso, é extremamente necessária a caracterização funcional desta molécula e uma investigação profunda do seu real papel, justificando o desenvolvimento do presente trabalho.

## 2.2. OBJETIVOS

### 2.2.1. Objetivo Geral

O objetivo geral deste capítulo é caracterizar funcionalmente o gene codificador de proteína *CD14*, em linhagens celulares humanas de câncer de mama triplo negativo, no intuito de esclarecer o papel de seu respectivo produto gênico na progressão do fenótipo tumoral, contribuindo, desta forma, para o futuro desenvolvimento de potenciais protocolos clínicos e terapêuticos relacionados ao câncer de mama triplo negativo.

### 2.2.2. Objetivos Específicos

- a) Ampliar o painel de linhagens celulares humanas de câncer de mama disponível em nosso laboratório, as quais representam os diferentes subtipos destes tumores e apresentam diferentes graus de malignidade, para avaliar a expressão do gene codificador de proteína *CD14* frente a essas linhagens;
- b) Analisar a função do gene codificador de proteína *CD14* através da superexpressão, via sistema lentiviral, na linhagem de câncer de mama triplo negativo Hs578T;
- c) Caracterizar as linhagens celulares geradas, *in vitro*, quanto à proliferação, ao crescimento de colônias dependente/independente de ancoragem e ao potencial migratório e invasivo;
- d) Analisar, nas linhagens geradas, a expressão de genes envolvidos na malignidade tumoral;

## 2.3. MATERIAIS E MÉTODOS

### 2.3.1. Linhagens Celulares e Condições de Cultivo

Para este trabalho, foram utilizadas sete linhagens celulares humanas de mama, as quais representam um modelo de progressão tumoral, como, também, os diferentes subtipos do carcinoma mamário. Todas as linhagens celulares foram obtidas da ATCC (*American Type Culture Collection*) e foram mantidas em meios de cultura específicos para cada linhagem. Abaixo são descritas as principais características de cada uma das linhagens celulares utilizadas neste trabalho:

- **MCF10A (ATCC: CRL 10317)**: linhagem epitelial não-tumorigênica derivada da doença fibrocística de mama;
- **MCF12A (ATCC: CRL 10782)**: linhagem epitelial não-tumorigênica derivada da doença fibrocística de mama e imortalizada espontaneamente;
- **MCF-7 (ATCC: HTB-22)**: linhagem epitelial derivada de carcinoma ductal invasivo de mama, a qual representa o subtipo molecular do carcinoma mamário Luminal A;
- **ZR-75-1 (ATCC: CRL-1500)**: linhagem epitelial derivada de adenocarcinoma ductal invasivo de mama, a qual representa o subtipo molecular do carcinoma mamário Luminal B;
- **SK-BR-3 (ATCC: HTB-30)**: linhagem epitelial derivada de adenocarcinoma metastático de mama, a qual representa o subtipo molecular do carcinoma mamário HER2;
- **MDA-MB-231 (ATCC: HTB-26)**: linhagem epitelial derivada de um adenocarcinoma metastático de mama, a qual representa o subtipo molecular do carcinoma mamário Triplo Negativo;

- **Hs578T (ATCC: HTB-126)**: linhagem epitelial derivada de carcinosarcoma mamário, a qual representa o subtipo molecular do carcinoma mamário triplo negativo;
- **293T**: linhagem derivada de células HEK293 oriundas de rim embrionário humano, as quais foram transformadas com o antígeno *large T* de SV-40, apresentando alta eficiência de transfecção.

As linhagens não-tumorigênicas de mama MCF10A e MCF12A foram cultivadas em meio DMEM/F12 (*Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12*) (Gibco) suplementado com 20ng/mL de EGF (Sigma-Aldrich), 10µ/mL de insulina (Sigma-Aldrich), 500ng/mL de hidrocortisona (Sigma-Aldrich), 10 µM de toxina colérica (Sigma-Aldrich) e 5% de soro fetal equino (SFE) (Gibco), conforme descrito por Soule e colaboradores [Soule *et al.*, 1990]. As linhagens MCF-7, ZR-75-1 e MDA-MB-231 foram cultivadas em meio RPMI (*Roswell Park Memorial Institute Medium*) (Gibco) suplementado com 10% de soro fetal bovino (SFB) (Cultilab/Vitrocell, Campinas, SP). Por fim, as linhagens SKBR-3 e Hs578T foram cultivadas em meio DMEM (*Dulbecco's Modified Eagle Medium*) (Gibco) suplementado com 10% de SFB. Todas as linhagens celulares foram cultivadas na presença dos antibióticos ampicilina (25mg/L) e estreptomicina (100mg/L) (Sigma-Aldrich) à temperatura de 37°C em atmosfera úmida de 2% ou 5% de CO<sub>2</sub>, para a manutenção do pH próximo ao fisiológico. Ao atingirem aproximadamente 80% da densidade de saturação, as culturas celulares foram lavadas com PBSA (*Phosphate Buffered Saline* – sem íons de cálcio e magnésio) e subcultivadas utilizando-se solução de tripsina. Os estoques celulares foram mantidos em meio de cultivo contendo 10% de DMSO (dimetilsulfóxido estéril) (Sigma-Aldrich) e armazenados em nitrogênio líquido a -190°C.

### 2.3.2. Ensaio de PCR quantitativo em tempo real (qRT-PCR)

#### 2.3.2.1. Extração do RNA total

A extração do RNA total das linhagens celulares foi realizada a partir de aproximadamente  $1 \times 10^6$  células, utilizando-se os reagentes do RNeasy Mini Kit (QIAGEN), conforme as especificações do fabricante, exceto que o tratamento com DNaseI foi realizado por 1h. As amostras de RNA foram quantificadas através de sua absorbância a 260nm e sua pureza foi avaliada através da razão 260/280nm em espectrofotômetro (*NanoDrop™ 1000 Spectrophotometer*, Thermo Scientific, Wilmington, USA), tendo-se, como pureza satisfatória, uma relação próxima de 2,0. O RNA purificado foi armazenado a  $-80^\circ\text{C}$ .

#### 2.3.2.2. Desenho dos *primers*

Os primers utilizados para amplificação dos genes nos experimentos de qRT-PCR foram desenhados com o auxílio do programa computacional *Primers Express 3.0* (Thermo Scientific) (Tabela 2). Para se obter um bom desempenho nas reações de qRT-PCR, eles foram desenhados com as seguintes características: amplificação de fragmentos de aproximadamente 150pb, conteúdo de CG entre 30 e 80%, incapacidade de formação de dímeros ou de estruturas secundárias e apresentar temperatura de anelamento entre  $58^\circ\text{C}$  e  $60^\circ\text{C}$ . Além disso, a fim de evitar uma eventual coamplificação de DNA genômico contaminante, a maior parte dos pares de primers foram desenhados em éxons diferentes do gene.

**Tabela 2: Sequência de *primers* utilizados nos ensaios de qRT-PCR.**

<b>Genes</b>	<b>Sequência dos <i>Primers</i> (5' → 3')</b>
<b><i>CD14</i></b>	F: CCGCTGGTGCACGTCTCT R: AATCTTCATCGTCCAGCTCACA
<b><i>GAPDH</i></b>	F: TGCACCACCAACTGCTTAGC R: GGCATGGACTGTGGTCATGAG

<b>HPRT1</b>	F: TGAGGATTTGGAAAAGGGTCT R: GAGCACACAGAGGGCTACAA
<b>HMBS</b>	F: TGGACCTGGTTGTTCACTCCTT R: CAACAGCATCATGAGGGTTTT

### 2.3.2.3. Transcrição reversa e síntese de cDNA

Para a transcrição reversa e síntese de cDNA fita simples foi utilizado 1 µg de RNA total de cada condição experimental. Cada amostra foi transcrita utilizando-se a enzima *SuperScript III* (Invitrogen), de acordo com as recomendações do fabricante. Os cDNAs sintetizados foram então utilizados como *template* para as reações de qRT-PCR.

Para a quantificação do produto formado durante a reação de amplificação, foi utilizado o reagente *SYBR®GreenDye* (Applied Biosystems). As reações de qRT-PCR foram realizadas em triplicatas em um volume final de 6µl no Termociclador ViiA7 Real-Time PCR System (Applied Biosystems). Como template, foram utilizados 1,5µl do cDNA sintetizado e diluído 30x em água Milli-Q®. Para cada reação, foram utilizados 1,5µl do conjunto de primers na concentração final previamente determinada e 3µl do reagente *SYBR®GreenDye*. As reações foram realizadas nas seguintes condições: 50°C por 2 min (etapa de ativação da enzima Uracil N-Glicosilase, AmpEraser), 95°C por 10 min (etapa de ativação da enzima DNA Polimerase, Taq Gold), 40 ciclos de 95°C de 15 seg (etapa de desnaturação) e 60°C por 1 min (etapa de anelamento dos oligonucleotídeos e extensão do amplicon). Para avaliar a especificidade das reações de amplificação e detectar a presença de eventual contaminação, uma reação de dissociação do produto amplificado foi realizada ao término de cada reação. Além disso, foram realizadas reações controles de amplificação na ausência de molde (cDNA) denominadas NTC (*Non-Template Control*).

Considerando a possibilidade de variação da concentração inicial de cDNA na análise dos dados provenientes de duas ou mais amostras, foi realizada a normalização na qual a expressão do gene-alvo foi determinada em função da expressão de genes referências. Assim, a quantificação das amostras foi realizada utilizando-se *GAPDH*, *HPRT1* e *HMBS* como referências.

#### 2.3.2.4. Determinação da concentração final de *primers*

A concentração final adequada de cada par de primers, específicos para os genes analisados, foi determinada através de reações variando-se a concentração final destes primers entre a mínima de 200nM e máxima de 600nM. Como *template*, utilizou-se uma mistura de cDNAs provenientes de diferentes linhagens celulares. Desta forma, foi possível determinar a menor concentração final de primers que resultou na menor formação de dímeros, ou inexistente, sem que houvesse variação no valor do Ct e no perfil da curva de ampliação do gene em relação às maiores concentrações analisadas.

#### 2.3.2.5. Determinação da eficiência dos *primers*

Para calcular a eficiência dos *primers*, foram realizadas reações de qRT-PCR utilizando-se, como *template*, uma mistura de cDNAs provenientes de diferentes linhagens de mama humana. Através da análise de regressão linear dos valores de Cts em função do logaritmo da respectiva diluição, foi possível obter o coeficiente angular da reta (a, em  $y = a.x + b$ ). Este coeficiente foi, então, utilizado para cálculo da eficiência de amplificação do produto pelos *primers*, na seguinte fórmula:

$$E_f = 10^{-1/\text{coeficiente angular}}$$

$$E_f (\%) = (E_f - 1) \times 100$$

### 2.3.2.6. Confirmação da expressão diferencial

A média dos Cts das triplicatas foi calculada inicialmente. Dado que a expressão do gene é analisada em relação a uma amostra que é tomada como referência, calculou-se, então, a diferença entre a média dos Cts da amostra referência e a média dos Cts da amostra estudada. Essa diferença foi definida como  $\Delta C_p$ . O Cálculo do  $\Delta C_p$  foi realizado para os dados do gene alvo e para os dados do gene de expressão constitutiva. A fórmula final para o cálculo da diferença de expressão dos genes entre as amostras analisadas, considerando-se que pode não ocorrer um ganho de duas vezes do produto amplificado a cada ciclo, dado que a eficiência de amplificação dos *primers* utilizados não é de 100%, sendo dada por:

$$\text{ratio} = \frac{E_{f(\text{gene alvo})}^{\Delta C_p(\text{gene alvo})}}{\text{Fator de Normalização}}$$

O Fator de Normalização foi calculado através da análise da expressão dos genes *GAPDH*, *HPRT1* e *HMBS* no programa computacional GeNorm [Vandesompele *et al.*, 2002]. De acordo com esta análise, foram classificados os dois genes mais estáveis, *HPRT1* e *GAPDH*, que foram, então, utilizados para calcular o controle referência utilizado como Fator de Normalização no cálculo da expressão diferencial.

### 2.3.3. Construção do vetor lentiviral pLV-eGFP-CD14 para ensaio de superexpressão

Para a superexpressão do gene CD14 humano, utilizamos o sistema lentiviral de expressão em células de mamífero, o qual é baseado no plasmídeo pLV-eGFP, que expressa o gene repórter *eGFP* (*Enhanced Green Fluorescent Protein*) [Tiscornia *et al.*, 2006]. Os vetores lentivirais foram construídos com base na sequência completa do cDNA correspondente ao *CD14*, o qual foi clonado nesse sistema de maneira sítio-dirigida para que apenas a fita de interesse seja expressa. A superexpressão do *CD14*

foi realizada em células Hs578T, linhagem na qual esse gene apresenta baixa expressão.

#### 2.3.3.1. Desenho dos *primers* para amplificação da região codificadora do *CD14*

Os *primers* para amplificação foram desenhados utilizando-se a sequência inicial e final do cDNA correspondente à sequência do cDNA de *CD14* humano (Tabela 3). Na extremidade 5' dos *primers forward* foi adicionado um sítio de restrição para a enzima *XbaI* (TCTAGA), enquanto na extremidade 3' dos *primers reverse* foi adicionado um sítio de restrição para a enzima *MluI* (ACGCGT), visando a inserção do gene de interesse nos respectivos sítios *XbaI* e *MluI* do vetor lentiviral.

**Tabela 3: *Primers* de Clonagem do *CD14***

<b>Genes</b>	<b>Sequência dos <i>Primers</i> (5'→ 3')</b>	<b>Temp. de Melting (°C)</b>
<b>hCD14-F-<i>XbaI</i></b>	TCTAGAATGGAGCGCGCTCCTGCTT	73°C
<b>hCD14-R-<i>MluI</i></b>	ACGCGTTTAGGCAAAGCCCCGG	72°C

#### 2.3.3.2. Amplificação da região codificadora do gene *CD14*

Cerca de 1,0µg de cDNA da linhagem celular não-tumorigênica de mama MCF10A foi utilizada como molde para amplificação da região codificadora do gene *CD14* humano. Para a reação de amplificação, foram utilizados os seguintes reagentes: 1X *High Fidelity Buffer* (Finnzymes); 2mM de dNTP (Fermentas); 1U de *Phusion® Hot Start High Fidelity DNA Polymerase* (Finnzymes); *primers* específicos na concentração 0,5µM (*primer forward* e *primer reverse*). As condições de ciclagem no termociclador *PTC- 225 Pelthier Thermal Cycler®* (MJ Research) foram: 98°C por 2 min; 40 ciclos de 98°C por 10 seg, 72°C por 30 seg, 72°C por 2 min e 72°C por 5 min para extensão final dos fragmentos. O produto de PCR foi analisado em gel de agarose 0,8% corado com Brometo de Etídeo (0,5µg/mL) e visualizado sob luz ultravioleta (UV) no fotodocumentador ImageQuant LAS 4000 (GE Healthcare).

#### 2.3.3.3. Clonagem da região codificadora do gene *CD14*

Os fragmentos correspondentes à região codificadora do gene *CD14* humano, amplificados por PCR, foram fracionados em gel de agarose para, então, serem cortados e purificados a partir do gel, utilizando-se o kit *QIAquick Gel Extraction Kit* (QIAGEN), de acordo com as recomendações do fabricante. Após a extração, o DNA foi quantificado no espectrofotômetro *Nanodrop ND-1000 Spectrophotometer*.

O fragmento de interesse foi subclonado no vetor *pGEM®-T Easy* (Promega), o qual é um sistema otimizado para a clonagem de produtos de PCR, sendo, por este motivo, utilizado como primeira abordagem para a clonagem dos produtos de amplificação do *CD14*, visando seu sequenciamento e posterior subclonagem nos vetores de expressão utilizados neste trabalho. Inicialmente o fragmento foi tratado com dATP para inserção de uma cauda poliA, a qual é importante para a ligação do mesmo no vetor *pGEM®-T Easy*. Para tanto, a seguinte reação foi realizada: 7 µL do fragmento de interesse purificado, 1x *Taq Polymerase Reaction Buffer* (Thermo Scientific), 2mM de Cloreto de Magnésio ( $MgCl_2$ ), 0,2mM de dATP (Thermo Scientific) e 5U da enzima *Taq Polymerase* (Thermo Scientific). A reação foi então incubada à 70°C por 30 min. Os fragmentos tratados com dATP foram então subclonados no vetor *pGEM®-T Easy*, utilizando-se a enzima T4-DNA ligase (kit *pGEM®-T Easy Vector System*). A reação foi incubada à 16°C por 18h, seguida da inativação enzimática à 80°C por 20 min. Alíquotas da reação de ligação foram utilizadas para transformação de bactérias *E. coli* XL1-Blue por eletroporação (eletroporador EC-100, EC Apparatus Corporation).

#### 2.3.3.4. Transformação por eletroporação de bactérias *Escherichia coli* eletrocompetentes com DNA plasmideal

Bactérias *E. coli* XL1-Blue eletrocompetentes foram transformadas por

eletroporação com as construções plasmideais de interesse como descrito a seguir: Uma alíquota de 3µL da reação de ligação foi adicionada a 100µL de bactérias. Esta mistura foi transferida para uma cubeta de eletroporação (*gap* entre os eletrodos de 2mm) (Bio-Rad). O eletroporador (EC-100, EC Apparatus Corporation) foi ajustado conforme o manual do aparelho para 2.500V, sendo então aplicado o pulso à cubeta. Em seguida, foi adicionado 900 µL de meio LB à suspensão bacteriana e esta foi incubada à 37°C sob agitação de 200-250rpm por 1h. As bactérias transformadas foram plaqueadas em meio LB-ágar 1,5% contendo 100µg/mL de ampicilina para obtenção de colônias resistentes ao antibiótico. As placas foram, então, incubadas à 37°C por 16h. Os clones bacterianos obtidos foram expandidos em meio LB líquido contendo 100µg/mL de ampicilina e incubados à 37°C por 16-18 horas sob agitação de 200-250rpm. Estoques celulares foram estabelecidos para manutenção em -80°C. Para isto, adicionou-se à suspensão celular uma solução de glicerol 30% na proporção de 1:1 (v:v).

#### 2.3.3.5. Preparação de DNA plasmideal em pequena escala

Inóculos de 5mL foram utilizados para preparações plasmideais em pequena escala. Para estas preparações foi utilizado o kit de purificação de DNA *Illustra PlasmidPrep Mini Spin Kit* (GE Healthcare), seguindo as recomendações do fabricante. As preparações de DNA resultantes foram quantificadas em espectrofotômetro (*Nanodrop Spectrophotometer ND-100*, Thermo Scientific, Wilmington, USA). A sequência íntegra do cDNA do gene *CD14* foi extraída do vetor *pGEM®-T Easy* através da digestão com as enzimas de restrição *XbaI* e *MluI* (Fermentas), sendo este fragmento, posteriormente, purificado do gel.

#### 2.3.3.6. Análise da sequência de nucleotídeos

A integridade da sequência do cDNA do gene *CD14* foi avaliada através de sequenciamento de DNA no aparelho *ABI Prism 3700 DNA Analyser* (Applied Biosystems) utilizando-se o Kit *BigDye® Terminator v3.1 Cycle Sequencing* (Applied Biosystems), conforme especificações do fabricante. As sequências obtidas no sequenciamento foram analisadas e alinhadas utilizando-se os programas *PhredPhrap* [Ewing, 1998; Ewing & Green., 1998] e *Blast* (NCBI) [<https://blast.ncbi.nlm.nih.gov/Blast.cgi>].

#### 2.3.3.7. Sub-clonagem da construção no vetor lentiviral de terceira geração

Uma das construções contendo a sequência completa de nucleotídeos do gene *CD14* foi utilizada para a subclonagem nos respectivos sítios de restrição do vetor lentiviral pLV-eGFP e esta construção foi então utilizada no ensaio funcional de superexpressão.

### 2.3.4. Geração da linhagem Hs578T superexpressando o gene *CD14*

#### 2.3.4.1. Produção de partículas lentivirais em células 293T

As partículas lentivirais recombinantes de superexpressão foram geradas por meio de transfecção da linhagem empacotadora de vírus 293T. Foi utilizado o vetor pLV-eGFP/CD14 para produção de partículas virais para o ensaio de superexpressão. Para cada construção de interesse foi utilizada uma placa de 6 poços (35mm de diâmetro) de células 293T em alta densidade de saturação (80-90%). A proporção dos DNAs plasmídeos para a produção dos lentivírus foi calculada de acordo com o protocolo descrito por Tiscornia e colaboradores [Tiscornia *et al.*, 2006]. Sendo assim, foi utilizado 1,1µg do vetor de transferência, 720ng do vetor pMDL, 280ng do vetor pREV e 390ng do vetor pVSVG. Estes plasmídeos foram diluídos em 150µL de meio

DMEM sem soro. Em outro tubo de microcentrífuga, 10µL do reagente de lipofecção (*Lipofectamine 2000* - Life Technologies) foram diluídos em 150µL de meio DMEM sem soro para cada ensaio de transfecção. Os conteúdos dos dois tubos foram misturados e incubados por 20 min à temperatura ambiente. Após este período, a mistura de lipossomos/DNA foi adicionada por gotejamento ao meio de cultura de células. Após 4-6h, foi efetuada a troca do meio para início das coletas de partículas virais em 24h. O sobrenadante de cada placa foi coletado no tempo de 24, 48 e 72hs após a transfecção. Para tanto, foi adicionado 1mL de DMEM suplementado com 10% SFB em cada poço. O sobrenadante de cada coleta foi centrifugado a 20.000xg por 10 min à temperatura ambiente para eliminar restos celulares e, em seguida, aliquoteados e congelados à -80°C.

#### 2.3.4.2. Titulação das partículas lentivirais eGFP positivas (eGFP<sup>+</sup>)

As partículas virais produzidas foram tituladas em células da linhagem 293T conforme descrito por Tiscornia e colaboradores [Tiscornia *et al.*, 2006]. Foram preparadas diluições  $10^{-1}$  e  $10^{-2}$  dos meios condicionados contendo as partículas virais. Para realização da transdução viral, misturou-se 20µL dos meios condicionados originais ou das diluições e 480µl de uma suspensão celular contendo  $1 \times 10^5$  células 293T em DMEM suplementado com 10% SFB. Imediatamente após a mistura, as células foram transferidas para poços de uma placa de 24 poços e incubadas por 48h, até o momento da contagem das células fluorescentes. Para contagem, estas células foram dissociadas e a porcentagem de células eGFP positivas foi analisada por citometria de fluxo.

O título de cada preparação viral (UFC: Unidades Formadoras de Colônias) foi obtido a partir da contagem da proporção de células fluorescentes, utilizando-se a seguinte relação:

Título (UFC) = (P x N / 100 x V) x /DF, onde:

P = % de células EGFP+

N = número de células no momento da transdução ( $10^5$ )

V = volume da diluição usado para a transdução

DF = fator de diluição

#### 2.3.4.3. Transdução viral da linhagem Hs578T com partículas lentivirais

As células da linhagem Hs578T foram transduzidas com a preparação lentiviral correspondente ao pLV-eGFP/CD14 e ao controle pLV-eGFP. Para a transdução, cada construção lentiviral foi misturada com uma suspensão de  $1 \times 10^5$  células para se obter uma multiplicidade de infecção de 10 (MOI: *Multiplicity of Infection*). As células foram, então, plaqueadas em placas de 24 poços e o meio de cultura foi trocado após 48h da transdução. As células eGFP+ foram selecionadas através da separação celular no citômetro de fluxo (*cell sorting*) e subcultivadas até atingir um número de células suficiente para congelamento em nitrogênio líquido.

#### 2.3.5. Citometria de Fluxo

Experimentos de citometria de fluxo foram realizados no citômetro FACSAriaII (BD, Biosciences). Os resultados foram analisados e processados no programa computacional FlowJo (Tree Star INC., USA) versão 7.6. Para o *cell sorting*, as células transduzidas com os vírus pLV-eGFP e pLV-eGFP/CD14 e eGFP+ foram sorteadas e coletadas em ambiente asséptico, mantidas em cultura para posterior caracterização molecular das linhagens geradas e para a realização de ensaios fenotípicos.

### 2.3.6. Confirmação da superexpressão do *CD14* nas linhagens transduzidas

Para avaliar o nível de mRNA correspondente ao gene *CD14* nas linhagens transduzidas, foram realizadas reações de qRT-PCR. A expressão proteica do CD14 foi avaliada por *Western Blot*. Os extratos proteicos totais foram obtidos incubando-se as células em tampão de lise RIPA (150mM NaCl, 1% NP-40, 0,5% SDS, 50mM Tris pH 8.2, 1mM EDTA) contendo inibidor de protease. A concentração das proteínas nos extratos foi determinada pelo método de Bradford (Bio-Rad Protein Assay, Bio-Rad, Hercules, CA, EUA), utilizando-se uma curva padrão de BSA (albumina sérica bovina). Alíquotas de 30ug dos extratos proteicos foram fracionados por eletroforese em gel SDS-PAGE e transferidos para membrana de nitrocelulose. Posteriormente, as membranas foram incubadas com o anticorpo primário anti-CD14 (Ab 45870, 1/500) (Abcam, Cambridge, UK). Depois da lavagem e incubação com o anticorpo secundário HRP conjugado (1/5000) (Sigma-Aldrich), as bandas proteicas foram detectadas no fotodocumentador ImageQuant LAS 4000 (GE Healthcare). Como controle endógeno foi utilizado o anticorpo anti-GAPDH (G8795, 1/10000) (Sigma-Aldrich).

### 2.3.7. Caracterização das linhagens celulares geradas

#### 2.3.7.1. Curva de crescimento celular

A taxa de proliferação da linhagem parental e das linhagens transduzidas foi determinado através do ensaio de curva de crescimento celular. Foram plaqueadas  $1 \times 10^4$  células/mL em cada poço de placas de 12 poços (22mm de diâmetro). As placas foram mantidas em condições adequadas de cultura e o meio de cultivo foi renovado a cada dois dias. Realizaram-se diferentes coletas para cada linhagem celular analisada, realizadas após 1, 3, 5, 7 e 9 dias do plaqueamento, em triplicata. A coleta consistiu na lavagem das placas com PBSA e posterior incubação com 0,2mL de

solução de tripsina. Depois de constatado que as células se encontravam em suspensão, adicionou-se 0,25mL de PBSA. Em seguida, foram adicionados 0,05mL de formaldeído 37% para fixação das células e posterior contagem. O número de células foi determinado com auxílio do hemocitômetro/citômetro Accuri (BD Bioscience).

#### 2.3.7.2. Ensaios de tumorigênese *in vitro*

O potencial tumorigênico, medido através do crescimento de colônias, da linhagem parental e das linhagens transduzidas foi analisado utilizando-se diferentes abordagens *in vitro*:

##### 2.3.7.2.1. Ensaio de crescimento de colônias em substrato sólido

O potencial tumorigênico da linhagem parental e das linhagens transduzidas foi inicialmente determinado através do ensaio de crescimento de colônias em substrato sólido. Este ensaio tem a finalidade de avaliar a capacidade das células de formarem colônias dependente de ancoragem, além de avaliar a viabilidade e a proliferação celular [Puck *et al.*, 1956]. Foram plaqueadas  $3 \times 10^2$  células/mL em cada poço de placas de 6 poços. As placas foram mantidas em condições adequadas de cultura e o meio de cultivo foi renovado sempre que necessário. As placas foram mantidas em cultura por 12 dias, tempo necessário para que fosse observada a formação de colônias. Ao término da segunda semana, o meio de cultivo foi removido e foram adicionados 0,5mL de cristal violeta 0,5% para fixação e coloração das colônias. O excesso de cristal violeta foi removido e as placas foram lavadas com água corrente cuidadosamente. Posteriormente as colônias foram quantificadas manualmente.

#### 2.3.7.2.2. Ensaio de crescimento de colônias em suspensão de agarose

O potencial tumorigênico da linhagem parental e das linhagens transduzidas foi também determinado indiretamente através do ensaio de crescimento de colônias em suspensão de agarose. Este ensaio tem a finalidade de avaliar a capacidade das células de formarem colônias independente de ancoragem, a qual é uma marca registrada do processo de tumorigênese [Borowicz et al., 2014]. Cada poço de placas de 24 poços (16mm de diâmetro) foi preenchido com 450µL de solução semi-sólida de agarose 0,6% e meio de cultura específico para a linhagem. Sobre esta camada semi-sólida, foram plaqueados 100µL de meio de cultura contendo células isoladas diluídas na concentração de  $10^5$  células/mL em meio de cultura, em triplicata. Rapidamente, para evitar a adesão das células, foram adicionados 500µL de solução semi-sólida de agarose 0,3% em meio de cultura específico para a linhagem. Após 21 dias do plaqueamento, as colônias foram fotografadas e quantificadas, no aumento de 40x, utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* (Life Technologies) e, posteriormente, fixadas adicionando-se 100µL de paraformaldeído 4%. Foram consideradas colônias apenas aquelas que apresentaram diâmetro maior do que 250µm.

#### 2.3.7.3. Ensaio de migração celular *in vitro*

O potencial migratório da linhagem parental e das linhagens transduzidas foi analisado utilizando-se diferentes abordagens *in vitro*:

##### 2.3.7.3.1. Ensaio de migração celular por *scratching* (*wound healing*)

O potencial migratório da linhagem parental e das linhagens transduzidas foi inicialmente determinado através do ensaio de migração celular por *scratching*. Foram plaqueadas  $2 \times 10^5$  células/mL em cada poço de placas de 24 poços, em triplicata. No

dia seguinte, após atingirem a confluência, o meio de cultura foi aspirado e um risco contínuo (*scratching*) foi feito na placa com uma ponteira de 200 $\mu$ L. O poço foi então lavado uma vez com PBSA para a remoção dos detritos celulares. Imagens do mesmo campo de cada poço foram obtidas do microscópio *EVOS FL Fluorescence Cell Image System* (Life Technologies) em aumento de 100x nos tempos de 0, 6 e 12 horas após o *scratching*. As análises foram realizadas medindo-se a área do risco em unidades arbitrárias (u.a) nas imagens utilizando-se o software ImageJ (NIH, MD, EUA) para cálculo da taxa de migração.

#### 2.3.7.3.2. Ensaio de migração celular por *Transwell*<sup>TM</sup>

O potencial migratório da linhagem parental e das linhagens transduzidas foi também determinado através do ensaio de migração celular em placas do tipo *Transwell*<sup>TM</sup> (BD Bioscience). Foram plaqueadas  $50 \times 10^3$  células na superfície superior da membrana porosa do inserto (tamanho do poro: 8 $\mu$ m), em 300 $\mu$ L de meio DMEM 0% SFB, e no poço abaixo do inserto, adicionou-se 700  $\mu$ L de meio DMEM 10% SFB, o qual atua como quimioatraente para as células. As células foram mantidas por 16h-18h a 37°C em atmosfera de 2% de CO<sub>2</sub>, seguidas por fixação em metanol gelado por 20 min e coradas com hematoxilina de Mayer por 20 min. As células que não migraram neste período foram removidas com o auxílio de hastes flexíveis de algodão. Após a secagem, as células aderidas a outra face do inserto foram contadas em microscópio *EVOS FL Fluorescence Cell Image System* (Life Technologies) em aumento de 200x. Quatros campos representativos de cada um dos insertos foram considerados, com duplicatas para cada tratamento.

#### 2.3.7.4. Ensaio de invasão celular *in vitro*

A capacidade invasiva da linhagem parental e das linhagens transduzidas foi também determinada através de ensaios de invasão *in vitro* em placas do *Transwell*<sup>TM</sup> (BD Bioscience), utilizando-se membrana basal reconstituível (*Matrigel*<sup>TM</sup>) como substrato para invasão. Foram plaqueadas  $100 \times 10^3$  células na superfície superior da membrana porosa do inserto (tamanho do poro: 8 $\mu$ m), em 300 $\mu$ L de meio DMEM 0% SFB, e no poço abaixo do inserto, adicionou-se 700  $\mu$ L de meio DMEM 10% SFB, o qual atua como quimioatraente para as células. As células foram mantidas por 16h-18h a 37°C em atmosfera de 2% de CO<sup>2</sup>, seguidas por fixação em metanol gelado por 20min e coradas com hematoxilina de Mayer por 20min. As células que não migraram neste período foram removidas com o auxílio de hastes flexíveis de algodão. As células capazes de degradar o *Matrigel*<sup>TM</sup> e migrar para a superfície oposta da membrana porosa foram então fixadas, coradas e contadas em microscópio *EVOS FL Fluorescence Cell Imagem System* (Life Technologies) em aumento de 200x. Quatro campos representativos de cada um dos insertos foram considerados, com duplicatas para cada tratamento.

#### 2.3.7.5. Análise da expressão de genes envolvidos na malignidade tumoral

Para entender melhor os possíveis efeitos da superexpressão do *CD14* nos processos de angiogênese, migração e invasão celular, foi investigada a expressão de diferentes genes relacionados a malignidade tumoral por qRT-PCR, conforme descrito no item 2.3.2. Os seguintes genes foram utilizados na análise: 1) Gene *XIAP* (anti-apoptótico); 2) Genes *TGF- $\beta$* , *SPARC*, *CXCL12*, *SATB1*, *E-CAD*, *N-CAD* (associados aos processos de tumorigênese e transição epitélio-mesenquimal); 3) Genes *VEGF-A* e *COL3A1* (associados à angiogênese tumoral).

**Tabela 4: Sequência de *primers* utilizados na análise da expressão de genes envolvidos na malignidade tumoral.**

<b>Gene</b>	<b>Sequência dos <i>Primers</i> (5' → 3')</b>
<b><i>XIAP</i></b>	F: GCCAGACTATGCCCATTTA R: CCACCACAACAAAAGCATTG
<b><i>TGF-β</i></b>	F: GGGACTATCCACCTGCAAGA R: CCTCCTTGGCGTAGTAGTCG
<b><i>SPARC</i></b>	F: GCAGCCCCTCAGCAAGAAG R: CAGTCACCTCTGCCACAGTTTC
<b><i>SATB1</i></b>	F: TGCAAAGGTTGCAGCAACCAAAGC R: AACATGGATAATGTGGGGCGGCCT
<b><i>CXCL12</i></b>	F: TCAGCCTGAGCTACAGATGC R: CTTTAGCTTCGGGTCAATGC
<b><i>E-CAD</i></b>	F: CCACCACGGGCTTGGATT R: GGAGGTGGTGAGAGAGACCT
<b><i>N-CAD</i></b>	F: GGGAAATGGAACTTGATGGCA R: CAGTTGCTAACTTCACTGAAAGGA
<b><i>VEGF-A</i></b>	F: GAGGAGCAGTTACGGTCTGTG R: TCCTTTCCTTAGCTGACACTTGT
<b><i>COL3A1</i></b>	F: TGATGGGATCCAATGAGGGAGA R: GAGTCTCATGGCCTTGCGTGTTT

#### 2.3.8. Análises Estatísticas

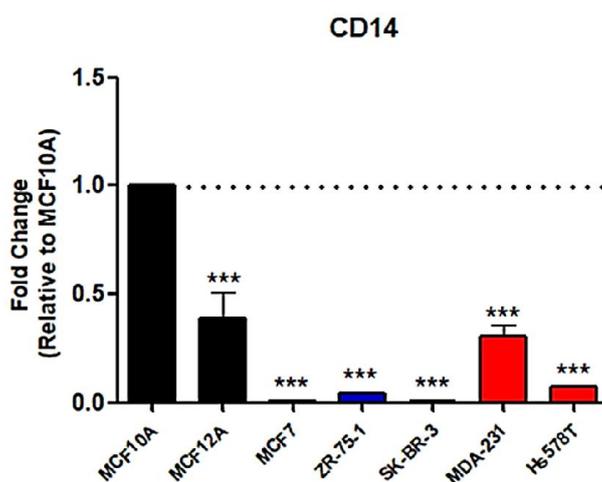
Para a análise da expressão relativa do mRNA nas linhagens celulares humanas de câncer de mama e nas linhagens geradas utilizou-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett. Para a análise das variáveis quantitativas e das comparações múltiplas, utilizou-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni. Todas as análises estatísticas foram realizadas utilizando-se o software *GraphPad Prism 5* (GraphPad, CA, EUA). Os dados são apresentados como a média ± SEM de três experimentos independentes, realizados em triplicata.

## 2.4. RESULTADOS

### 2.4.1. Análise da expressão do gene *CD14* no modelo de progressão tumoral do carcinoma mamário

Inicialmente, estabelecemos um painel de linhagens celulares humanas de câncer de mama, as quais representam os seus diferentes subtipos e apresentam diferentes graus de malignidade e de potencial metastático. Já foi relatado na literatura, pelo nosso grupo de pesquisa, que a expressão do *CD14* está correlacionada com o grau de malignidade em linhagens humanas de mama [Lobba *et al*, 2012]. Estes dados foram ampliados para outras linhagens celulares (MCF12A, ZR-75-1 e SK-BR-3) e confirmados no presente trabalho.

A partir dos dados de qRT-PCR, foi observado um elevado nível de expressão relativa do *CD14* na linhagem não-tumorigênica MCF10A, e um baixo nível de expressão nas linhagens tumorigênicas triplo negativas MDA-MB-231 e Hs578T (Figura 6), o que sugere um possível papel supressor tumoral para esse marcador, conforme descrito anteriormente por Lobba e colaboradores [2012].

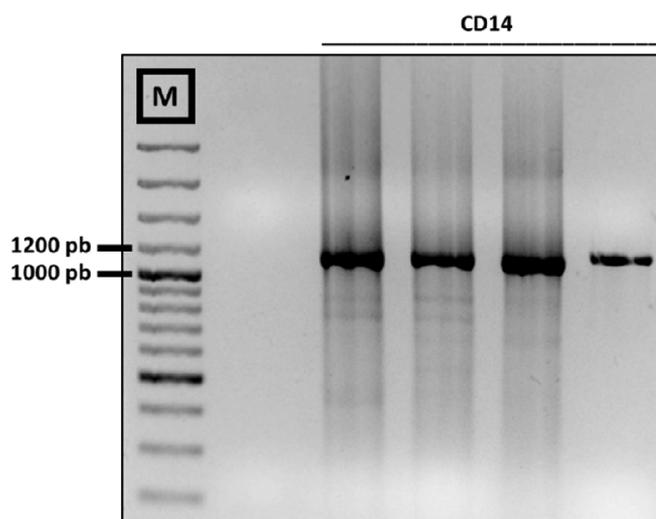


**Figura 6: Análise da expressão do gene *CD14* no painel de linhagens celulares humanas de câncer de mama.** O RNA total foi extraído conforme descrito em Materiais e Métodos e os níveis de expressão do *CD14* foram analisados por qRT-PCR utilizando-se *primers* específicos. O gráfico representa a razão de fold change relativo à expressão na linhagem MCF10A  $\pm$  SEM. O gene endógeno utilizado para normalização foi o HPRT1. As análises estatísticas foram feitas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett (\*\*\*)  $P < 0.0001$ ). Foram realizados três experimentos independentes, em triplicatas.

Baseado nestes resultados, resolvemos investigar o papel do gene *CD14* na progressão tumoral do câncer de mama triplo negativo. Para tanto, superexpressamos o gene codificador de proteína *CD14* na linhagem triplo negativa Hs578T, a qual apresenta elevado potencial tumorigênico e metastático.

#### 2.4.2. Construção do vetor lentiviral pLV-EGFP/*CD14* para ensaio de superexpressão do gene *CD14*

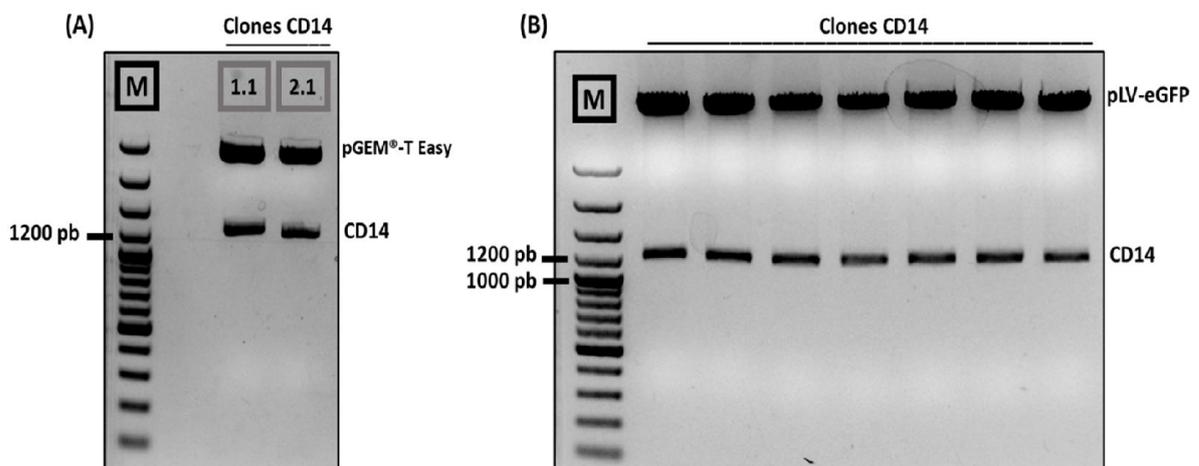
Para amplificação da região codificadora do gene *CD14*, foi utilizado, como molde, cDNA da linhagem celular não-tumorigênica de mama MCF10A e primers específicos, desenhados conforme descrito em Materiais e Métodos. Através da técnica de PCR, foi possível amplificar os fragmentos de cDNA de 1.128pb do *CD14* (Figura 7).



**Figura 7: Amplificação do cDNA de *CD14* humano.** Gel de agarose 0,8% para visualização do fragmento amplificado de *CD14*. M: Marcador de Peso Molecular.

A sequência codificadora do *CD14* foi purificada do gel e subclonada em vetor *pGEM®-T Easy*. Através da técnica de PCR, e após a digestão do plasmídeo *pGEM/CD14* com enzimas de restrição específicas, foi verificada a presença do inserto de interesse em dois clones bacterianos gerados. Os fragmentos obtidos apresentaram o tamanho esperado, levando-se em conta que os *primers* para o vetor

*pGEM®-T Easy* aumentam o tamanho do fragmento em cerca de 280pb (Figura 8A). Ambos os clones, os quais apresentaram o gene de interesse, foram submetidos ao sequenciamento e a análises de Bioinformática, com o objetivo de confirmar a integridade da sequência do gene *CD14*. Após a confirmação, os fragmentos do *CD14* foram purificados do gel para a subclonagem nos respectivos sítios de restrição do vetor lentiviral pLV-EGFP visando à superexpressão deste gene. Através da técnica de PCR, e após a digestão do plasmídeo pLV-EGFP/CD14 com enzimas de restrição específicas (*XbaI* e *MluI*), foi verificada a presença do inserto de interesse em diversos clones bacterianos gerados (Figura 8B).

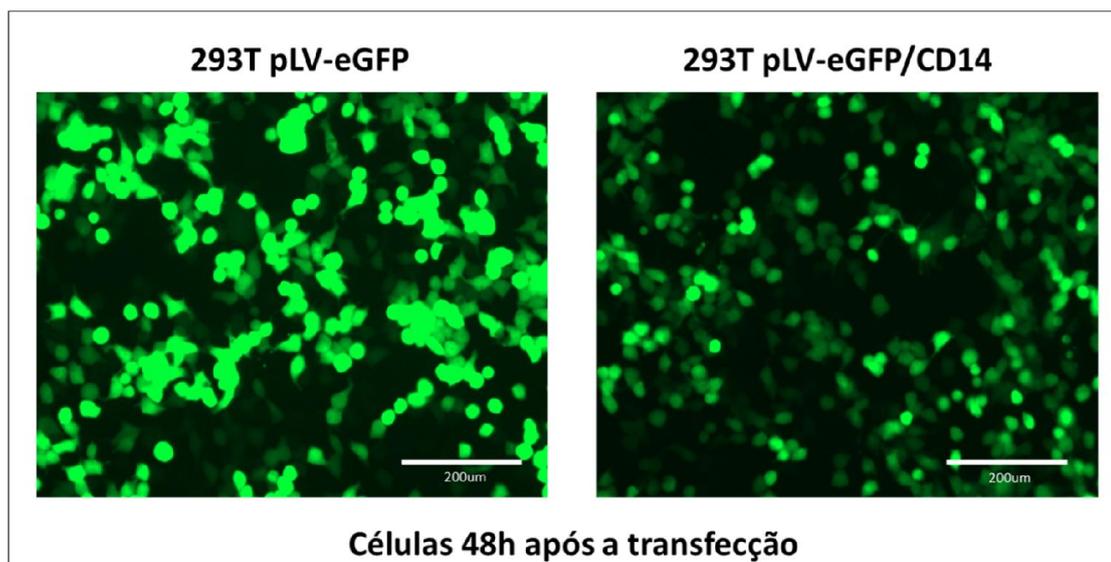


**Figura 8: Digestão dos plasmídeos recombinantes para confirmação dos transfectantes.** (A) Digestão do DNA plasmídeo obtido de dois clones bacterianos pGEM/CD14. (B) Digestão do DNA plasmídeo obtido de dois clones bacterianos pLV-eGFP/CD14. Os clones foram digeridos com as enzimas *XbaI* e *MluI* e, em seguida, visualizados em gel de agarose 1,2% para detecção dos clones transformantes contendo o gene de interesse. M: Marcador de peso molecular.

#### 2.4.3. Geração da linhagem celular Hs578T superexpressando o gene *CD14*

De posse de um dos clones contendo o plasmídeo recombinante pLV-eGFP/CD14, as partículas virais recombinantes foram geradas por meio de transfecção da linhagem 293T, conforme descrito em Materiais e Métodos. Após 48h e através de microscopia de fluorescência, as células transfectadas foram analisadas quanto à expressão do gene repórter eGFP. A presença da fluorescência nas células

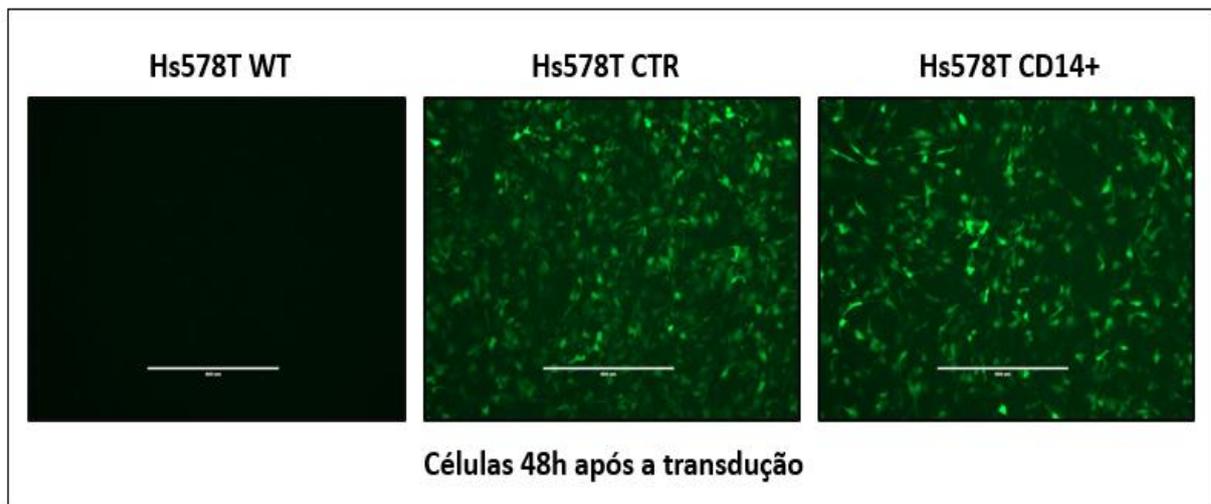
indica que as partículas lentivirais foram geradas de forma eficiente (Figura 9). O meio condicionado foi coletado conforme descrito em Materiais e Métodos, nos tempos de 24, 48 e 72h após a transfecção. Os sobrenadantes de cada coleta foram unidos, centrifugados a 20.000xg por 10 min à temperatura ambiente e, então, utilizados no processo de transdução celular.



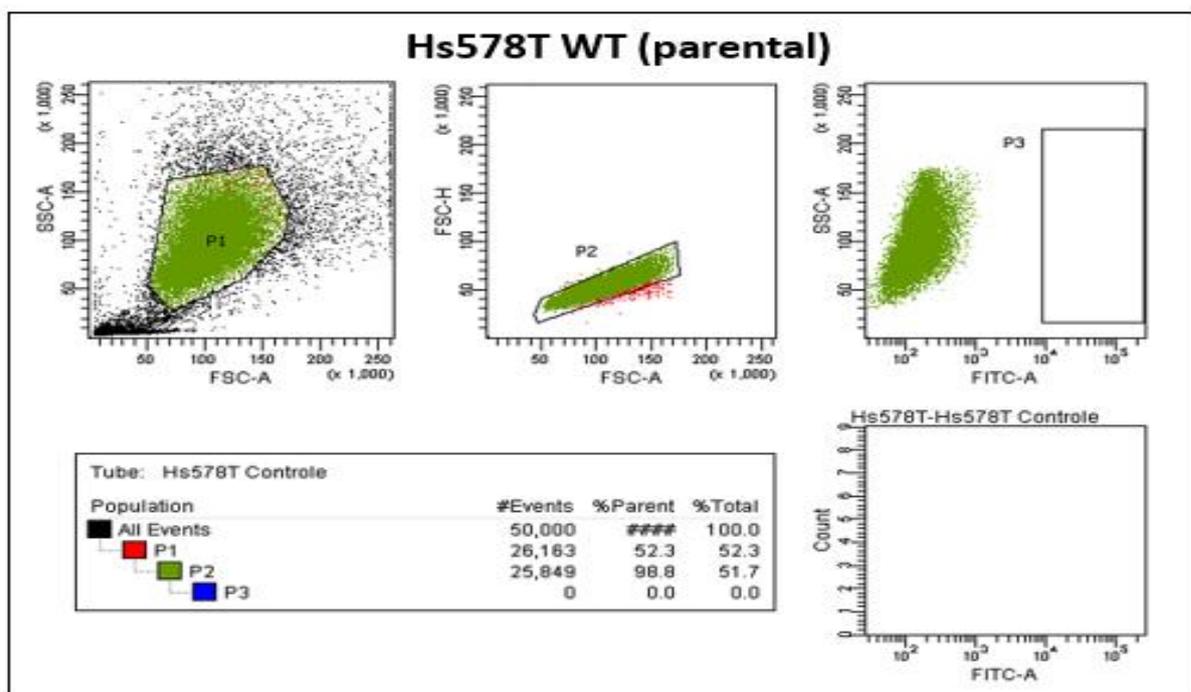
**Figura 9: Células 293T transfectadas com as construções para produção de partículas lentivirais.** Os vetores pLV-eGFP e pLV-eGFP/CD14 e os outros vetores para geração das partículas lentivirais foram co-transfectados em células 293T. Após 48h de transfecção, as células foram observadas ao microscópio de fluorescência utilizando-se microscopia de fluorescência a 509nm para detecção de células eGFP<sup>+</sup> indicando o sucesso da transfecção e geração das partículas lentivirais. Micrografias obtidas utilizando-se o microscópio *EVOS FL Cell Imaging Systems* (Thermo Scientific), no aumento de 200X. Barra: 200µm.

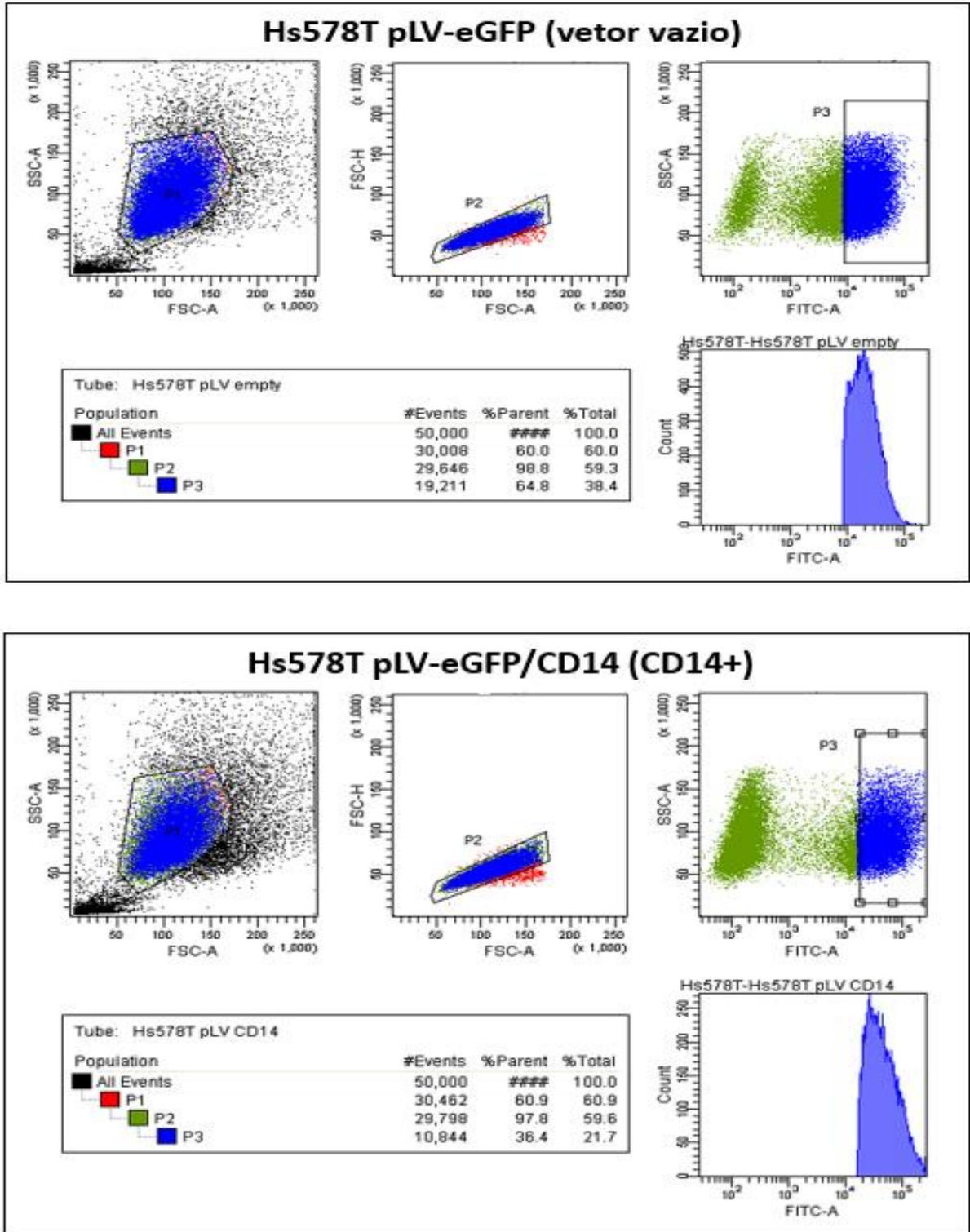
Células da linhagem 293T foram transduzidas para realizar o cálculo do título das partículas lentivirais. Após 48h, as células foram dissociadas e a porcentagem de células eGFP positivas foi analisada por citometria de fluxo. Observou-se que cerca de 3,8% e 3,3% das células transduzidas com o vetor controle pLV-eGFP e com o vetor pLV-eGFP/CD14 eram eGFP positivas, respectivamente. Conforme descrito em Materiais e Métodos, a partir destas porcentagens, foi realizado o cálculo do título lentiviral. Para transduzir a linhagem tumorigênica Hs578T (Hs WT) foi utilizado multiplicidade de infecção (MOI) de 10 das partículas lentivirais contendo os vetores pLV-eGFP e pLV-eGFP/CD14, visando, assim, obter uma linhagem controle e uma

linhagem superexpressando o *CD14* (Figura 10). Após a transdução, observou-se aproximadamente 38,4% de células eGFP positivas na linhagem transduzida com o vetor controle pLV-eGFP (Hs CTR) e cerca de 21,7% na linhagem transduzida com o vetor pLV-eGFP/CD14 (Hs CD14+). Apesar do baixo rendimento, foi possível selecionar a subpopulação eGFP positiva dessas linhagens através da técnica de “*cell sorting*” (Figura 11).



**Figura 10: Células Hs578T transduzidas com as construções lentivirais.** Os vetores pLV-eGFP e pLV-eGFP/CD14 foram transduzidos em células Hs578T. Após 48h de transdução, as células foram observadas ao microscópio de fluorescência utilizando-se microscopia de fluorescência a 509nm para detecção de células eGFP<sup>+</sup> indicando o sucesso da transdução. Micrografias obtidas utilizando-se o microscópio EVOS FL *Cell Imaging Systems* (Thermo Scientific), no aumento de 100X. Barra: 400μm.

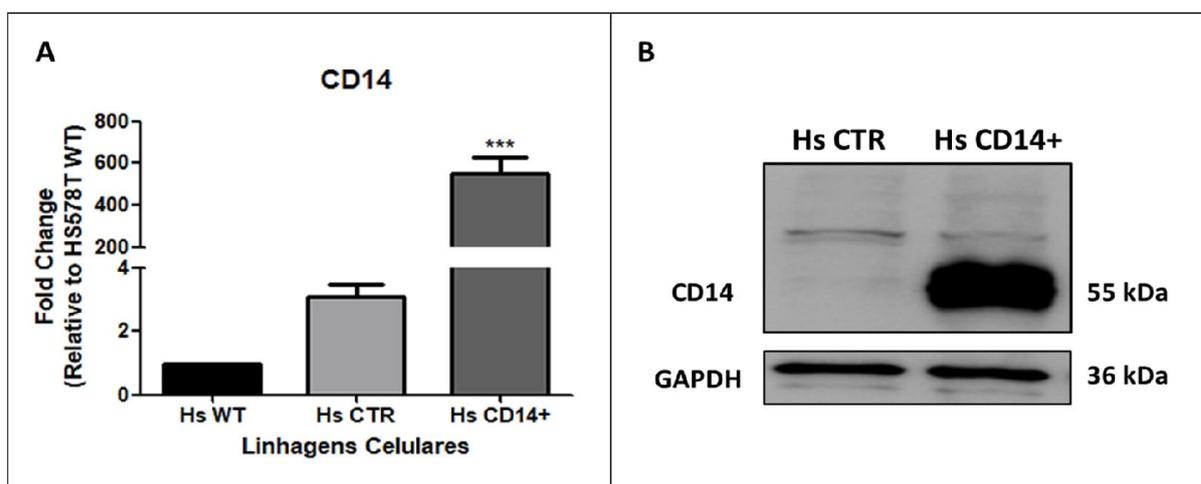




**Figura 11: Separação das células Hs578T fluorescentes transduzidas com as construções pLV-eGFP (Hs CTR) e pLV-eGFP/CD14 (Hs CD14+) por citometria de fluxo.** As células Hs578T foram transduzidas utilizando-se MOI=10 das partículas lentivirais. Após a transdução, as células foram dissociadas e as células eGFP<sup>+</sup> foram isoladas do restante da população através de citometria de fluxo, utilizando-se o FACSaria II, de acordo com o *gate* da imagem. A compensação foi realizada antes de cada experimento de maneira automática no próprio programa do citômetro FACSaria II.

#### 2.4.4. Confirmação da superexpressão do *CD14* na linhagem Hs578T transduzida

O nível de mRNA correspondente ao *CD14* nas linhagens Hs578T transduzidas foi confirmado através de qRT-PCR (Figura 12 A) e a expressão proteica foi confirmada por *Western Blot* (Figura 12 B).



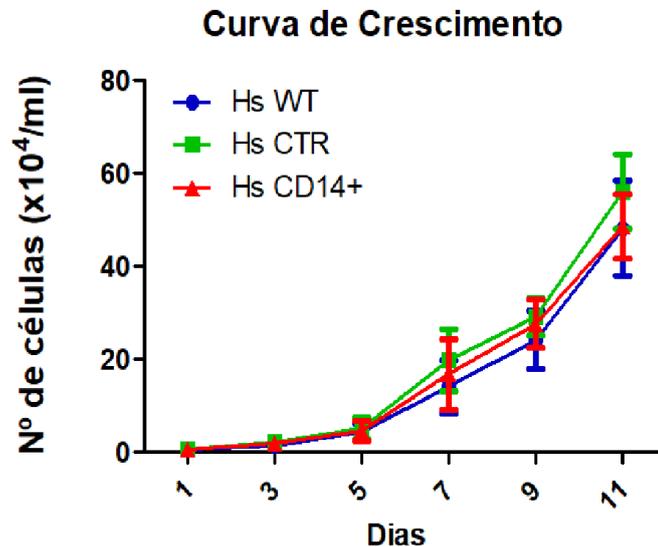
**Figura 12: Análise da expressão do gene *CD14* nas linhagens geradas.** A) Análise da expressão do *CD14*, através de qRT-PCR, nas linhagens Hs578T parental (Hs WT) e naquelas transduzidas com o vetor vazio (pLV-eGFP - Hs CTR) e com o vetor de superexpressão do *CD14* (pLV-eGFP/*CD14* - Hs CD14+). O gráfico representa a razão de *fold change* relativo à expressão na linhagem Hs WT  $\pm$  SEM. O gene endógeno utilizado para normalização foi o HPRT1. As análises estatísticas foram feitas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett (\*\*\*)  $P < 0.0002$ ). Foram realizados três experimentos independentes, em triplicatas. B) Análise da expressão de *CD14* por Western-Blot nas linhagens Hs578T transduzidas com vetor vazio (pLV-eGFP - Hs CTR) e com o vetor de superexpressão do *CD14* (pLV-eGFP/*CD14* - Hs CD14+). Os extratos proteicos das células fluorescentes previamente selecionadas por *cell sorting* foram fracionados em gel SDS-PAGE, transferidos para a membrana de nitrocelulose e incubados com o anticorpo anti-*CD14*.

#### 2.4.5. Caracterização das linhagens celulares geradas

##### 2.4.5.1. Curva de crescimento celular

Para investigar a influência da proteína *CD14* nas taxas de proliferação celular foram realizados ensaios *in vitro* de curva de crescimento celular. Foram realizados três experimentos independentes, em triplicatas, com a linhagem parental Hs578T (Hs WT), com a linhagem controle (Hs CTR) e com aquela transduzida com o vetor de superexpressão do gene *CD14* (Hs CD14+) (Figura 13). Conforme observado, não

houve diferença significativa de crescimento entre as linhagens, indicando que a superexpressão do *CD14* não influencia na taxa de proliferação celular.



**Figura 13: Análise da taxa de proliferação de células Hs578T superexpressando o gene *CD14*.** O ensaio de curva de crescimento foi realizado inoculando-se  $1 \times 10^4$  células em placas de 35mm de diâmetro. Após 1, 3, 5, 7, 9 e 11 dias de cultivo, estas células foram coletadas e contadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância Two-Way ANOVA, seguido de teste de comparação múltipla Bonferroni.

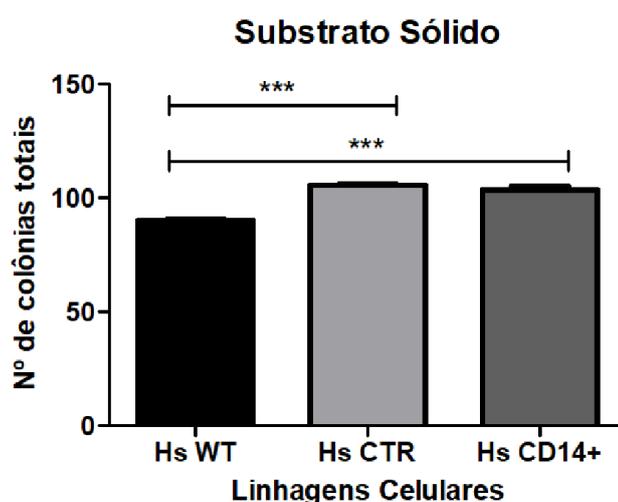
#### 2.4.5.2. Ensaios de tumorigênese *in vitro*

Para investigar a influência da proteína CD14 no processo de tumorigênese, a capacidade de formação de colônias das células foi avaliada através de duas abordagens *in vitro*: ensaio de crescimento de colônias em substrato sólido e em suspensão de agarose. Foram realizados três experimentos independentes, em triplicatas, com a linhagem parental Hs578T (Hs WT), com a linhagem controle (Hs CTR) e com aquela transduzida com o vetor de superexpressão do gene *CD14* (Hs CD14+).

##### 2.4.5.2.1. Ensaio de crescimento de colônias em substrato sólido

Para investigar a influência da proteína CD14 na tumorigênese, medida através da capacidade de formação de colônias dependente de ancoragem, foram realizados ensaios *in vitro* de crescimento de colônias em substrato sólido (Figura 14). Pode-se

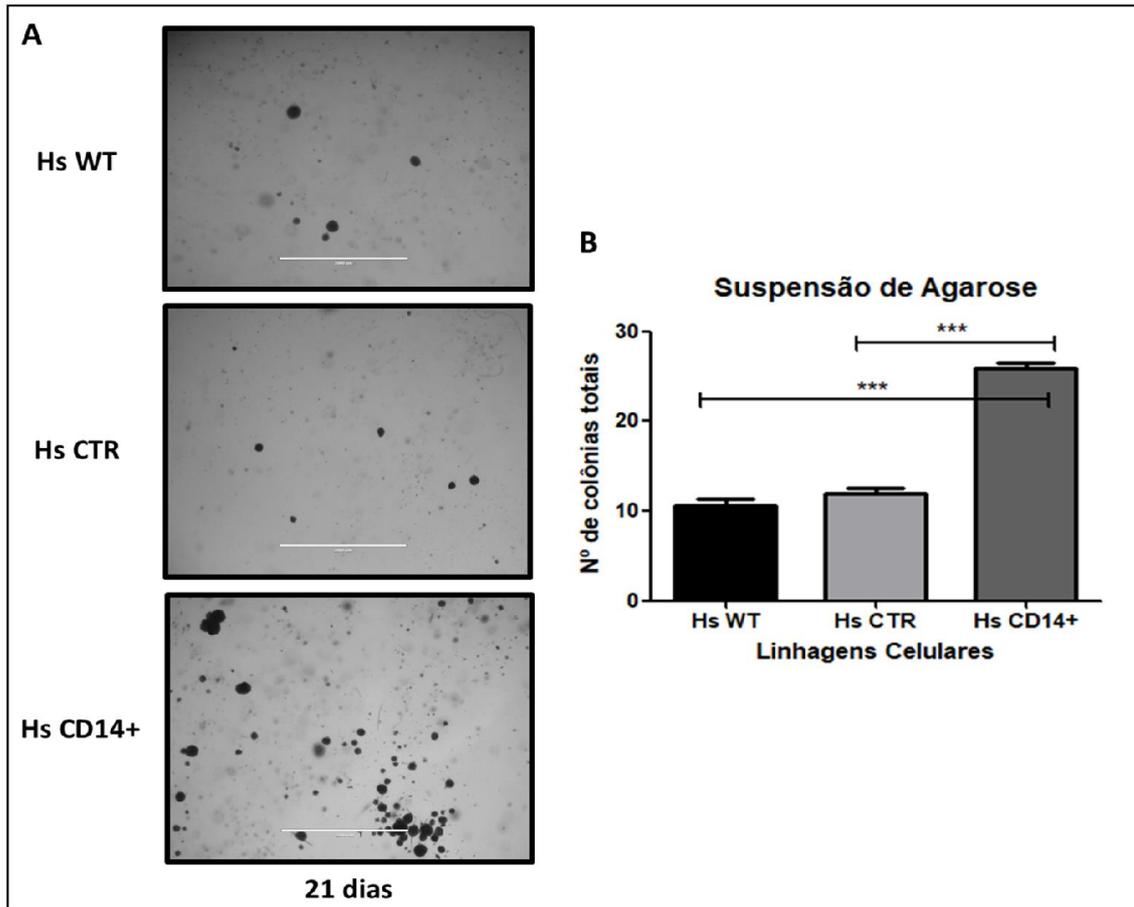
observar que as linhagens Hs CTR e Hs CD14+ dão origem a um número maior de colônias quando comparadas com a linhagem parental, indicando que estas linhagens apresentaram uma maior capacidade de formar colônias. No entanto, nenhuma diferença significativa foi observada quando comparamos a linhagem Hs CD14+ com a linhagem Hs CTR. Diante deste fato, não podemos afirmar se o aumento na capacidade de formar colônias se deve à presença do vetor ou da proteína CD14.



**Figura 14: Análise do potencial tumorigênico de células Hs578T superexpressando o gene *CD14* através do ensaio de crescimento de colônias em substrato sólido.** O número de colônias foi quantificado manualmente ao final de cada experimento. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni ( $***P < 0,05$ ).

#### 2.4.5.2.2. Ensaio de crescimento de colônias em suspensão de agarose

Para investigar a influência da proteína CD14 na tumorigênese, medida através da capacidade de formação de colônias independente de ancoragem das células, foram realizados ensaios *in vitro* de crescimento de colônias em suspensão de agarose. Houve crescimento de colônias nas três linhagens analisadas (Figura 15 A e 15 B), mas, diferentemente do esperado, a linhagem Hs CD14+ foi capaz de formar um número maior de colônias, em relação à linhagem Hs WT e Hs CTR, indicando que a superexpressão do *CD14* nestas células aumentou ainda mais o seu potencial tumorigênico.



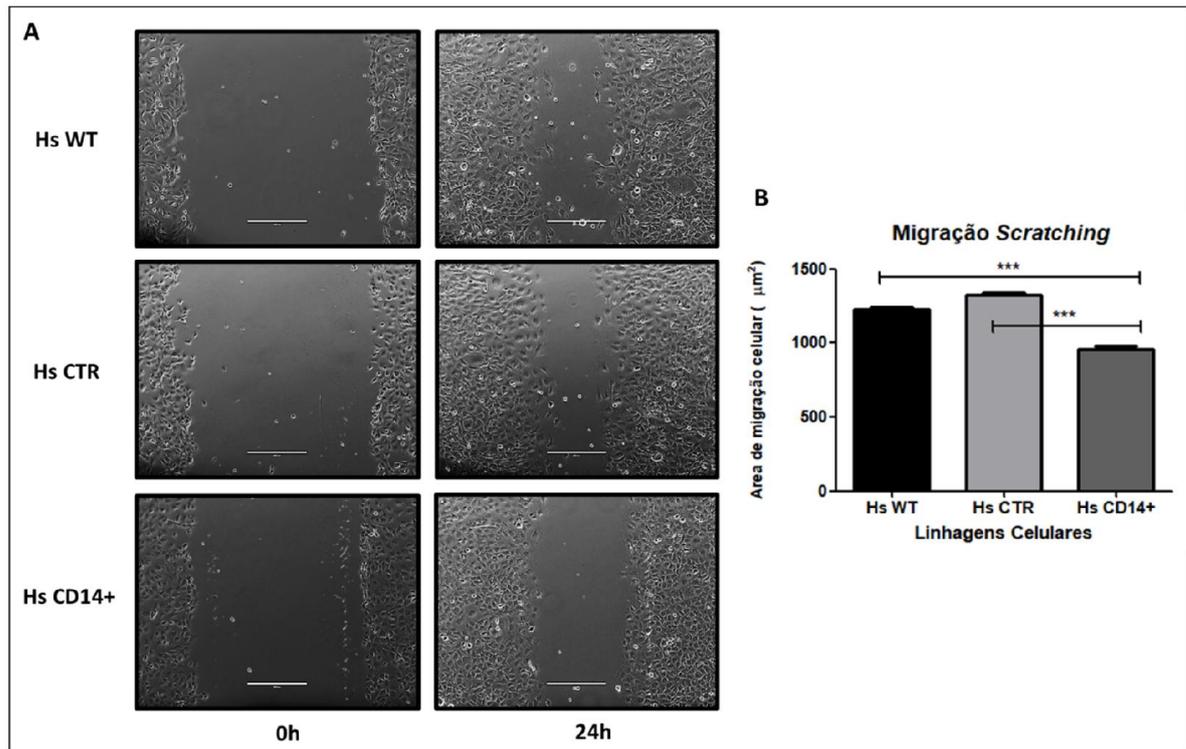
**Figura 15: Análise do potencial tumorigênico de células Hs578T superexpressando o gene *CD14* através do ensaio de crescimento de colônias em suspensão de agarose.** A) Foto representativa referente à cada linhagem celular ao final do experimento de 21 dias. B) Quantificação do número total de colônias ao final do experimento. Foram consideradas colônias apenas aquelas que apresentaram diâmetro maior do que 250µm. As colônias foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 40x. Barra: 2000µm. Os dados são apresentados como a média ± SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\*P<0,05).

#### 2.4.5.3. Ensaio de migração celular *in vitro*

Para investigar a influência da proteína CD14 no processo de migração celular, a capacidade migratória destas células foi avaliada através de duas abordagens *in vitro*: migração por *scratching* e por *transwell*. Foram realizados três experimentos independentes, em triplicatas, com a linhagem parental Hs578T (Hs WT), com a linhagem controle (Hs CTR) e com aquela transduzida com o vetor de superexpressão do gene *CD14* (Hs CD14+).

#### 2.4.5.3.1 Ensaio de migração celular por *scratching* (*wound healing*)

Nesta abordagem, pode-se observar que as células Hs CD14+ apresentam uma capacidade migratória significativamente menor, no tempo de 24 horas após o *scratching*, quando comparada com a Hs WT e a Hs CTR, indicando uma possível interferência da proteína CD14 no processo de migração celular (Figura 16).

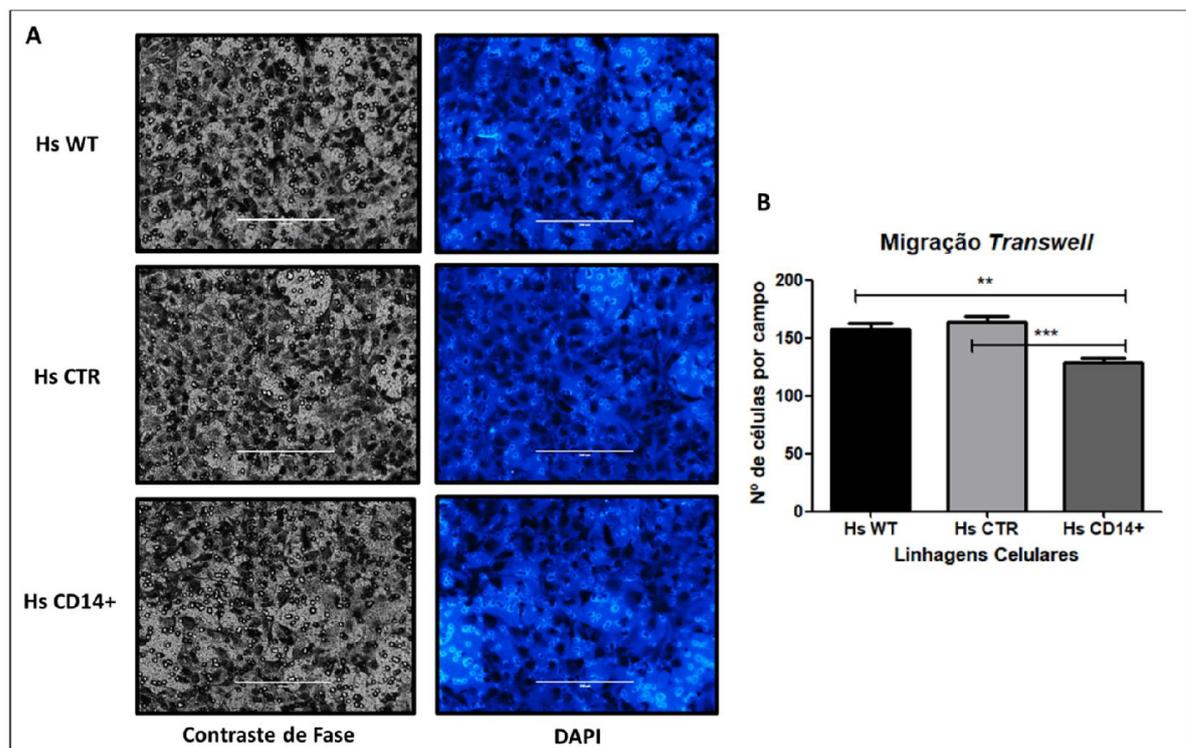


**Figura 16: Análise da capacidade migratória de células Hs578T superexpressando o gene CD14.** A) Fotos representativas referentes à cada linhagem celular ao final do experimento, nos tempos de 0h e 24h. B) O gráfico representa a quantificação da fração de área do risco preenchida pelas células 24h após o *scratching*. As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 100x. Barra: 400µm. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizado em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\*P<0,0001).

#### 2.4.5.3.2. Ensaio de migração celular por *Transwell*<sup>TM</sup>

Nesta abordagem, as células foram plaqueadas sobre insertos do tipo *Transwell*<sup>TM</sup> em placas de 24 poços e mantidas em cultura por 16h-18h, conforme descrito em Materiais e Métodos. Corroborando o resultado obtido no ensaio de migração por *scratching*, observou-se uma diminuição significativa do potencial migratório na linhagem Hs578T superexpressando o *CD14* (Hs CD14+), quando

comparada com a linhagem parental (Hs WT) e a linhagem controle (Hs CTR) (Figura 17). Baseado nos resultados de ambos os experimentos de migração celular *in vitro*, conclui-se, portanto, que a superexpressão do *CD14* influencia o potencial migratório das células Hs578T.

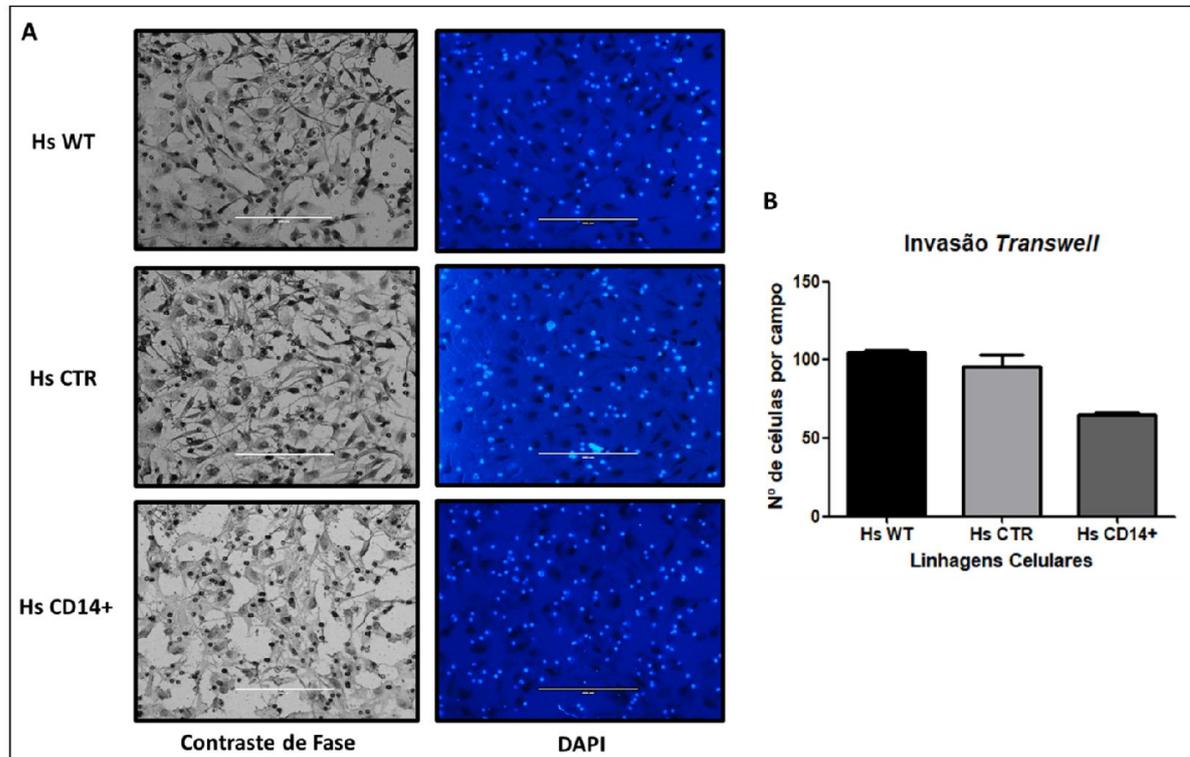


**Figura 17: Análise do potencial migratório de células Hs578T superexpressando o gene *CD14*.** A) Fotos representativas referente à cada linhagem celular ao final do experimento, em campo de contraste de fase e em DAPI. B) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 200x. Barra: 200µm. Os dados são apresentados como a média ± SEM de três experimentos independentes, realizados em duplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,0007$ ).

#### 2.4.5.4. Ensaio de invasão celular *in vitro*

Para investigar a influência da proteína CD14 no processo de invasão celular foram realizados ensaios de invasão celular *in vitro*, conforme descrito em Materiais e Métodos. Foi realizado um único experimento independente, em duplicatas, com a linhagem parental Hs578T (Hs WT), com a linhagem controle (Hs CTR) e com aquela transduzida com o vetor de superexpressão do gene *CD14* (Hs CD14+). Pode-se observar que as células Hs CD14+ apresentam um potencial invasivo

significativamente menor, quando comparada com a Hs WT e a Hs CTR (Figura 18). Como apenas um único experimento independente foi realizado, não podemos concluir se, de fato, esta alteração é devida à superexpressão do *CD14* nestas células.



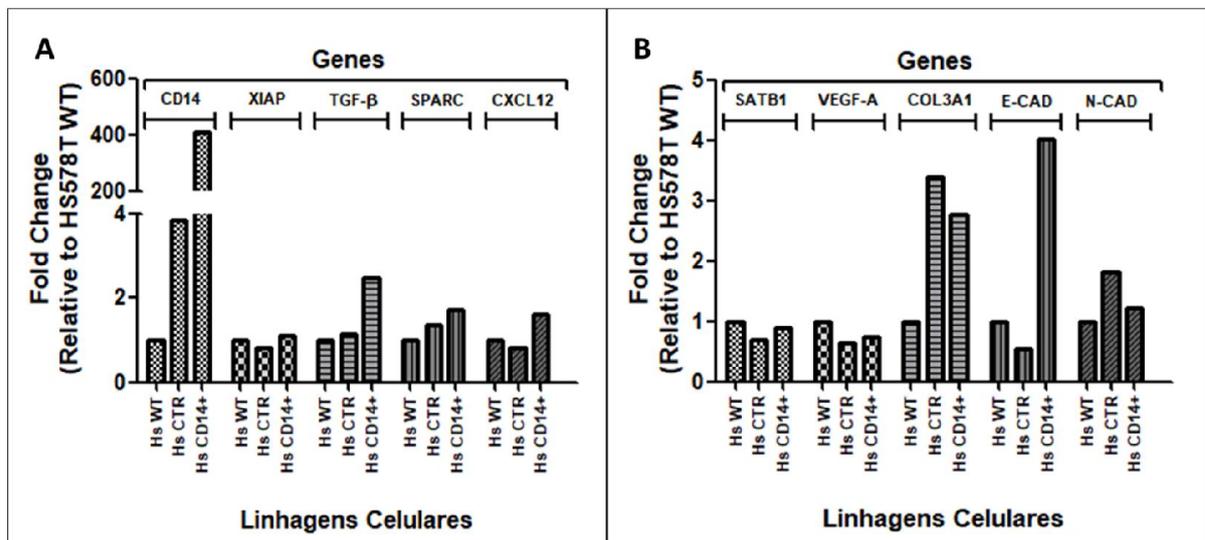
**Figura 18: Análise do potencial invasivo de células Hs578T superexpressando o gene *CD14*.** A) Fotos representativas referente à cada linhagem celular ao final do experimento, em campo de contraste de fase e em DAPI. B) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 200x. Barra: 200µm. Os dados são apresentados como a média  $\pm$  SEM de um único experimento independente, realizado em duplicata.

#### 2.4.5.5. Análise da expressão de genes envolvidos na malignidade tumoral

Para compreender os possíveis efeitos da superexpressão do *CD14* nos processos de apoptose, angiogênese, migração e invasão celular, investigamos a expressão de diferentes genes envolvidos na malignidade tumoral por qRT-PCR. Foram analisados os seguintes genes: 1) Gene *XIAP* (anti-apoptótico); 2) Genes *TGF- $\beta$* , *SPARC*, *CXCL12*, *SATB1*, *E-CAD*, *N-CAD* (associados aos processos de tumorigênese e transição epitélio-mesenquimal); 3) Genes *VEGF-A* e *COL3A1* (associados à angiogênese tumoral). Como apenas um experimento independente foi

realizado, os resultados do mesmo foram utilizados para nos dar uma visão inicial e geral de possíveis marcadores/vias alterados com a superexpressão do *CD14* na linhagem triplo negativa Hs578T.

Conforme observado na Figura 19, praticamente não houve alteração da expressão, nas linhagens analisadas, do gene anti-apoptótico *XIAP*, dos genes associados aos processos de tumorigênese e transição epitélio-mesenquimal *SPARC*, *SATB1* e *N-CAD*, e por fim, do gene angiogênico *VEGF-A*. Para os genes associados aos processos de tumorigênese e transição epitélio-mesenquimal *TGF- $\beta$* , *CXCL12* e *E-CAD* foi observado o aumento da expressão na linhagem superexpressando o gene *CD14* (Hs CD14+), quando comparado com as linhagens parental (Hs WT) e controle (Hs CTR). Por fim, o gene angiogênico *COL3A1* apresentou aumento da expressão na linhagem Hs CD14+, embora o aumento também tenha sido observado na linhagem Hs CTR.



**Figura 19: Análise da expressão de genes envolvidos na malignidade tumoral em células Hs578T superexpressando o gene codificador de proteína CD14.** (A) Análise da expressão dos genes *LINC01133*, *XIAP*, *TGF- $\beta$* , *SPARC* e *CXCL12*. B) Análise da expressão dos genes: *SATB1*, *VEGF-A*, *COL3A1*, *E-CAD* e *N-CAD*. Os genes foram analisados por qRT-PCR, nas linhagens Hs578T parental (Hs WT) e naquelas transduzidas com o vetor vazio (Hs CTR) e com o vetor de superexpressão do *CD14* (Hs CD14+). O gráfico representa a razão de *fold change* relativo à expressão na linhagem Hs WT. O gene endógeno utilizado para normalização foi o *HPRT1*. Os dados apresentados são referentes a apenas um experimento independente realizado em triplicata.

## 2.5. DISCUSSÃO

Ocupando posições de destaque, tanto na incidência quanto na mortalidade, o câncer de mama é considerado um principais problemas de saúde pública mundial. Atualmente, é o tipo de câncer maior incidência no Mundo, além de ocupar a 5ª posição em relação ao número de óbitos. Dentre os subtipos destes tumores identificados, o câncer de mama triplo negativo (TNBC) é considerado um desafio para a clínica, devido à sua heterogeneidade, às suas elevadas taxas de mortalidade e recidiva, associadas à escassez de tratamento direcionados eficazes. Nos últimos anos, um grande esforço tem sido feito para identificar novos biomarcadores e terapias direcionadas, mas poucos deles se mostraram úteis em ensaios clínicos [Cocco *et al.*, 2020], fazendo com que a quimioterapia citotóxica continue sendo a principal opção de tratamento para a maior parte dos pacientes com TNBC. Em busca de identificar e caracterizar novos marcadores proteicos que auxiliem na identificação e no manejo das pacientes, neste trabalho, aplicou-se técnicas de Biologia Celular e Molecular para analisar o potencial papel do marcador proteico de células-tronco CD14 no desenvolvimento e na progressão do câncer de mama triplo negativo (TNBC).

Inicialmente, analisamos a expressão do gene codificador de proteína *CD14*, frente a um painel de linhagens celulares humanas de câncer de mama que apresentam diferentes graus de malignidade e potencial metastático, sendo duas linhagens não-tumorigênicas (MCF12A e MCF10A), uma linhagem referente ao subtipo molecular luminal A (MCF-7), uma linhagem referente ao subtipo molecular luminal B (ZR-75-1), uma linhagem referente ao subtipo molecular HER2+ (SK-BR-3) e duas linhagens referentes ao subtipo molecular basal-like/triplo negativo (MDA-MB-231 e Hs578T). Corroborando os dados obtidos anteriormente por Lobba e

colaboradores [2012], a expressão do *CD14* mostrou-se elevada nas linhagens não-tumorigênicas MCF12A e MCF10A e baixa nas linhagens triplo negativas MDA-MB-231 e Hs578T, sugerindo um potencial papel de supressor tumoral para este gene.

Esses resultados corroboram os dados descritos na literatura, com diversos estudos tendo mostrado a expressão do *CD14* em células do tecido mamário e no leite materno. O primeiro relato da expressão de *CD14* foi em sua forma solúvel no leite materno, em um trabalho que associou elevados níveis de sCD14 com o controle da homeostase intestinal neonatal no primeiro encontro com um ambiente repleto de microrganismos, tendo a sCD14 um papel importante na ativação induzida por LPS das células CD14- da membrana, como as células epiteliais intestinais neonatais [Labéta *et al.*, 2000]. Posteriormente, foi descrito a presença da sCD14, em conjunto com a IL-12, no leite materno, sendo ambas importantes por promover o desenvolvimento de linfócitos Th1, evitando a produção excessiva de IgE e, conseqüentemente, quadros de inflamação alérgica [Camporota *et al.*, 2001]. Também já foi descrito, em diversos trabalhos que a ligação de LPS à sCD14 leva à ativação de vias que contribuem para a proteção contra a mastite no tecido mamário [Lee *et al.*, 2003; Zheng *et al.*, 2006; Wall *et al.*, 2009]. Em outro trabalho, a mCD14, em conjunto com LBP e STAT3, foi descrito por participar de uma cascata imune e na resposta de fase aguda no processo de involução da glândula mamária, em células epiteliais luminais mamárias de camundongos [Stein *et al.*, 2004]. A expressão do gene *CD14* no epitélio luminal mamário durante a involução indica a aquisição de um fenótipo fagocítico, o qual pode facilitar a remoção de células epiteliais apoptóticas eliminadas na luz alveolar [Monks *et al.*, 2008]. A proteína CD14 também já foi identificada, em diversos trabalhos, como um marcador de células progenitoras luminais mamárias [Asselin-Labat *et al.*, 2007; Asselin-Labat *et al.*, 2011; Bach *et al.*,

2017], as quais foram caracterizadas por Shehata e colaboradores, que mostraram a expressão do gene *CD14* em 86% das células progenitoras examinadas, confirmando, assim, sua relevância no tecido mamário (Shehata *et al.*, 2012). A partir destes resultados, *CD14* pode ser considerado como um marcador promissor e seu papel funcional durante a progressão do câncer de mama, especificamente o subtipo triplo negativo (TNBC), dever ser melhor investigado.

Para tentar descobrir o potencial papel da proteína *CD14* no modelo celular de câncer de mama triplo negativo, linhagens celulares derivadas a partir da mama humana superexpressando esta proteína foram geradas. Para tanto, utilizamos o sistema lentiviral de superexpressão em células de mamífero pLV-eGFP e a linhagem celular humana de câncer de mama triplo negativo Hs578T. A superexpressão do *CD14* foi confirmada com êxito, tanto ao nível de mRNA quanto de proteína. O êxito na superexpressão de genes utilizando este sistema lentiviral já foi relatada em outros trabalhos do nosso grupo. Trombetta-Lima e colaboradores utilizaram este sistema de superexpressão para caracterizar variantes de *splicing* do gene supressor tumoral *RECK* em linhagens celulares humanas de astrocitoma [Trombetta-Lima *et al.*, 2015]. Lobba e colaboradores também utilizaram este sistema para caracterizar o marcador proteico de células-tronco *CD90* em linhagens celulares humanas de câncer de mama [Lobba *et al.*, 2018].

O primeiro resultado de caracterização funcional mostrou que a superexpressão do *CD14* não alterou a taxa de proliferação das células indicando que a proteína não influencia neste processo. Estes resultados estão de acordo com os resultados da literatura, pois até o momento não foi descrito o envolvimento da proteína *CD14* neste processo no carcinoma mamário. No entanto, já foi relatado na literatura que a ativação do *CD14* em células de câncer de bexiga foi responsável por

orquestrar a inflamação promotora de tumor e conduzir à proliferação celular, através da indução da sinalização das proteínas adaptadoras MyD88 (proteína de resposta primária de diferenciação mielóide 88) e TRIF (proteína adaptadora contendo domínio TIR indutora de interferon  $\beta$ ) [Cheah *et al.*, 2015]. Essas proteínas adaptadoras iniciam um processo de transdução de sinal para o núcleo ativando as vias NF- $\kappa$ B e MAPK, responsáveis por estimular a transcrição de fatores pro-inflamatórios, como a IL-6, IL-8, IL-1 $\beta$  e TNF- $\alpha$  [Granucci & Zanoni, 2015].

Outro parâmetro analisado neste trabalho foi o papel do *CD14* na tumorigênese. Inicialmente, analisamos a capacidade de formação de colônias dependente de ancoragem, visto através do ensaio clonogênico, no qual observou-se um pequeno aumento no número de colônias nas linhagens controle (Hs CTR) e superexpressando o *CD14* (Hs CD14+) quando comparadas com a linhagem parental (Hs WT). Diante disto, o aumento da capacidade de formar colônias nestas linhagens estaria hipoteticamente relacionado ao vetor e não à superexpressão da proteína CD14. Uma possível hipótese que poderia explicar este resultado é a inserção das sequências vetoriais em regiões genômicas destas células responsáveis por regular o processo de formação de colônias dependentes de ancoragem.

Uma outra característica analisada, foi a capacidade de formação de colônias independente de ancoragem, através do ensaio de crescimento em suspensão de agarose. Considerado uma marca registrada do processo de tumorigênese, o crescimento independente de ancoragem avalia a capacidade das células transformadas de crescerem independentemente de uma superfície sólida. O ensaio de crescimento em suspensão de agarose é uma das técnicas mais amplamente utilizadas para avaliar a tumorigênese *in vitro* [Borowicz *et al.*, 2014]. Nossos resultados, indicaram um aumento considerável no número de colônias na linhagem

Hs CD14+ quando comparada com as linhagens Hs WT e Hs CTR, sugerindo um aumento do potencial tumorigênico nestas células. As vias que levam ao aumento do potencial tumorigênico destas células ainda precisam ser desvendadas, mas a união de algumas hipóteses podem ser levantadas para justificá-lo. O envolvimento do CD14 na capacidade de formação de colônias independente de ancoragem já foi descrito por Pal e colaboradores, os quais mostraram que células progenitoras luminais CD14+/CD55+ apresentam uma maior capacidade de formação de colônias independentes de ancoragem quando comparadas com progenitoras luminais CD14-/CD55+ [Pal *et al.*, 2017], sendo esta uma das hipóteses que podem justificar o aumento do potencial tumorigênico observado na linhagem Hs CD14+.

A segunda hipótese baseia-se no eixo de sinalização envolvendo TLR2/4/6-MyD88-IRAK1 dependente de CD14. Receptores *toll-like* (TLR) são membros de uma família de receptores de reconhecimento padrão (PRRs) responsáveis por reconhecer padrões moleculares associados a patógenos (PAMPs), como o lipopolissacarídeo bacteriano (LPS), ou padrões moleculares associados ao dano (DAMPs), como alguns proteoglicanos presentes na matriz extracelular (por exemplo, biglican, versican e decorin) [Takeda & Akira, 2005]. Tanto os PAMPs quanto os DAMPs necessitam do co-receptor CD14 para ativação da sinalização mediada por TLR2/4/6. A proteína CD14 atua por meio da transferência de moléculas de LPS ou pelo sequestro de DAMPs presentes na matriz extracelular, apresentando-os para os TLRs e ativando as cascatas de sinalização. A ativação do eixo de sinalização TLR2/4/6-MyD88-IRAK1 por estas moléculas leva à ativação das vias NF- $\kappa$ B e MAPK, responsáveis por estimular a transcrição de fatores pro-inflamatórios, como IL-1 $\beta$ , IL-6, IL-8, IL-17, TNF- $\alpha$  e a catelicidina humana (hCAP)18/LL-37. Estes fatores são responsáveis pela inibição da apoptose e pela instalação da inflamação crônica, levando ao

desenvolvimento tumoral [Roedig *et al.*, 2020]. Altos níveis de expressão de DAMPs (biglican e versican) [Ricciardelli *et al.*, 2002; Skandalis *et al.*, 2011; Zhao *et al.*, 2020], de IRAK1 [Scheeran *et al.*, 2014], de citocinas inflamatórias e catelicidina humana (hCAP)18/LL-37 [Heilborn *et al.*, 2005] tem sido descritos em diversos tipos de câncer, entre os quais o câncer de mama triplo negativo, e estão associados à indução da inflamação tumoral e, conseqüentemente, ao aumento do potencial tumorigênico. Scheeran e colaboradores mostraram que a sinalização do TLR impulsiona o crescimento tumoral, *in vitro*, em linhagens triplo negativas de câncer de mama (MDA-MB-231 e MDA-MB-468). Além disto, o bloqueio da sinalização do TLR em tumores triplo negativos *in vivo* inibiu o crescimento tumoral, demonstrando que a via é utilizada por tumores *in vivo* [Scheeran *et al.*, 2014]. Diante disto, a superexpressão do *CD14* na linhagem triplo negativa Hs578T pode ter levado ao aumento do sequestro de DAMPs presentes na matrix extracelular, ativando, ainda mais, a sinalização dos TLRs e levando ao aumento do potencial tumorigênico observado nestas células.

Outro fator que pode contribuir para o aumento do potencial tumorigênico observado na linhagem Hs CD14+ é o aumento da expressão dos marcadores de malignidade TGF- $\beta$  e CXCL12, classicamente descritos por promover o crescimento primário de diversos tipos de câncer, incluindo o câncer de mama [Muraoka *et al.*, 2003; Luker *et al.*, 2012; Wu *et al.*, 2015]. O papel do TGF- $\beta$  na progressão do TNBC é pouco conhecido e os dados disponíveis na literatura são controversos. Alguns estudos mostram que a expressão elevada do TGF- $\beta$  está associado com alto grau histológico e redução da taxa de sobrevida livre da doença [Ding *et al.*, 2016; Zhang *et al.*, 2017]. Outros estudos mostram que a ausência ou baixos níveis de expressão estão associados com o alto grau histológico destes tumores [Cheung *et al.*, 2015]. Além disto, um outro estudo relatou que não existe correlação dos níveis de expressão

do TGF- $\beta$  com a progressão dos TNBC [Hachim *et al.*, 2016]. No entanto, alguns estudos tem sugerido que o TLR4 pode diretamente aumentar a sinalização do TGF- $\beta$  em casos de fibrose [Seki *et al.*, 2007; Bhattacharyya *et al.*, 2013]. Desta forma, um estudo mais aprofundado das vias de atuação do TGF- $\beta$  neste subtipo tumoral é importante para compreender de forma precisa o seu papel.

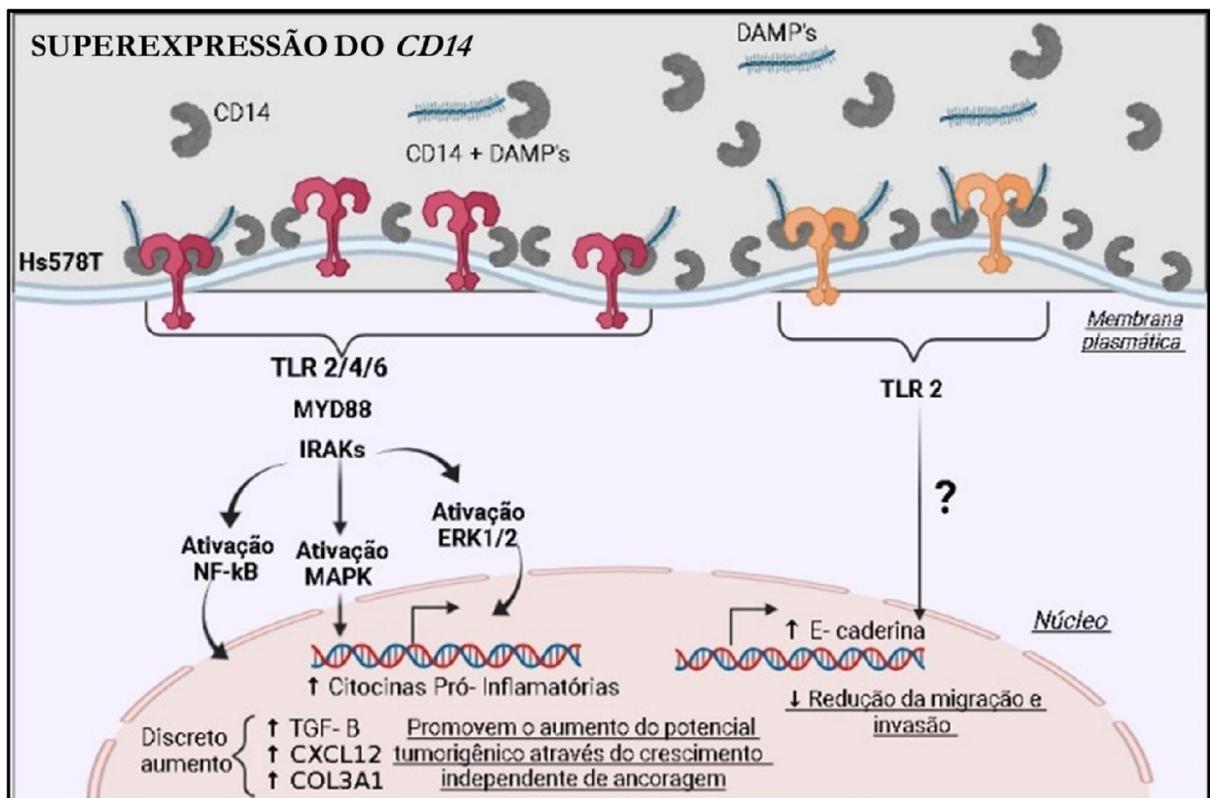
O aumento da capacidade de formação de colônias independente de ancoragem, observado na linhagem Hs CD14+ pode também estar relacionado com a redução da capacidade de migração e invasão observada nestas células. Observou-se uma redução significativa na capacidade de migração, visto através dos ensaios de *scratching* e *transwell*, e invasão (embora apenas um único experimento independente tenha sido realizado) destas células. Hipoteticamente, o aumento do potencial tumorigênico poderia ser um mecanismo compensatório para a significativa redução do potencial metastático observado.

Migração e invasão são características relacionadas ao processo de transição epitélio-mesenquimal (EMT), o qual é de extrema importância para a ocorrência de metástases. A EMT é um processo no qual células epiteliais normalmente polarizadas sofrem modificações morfológicas, bioquímicas e moleculares resultando na aquisição de um fenótipo de células mesenquimais, exibindo, assim, uma maior capacidade de migração, invasão e resistência à apoptose. Uma classe importante de moléculas que regulam este processo é a superfamília das integrinas, glicoproteínas de membrana, que são importantes por regular a integração dos meios intra e extracelulares. Dentre as integrinas, os membros da família das caderinas são importantes na formação das junções aderentes responsáveis pela adesão célula-célula. A alteração da expressão das caderinas, tais como a E-caderina e N-caderina, é um fenômeno amplamente descrito durante a progressão tumoral [Boyer *et al.*, 2000].

Em nosso trabalho, não se observou mudança no nível de expressão de *N-caderina*, mas observou-se um expressivo aumento na expressão da *E-caderina* na linhagem HsCD14+. A *E-caderina* é uma proteína descrita por ser supressora tumoral e a perda ou baixos níveis de sua expressão, em células tumorais, está associado ao processo de transição epitelial-mesenquimal e, conseqüentemente, á progressão das metástases [Nieto, 2011; Huang et al., 2012]. Níveis reduzidos de *E-caderina* foram observados em 238 de 574 casos de câncer de mama, sendo significativamente frequente em TNBC quando comparado com os outros subtipos moleculares [Kashiwagi et al., 2010; Liu et al., 2017]. A linhagem celular Hs578T apresenta um elevado potencial tumorigênico e metastático característico dos TNBC. Diante disto, o aumento da expressão de *E-caderina* na linhagem HsCD14+ justifica a redução do potencial migratório e invasivo observado na mesma. Além disto, já foi descrito na literatura que a estimulação do TLR2 aumenta fortemente a expressão de *E-caderina* em camundongos CD14<sup>-/-</sup> e WT, sendo assim, um dos possíveis fatores que levaram ao aumento da expressão de *E-caderina* na linhagem HsCD14+ [Basic et al., 2018].

Diante disto, pela primeira vez, o marcador proteico de células-tronco CD14 foi caracterizado funcionalmente, no presente, como uma importante molécula envolvida no desenvolvimento do câncer de mama triplo negativo. Mecanismos hipotéticos de possíveis vias de atuação deste marcador foram correlacionados com características importantes na progressão tumoral, os quais necessitam ser aprofundados no intuito de fornecer informações detalhadas sobre as bases biológicas deste subtipo tumoral, além de contribuir para o desenvolvimento de novas estratégias terapêuticas direcionadas (Figura 20). Os dados descritos aqui sugerem que a proteína CD14 possui um papel importante no controle do potencial tumorigênico e metastático na linhagem celular triplo negativa Hs578T. Além disto, descrevemos o importante papel

da via de sinalização TLR2/4/6-MyD88-IRAK1 ativada pelo CD14 no TNBC, tornando-a um importante alvo para o desenvolvimento de novas drogas direcionadas que tratem os pacientes deste subtipo tumoral. A possível combinação de estratégias terapêuticas que combinem a superexpressão do gene *CD14* e a inibição de moléculas importantes à jusante na via de sinalização dos TLRs, tais como MyD88 e IRAK1, poderia ser efetiva e levar à diminuição da malignidade do TNBC.



**Figura 20: Mecanismos hipotéticos ativados devido à superexpressão do gene *CD14* na linhagem triplo negativa Hs578T.** A superexpressão do gene *CD14* leva ao aumento da expressão da proteína CD14, tanto na sua forma ancorada à membrana quanto na sua forma solúvel. O aumento da quantidade da proteína CD14 leva ao aumento do sequestro de DAMPs presentes na matriz extracelular e, assim, a uma ativação excessiva da via de sinalização TLR2/4/6-MyD88-IRAK. A ativação desta via é responsável por aumentar, ainda mais, a produção de citocinas pró-inflamatórias e, com isto, promover o aumento do potencial tumorigênico observado nestas células. O leve aumento da expressão de *TGF-β*, *CXCL12* e *COL3A1* pode também contribuir para o aumento do potencial tumorigênico. Por fim, a ativação excessiva do TLR2 leva ao aumento da expressão de E-caderina e, conseqüentemente, à redução do potencial migratório e invasivo observados nas células.

## 2.6. CONCLUSÃO

No presente trabalho, foi possível superexpressar o gene codificador de proteína *CD14* na linhagem Hs578T, caracterizá-lo funcionalmente e descobrir sua relação com a malignidade do câncer de mama triplo negativo. De acordo com os dados obtidos, foi observado que o gene *CD14* apresenta baixa expressão na linhagem celular Hs578T, exibindo um importante papel na regulação de processos relacionados com a progressão tumoral, tais como o potencial tumorigênico e metastático. Portanto, demonstramos que a superexpressão do *CD14* reduziu a migração e invasão celular, mas, em contrapartida, aumentou a capacidade de formação de colônias independente de ancoragem desta linhagem. Diante disto, sugerimos que o nível basal de expressão do *CD14* naturalmente observado na linhagem Hs578T pode contribuir para o desenvolvimento primário do tumor, atuando como um oncogene. Juntos, nossos dados também sugerem um papel importante da via de sinalização CD14-TLR2/4/6-MyD88-IRAK1 na progressão do TNBC, tornando-a um potencial alvo terapêutico para o tratamento deste subtipo tumoral. Por fim, os dados obtidos em nosso estudo contribuem para o enriquecimento de informações relacionadas à Biologia do câncer de mama triplo negativo, auxiliando, desta forma, no desenvolvimento de potenciais protocolos clínicos e terapêuticos.

## 2.7. PRÓXIMOS PASSOS

Os próximos passos abaixo são importantes para complementação dos resultados obtidos neste trabalho:

- Finalizar os experimentos de invasão celular *in vitro*;
- Realizar reações de imunofluorescência no intuito de identificar possíveis alterações morfológicas nas células Hs578T superexpressando o gene *CD14*.
- Realizar ensaios de tumorigênese *in vivo* utilizando as células Hs578T superexpressando o gene *CD14*, avaliando parâmetros histopatológicos nos possíveis tumores gerados.
- Caracterizar, por Western-Blot e RNA-Seq, o mecanismo de ação do CD14 na progressão tumoral do câncer de mama triplo negativo.
- Caracterizar funcionalmente o *CD14* na linhagem não tumorigênica MCF10A através do nocauteamento gênico utilizando o sistema lentiviral CRISPR/Cas9.
- Realizar experimentos de inibição de alvos moleculares importantes da via de sinalização dos TLRs nas células Hs578T superexpressando o *CD14*.
- Analisar a expressão do *CD14* em amostras de pacientes normais e portadoras do câncer de mama triplo negativo.

### **3. CAPÍTULO II:**

**Identificação e caracterização funcional (superexpressão indutível e nocauteamento) do gene não codificador de proteína *LINC01133* em linhagens celulares humanas de câncer de mama triplo negativo**

### 3.1. INTRODUÇÃO

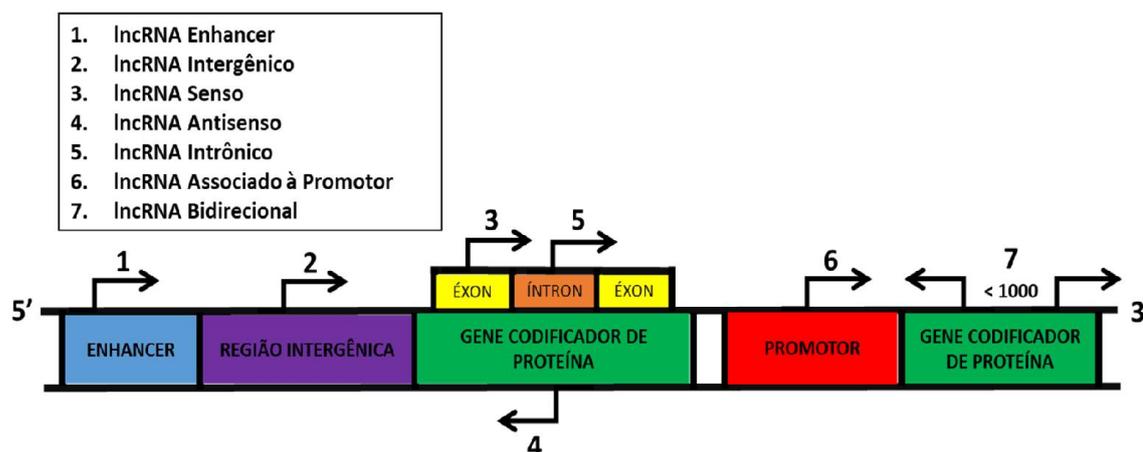
Nas últimas décadas, embora estudos genéticos tenham evoluído consideravelmente, o papel funcional dos RNAs permaneceu focado, por muito tempo, exclusivamente, na produção de proteínas [Gingeras, 2007]. Acreditava-se que o transcriptoma dos eucariotos era composto por RNAs mensageiros (mRNAs), RNAs transportadores (tRNAs), RNAs ribossomais (rRNAs), e pequenos RNAs nucleares (*small nuclear RNAs*, snRNA; *small nucleolar RNAs*, snoRNA), os quais exercem funções estruturais e catalíticas importantes, seja no processamento do mRNA ou na tradução das proteínas. Uma das maiores surpresas da Biologia moderna foi a descoberta de que o genoma humano possui apenas ~20,000 genes codificadores de proteína, representando menos do que 2% das sequências genômicas totais [Ponting & Belgard, 2010]. Com o advento de tecnologias de sequenciamento do genoma e do transcriptoma, foi estimado que pelo menos 90% do genoma é ativamente transcrito em RNAs [Costa, 2010]. Com isso, foi descoberto que o transcriptoma humano era bem mais complexo do que uma coleção de genes codificadores de proteínas e suas variantes de *splicing*, incluindo, também, uma parcela expressiva de RNAs não codificadores de proteínas (ncRNAs), os quais podem desempenhar papéis biológicos importantes no desenvolvimento celular e no metabolismo [Wilusz *et al.*, 2009; Ponting *et al.*, 2009; Mercer *et al.*, 2009].

Os ncRNAs são divididos em duas grandes classes, baseadas no tamanho do transcrito: ncRNAs curtos e ncRNAs longos (lncRNAs). NcRNAs curtos possuem menos do que 200nt de comprimento e incluem os bem documentados microRNAs (miRNAs), os pequenos RNAs de interferência (siRNAs), os RNAs que interagem com PIWI (piRNAs), os RNAs iniciadores da transcrição (tiRNAs), entre outros. LncRNAs são transcritos semelhantes a mRNAs que apresentam comprimentos maiores do que

200 nucleotídeos, com a maioria variando de 1.000 a 10.000 nucleotídeos, e não são traduzidos [Novikova *et al.*, 2013]. Eles são transcritos, principalmente, pela RNA polimerase II, frequentemente são poliadenilados, passam por *splicing* de éxons (indicando que é necessário processamento pós-transcrição para formar transcritos maduros) e os *loci* genômicos de muitos deles apresentam marcas de cromatina, consistentes com aquelas de genes transcritos, em seus promotores e nos corpos de genes, como *H3K9ac*, *H3K4me3* e *H3K36me3* [Hansji *et al.*, 2014]. Além disso, podem gerar estruturas secundárias tridimensionais, permitindo-lhe desempenhar papéis duplos semelhantes a RNAs e a proteínas. Diferentemente dos mRNAs, os lncRNAs são predominantemente localizados no núcleo, exibindo baixos níveis de expressão, além desta ser tecido-específica [Cabili *et al.*, 2011; Wang & Chang, 2011; Rutenberg-Schoenberg *et al.*, 2016]. Estima-se que o número total de lncRNAs é de aproximadamente 120.000, considerando diferentes espécies [Necsulea *et al.*, 2014].

Algumas classificações vêm sendo adotadas para classificar este grupo de RNAs. A classificação mais amplamente utilizada baseia-se na sua localização genômica, especificamente na posição do cromossomo onde ele é transcrito em relação aos genes codificadores de proteínas vizinhos, classificando-os em sete categorias. Os lncRNAs intergênicos (lincRNAs) estão localizados entre dois genes codificadores de proteínas. LncRNAs intrônicos estão localizados dentro de íntrons de genes codificadores de proteínas. LncRNAs bidirecionais estão localizados dentro de 1 kb de promotores na direção oposta em relação ao transcrito que codifica a proteína. LncRNAs senso são transcritos a partir da fita senso de genes codificadores de proteínas, podendo sobrepor-se a um ou vários íntrons e éxons. LncRNAs antisenso são transcritos a partir da fita antisenso de genes codificadores de proteínas, podendo sobrepor-se a um ou vários íntrons e éxons da sequência senso. Os lncRNAs

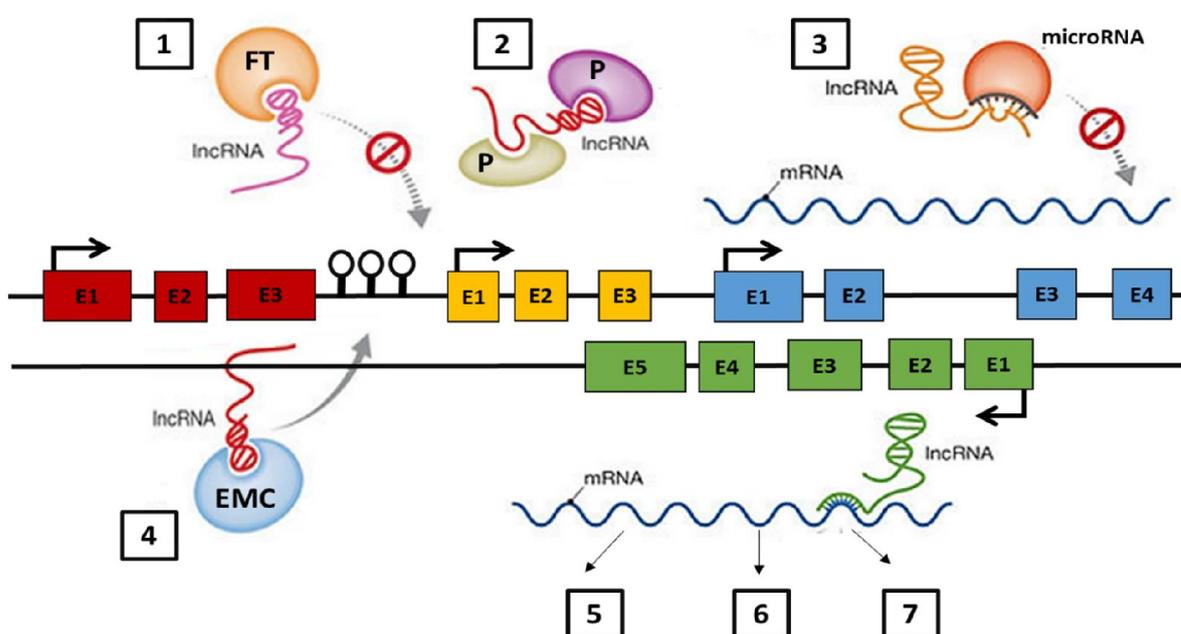
*Enhancers* (lncRNAs) são geralmente menores do que 2 kb, sendo transcritos a partir de regiões de *enhancers* do genoma. Por fim, os lncRNAs associados a promotores são transcritos a partir de regiões promotoras de genes codificadores de proteína (Figura 21) [Devaux *et al.*, 2015; Hauptman & Glavac, 2016]. Os lncRNAs também podem ser classificados, com base no efeito exercido sobre as sequências de DNA, em lncRNAs de ação cis (cis-lncRNAs) e de ação trans (trans-lncRNAs), como, também, com base no seu mecanismo de função, em: reguladores transcricionais, reguladores pós-transcricionais e outros mecanismos de funções [Ma *et al.*, 2013; Wang & Chang, 2011].



**Figura 21: Classificação mais amplamente utilizada dos RNAs longos não codificadores (lncRNAs).** Esta classificação é baseada na localização genômica da transcrição dos lncRNAs, levando em consideração os genes codificadores de proteínas vizinhos [adaptado de Hauptman & Glavač, 2016].

Por apresentarem características bastante específicas, esse grupo de ncRNAs exibem papéis como moléculas funcionais, regulando a expressão gênica em vários níveis, seja na transcrição, na tradução ou na pós-tradução. Os lncRNAs exercem suas funções por meio de vários mecanismos, tais como mediando interações intercromossômicas, servindo como iscas para fatores de transcrição (impedindo sua ligação a sequências de DNA alvo), agindo como esponjas de microRNAs endógenos (impedindo sua ligação a mRNAs alvos), interagindo com mRNAs (modulando o

*splicing*, inibindo a tradução ou levando à sua degradação), interagindo com proteínas dos complexos ribonucleoproteicos, recrutando enzimas modificadoras da cromatina que modulam componentes epigenéticos, entre outros (Figura 22) [Hu *et al.*, 2012]. Devido a isto, qualquer alteração nos níveis de expressão destas moléculas pode levar ao desenvolvimento de várias doenças, incluindo o câncer [Djebali *et al.*, 2012; Spizzo *et al.*, 2012; Huang *et al.*, 2016; Mitobe *et al.*, 2018].



**Figura 22: RNAs longos não codificadores de proteínas são considerados importantes reguladores da expressão gênica.** Os lncRNAs podem atuar por meio de diversos mecanismos: 1) servindo como iscas para fatores de transcrição; 2) interagindo com proteínas reguladoras dos complexos ribonucleoproteicos; 3) agindo como esponjas de microRNAs; 4) recrutando enzimas modificadoras da cromatina; 5) interagindo com mRNAs inibindo sua tradução; 6) interagindo com mRNAs modulando o *splicing*; e 7) interagindo com mRNAs levando a sua degradação [adaptado de Hu *et al.*, 2012]. E: éxon; FT: fatores de transcrição; P: proteínas; EMC: enzimas modificadoras da cromatina.

Diferentes lncRNAs foram descritos como sendo diferencialmente expressos na carcinogênese, com sua superexpressão ou nocauteamento podendo resultar em respostas fenotípicas alteradas relacionadas à transformação maligna, como mudanças no ciclo celular, na apoptose, no reparo do DNA, na proliferação, na angiogênese, na migração e na invasão [Su *et al.*, 2015]. Da mesma forma que os genes codificadores de proteínas, esses lncRNAs podem ser classificados como

oncogenes ou supressores tumorais, abrindo oportunidades para sua aplicação como biomarcadores, seja para o diagnóstico, prognóstico ou como possíveis alvos terapêuticos [Yarmishyn & Kurochkin, 2015].

Com o advento e aperfeiçoamento das mais diversas técnicas da Biologia Molecular, cresce o número de exemplos de lncRNAs envolvidos no câncer, alguns dos quais foram identificados no câncer de mama triplo negativo (TNBC). Em relação aos lncRNAs oncogênicos identificados neste subtipo de câncer de mama, pode-se citar como exemplos o HOTAIR, MALAT1, ARNILA, H19, TROJAN, LINC00339, MEG3, NRAD, DANCR, entre outros [Wang *et al.*, 2017; Volavat *et al.*, 2020; Zhang *et al.*, 2021]. HOTAIR é um dos lncRNAs reguladores mais estudados, tendo sido descrito por induzir a migração e a invasão de linhagens celulares de TNBC, atuando como importante marcador de metástase. Pacientes com elevada expressão do HOTAIR tendem a ter baixas taxas de sobrevivência. A detecção deste lncRNA em amostras de sangue de paciente com TNBC pode fornecer informações prognósticas e facilitar monitoramento da resposta ao tratamento [Gupta *et al.*, 2010; Wu *et al.*, 2014]. MALAT1 também é bastante estudado, tendo sido identificado com elevada expressão em TNBC, induzindo a proliferação e migração e inibindo a apoptose, *in vitro*, e induzindo o crescimento e metástase, *in vivo* [Jin *et al.*, 2016]. Estudos recentes também indicaram que MALAT1 desempenha papéis cruciais na progressão do tumor em vários subtipos de câncer de mama [Jadaliha *et al.*, 2016]. O lncRNA ARNILA foi descrito por desempenhar um papel fundamental na invasão e metástase de TNBC. Ligando-se competitivamente ao miR-204, o qual normalmente inibe a expressão de SOX4, ele pode promover a transição epitélio-mesenquimal (EMT). Sua elevada expressão sequestra o miR-204, liberando SOX4 e induzindo a EMT e,

conseqüentemente, a metástase. Além disso, foi observado que o silenciamento de ARNILA inibe a invasão, migração e metástase [Yang *et al.*, 2018].

Em contraste, lncRNAs supressores tumorais também foram identificados por desempenhar papéis importantes na progressão do TNBC, como, por exemplo, o GAS5, SONE, XIST, RMST, ZFAS1, entre outros [Hansji *et al.*, 2014; Zhang *et al.*, 2021]. GAS5 é um importante regulador do ciclo celular e apresenta baixa expressão em amostras de câncer de mama. Foi observado que sua superexpressão é responsável por inibir a proliferação celular e invasão e promover a apoptose em linhagens celulares de TNBC [Mourtada-Maarabouni *et al.*, 2009; Li *et al.*, 2018]. Além disto, foi descrito que os níveis de expressão deste lncRNA em pacientes com TNBC está associado à resistência tumoral a vários medicamentos quimioterápicos, incluindo adriamicina, paclitaxel e cisplatina, podendo, portanto, ser um potencial biomarcador para monitoramento do prognóstico de pacientes com TNBC [Li *et al.*, 2019]. O lncRNA SONE é responsável por inibir a proliferação e a migração em linhagens celulares de TNBC, sendo regulado positivamente por TP53 e negativamente por c-MYC. Ele é importante na indução de diversos miRNAs supressores tumorais, incluindo miR-34a, miR-15, miR-16 e let-7a [Youness *et al.*, 2019]. Diante desse cenário, a busca e caracterização funcional de novos lncRNAs que possam estar relacionados à progressão do câncer de mama triplo negativo, é de extrema importância, tanto na clínica médica, para o diagnóstico correto e avaliação do prognóstico, quanto para a eleição de uma possível estratégia terapêutica, justificando o desenvolvimento do presente trabalho.

## 3.2. OBJETIVOS

### 3.2.1. Objetivo Geral

O objetivo geral deste capítulo é caracterizar funcionalmente o gene não codificador de proteína *LINC01133*, em linhagens celulares humanas de câncer de mama triplo negativo, no intuito de esclarecer o papel de seu respectivo produto gênico na progressão do fenótipo tumoral, contribuindo, desta forma, para o futuro desenvolvimento de potenciais protocolos clínicos e terapêuticos relacionados ao câncer de mama triplo negativo.

### 3.2.2. Objetivos Específicos

- a) Identificar, utilizando abordagem *in silico*, possíveis novos RNAs longos não codificadores (lncRNAs) associados ao câncer de mama triplo negativo;
- b) Analisar a expressão de 10 lncRNAs selecionados, em nosso painel de linhagens celulares humanas de câncer de mama, por qRT-PCR;
- c) Selecionar um dos lncRNAs analisados, o qual se apresentou diferencialmente expresso entre as linhagens utilizadas;
- d) Analisar a função do gene não codificador do lncRNAs selecionado através da superexpressão, via sistema lentiviral indutível, na linhagem não-tumorigênica MCF10A e do nocauteamento, via sistema CRISPR/Cas9, na linhagem de câncer de mama triplo negativo Hs578T;
- e) Caracterizar as linhagens celulares geradas, *in vitro*, quanto à proliferação, ao crescimento de colônias dependente/independente de ancoragem e ao potencial migratório e invasivo;
- f) Analisar, nas linhagens celulares geradas, possíveis alterações morfológicas através de reações de imunofluorescência;

- g) Analisar, nas linhagens celulares geradas, a expressão de genes envolvidos na malignidade tumoral;

### 3.3. MATERIAIS E MÉTODOS

#### 3.3.1. Linhagens Celulares e Condições de Cultivo

Conforme descrito no Capítulo I (item 2.3.1).

#### 3.3.2. Identificação *in silico* de lncRNAs associados ao câncer de mama

Baseado em dados de artigos publicados na literatura científica e com auxílio do website *UCSC Genome Browser* [<https://genome.ucsc.edu/>], cerca de dez lncRNAs foram selecionados para posterior análise de expressão gênica no painel de linhagens de mama citado no tópico anterior. Foram estabelecidos critérios para a escolha destes dez alvos: na literatura eles devem possuir associação com qualquer tipo de tumor sólido; podem ser oncogênicos ou supressores tumorais; devem ser classificados como RNAs longos não codificadores intergênicos, facilitando a realização do CRISPR/Cas9; e por fim, devem apresentar comprimento menor do que 3.5 kb.

#### 3.3.3. Ensaio de PCR quantitativo em tempo real (qRT-PCR)

##### 3.3.3.1. Extração do RNA total

Conforme descrito no Capítulo I (item 2.3.2.1).

##### 3.3.3.2. Desenho dos primers

Os *primers* utilizados para amplificação dos genes nos experimentos de qRT-PCR foram desenhados com o auxílio do programa computacional *Primers Express 3.0* (Thermo Scientific) (Tabela 5). Para se obter um bom desempenho nas reações de qRT-PCR, eles foram desenhados com as seguintes características: amplificação de fragmentos de 90 à 150pb, conteúdo de CG próximos entre os *primers forward* e *reverse* (entre 30 e 80%), incapacidade de formação de dímeros ou de estruturas secundárias e apresentar temperatura de anelamento acima de 60°C. Além disso, a

fim de evitar uma eventual coamplificação de DNA genômico contaminante, a maior parte dos pares de primers foram desenhados em éxons diferentes do gene.

**Tabela 5: Sequência de *primers* utilizados nos ensaios de qRT-PCR.**

<b>Genes</b>	<b>Sequência dos <i>Primers</i> (5' → 3')</b>
<b>BANCR</b>	F: AGCTGACAGGGTCCATCCA R: GAGCAGCAGCCAGAGATCA
<b>CCAT1</b>	F: GAAAGGTGCCGAGACATGAA R: GGAGAAAACGCTTAGCCATACA
<b>CYTOR</b>	F: AGCCTCCATCCACATTCCA R: GGGGCTGAGTCGTGATTTT
<b>FOXCUT</b>	F: CGGATTGTTAATGCAAGACGA R: GGAATGCGGGAAGTGAATCT
<b>GHET1</b>	F: CAACAAGGATCGGGCTGTT R: CCCTTGCCTGCACCTTTAC
<b>HOTAIRM1</b>	F: ACATCGCGTTGTCATTGGA R: CCAGCTGTTGATGGGTTCA
<b>HULC</b>	F: TCTGCAAGCCAGGAAGAGTC R: TTCCGGCCTTTACTTCAGAGT
<b>LINC01133</b>	F: CCACAGCCTGGCTCTACCT R: TTGGGGTCTGTCCTTGTT
<b>MA-LINC1</b>	F: TTGCGAAACAGTTCAACTTCC R: AAGCCAGGAAAGAAGCCAAG
<b>TERC</b>	F: GCGAAGAGTTGGGCTCTGT R: GGACTCGCTCCGTTCCCTCT
<b>GAPDH</b>	F: TGCACCACCAACTGCTTAGC R: GGCATGGACTGTGGTCATGAG
<b>HPRT1</b>	F: TGAGGATTTGGAAAAGGGTCT R: GAGCACACAGAGGGCTACAA
<b>HMBS</b>	F: TGGACCTGGTTGTTCACTCCTT R: CAACAGCATCATGAGGGTTTT

### 3.3.3.3. Transcrição reversa e síntese de cDNA

Conforme descrito no Capítulo I (item 2.3.2.3), com as seguintes modificações: para a transcrição reversa e síntese de cDNA fita simples foi utilizado 1,5 µg de RNA total de cada condição experimental. Cada amostra foi transcrita utilizando-se a enzima *SuperScript III* (Invitrogen), utilizando-se oligonucleotídeos iniciadores oligo-dT(20), segundo as recomendações do fabricante, com incubação à 50°C. Uma pré-

incubação de 3 min a 85°C foi introduzida no protocolo, antes da fase de anelamento, para desnaturar completamente o RNA. Após o término da fase de anelamento, a reação foi mantida à 50°C, para a adição da enzima e do tampão pré-aquecidos à 50°C, eliminando-se a etapa de resfriamento descrita no protocolo do fabricante. Os cDNAs sintetizados foram então utilizados como *template* para as reações de qRT-PCR.

#### 3.3.3.4. Determinação da concentração final de primers

Conforme descrito no Capítulo I (item 2.3.2.4).

#### 3.3.3.5. Determinação da eficiência dos *primers*

Conforme descrito no Capítulo I (item 2.3.2.5).

#### 3.3.3.6. Confirmação da expressão diferencial

Conforme descrito no Capítulo I (item 2.3.2.6).

#### 3.3.4. Seleção de RNA longo não codificador para análise funcional

Baseado nos dados de expressão obtidos a partir do *screening* feito por qRT-PCR, um RNA longo não codificador foi selecionado para análise funcional. O alvo escolhido foi o LINC01133.

#### 3.3.5. Construção do vetor lentiviral pLVX-TetOne-Puro-LINC01133 para ensaio de superexpressão

Para a superexpressão do gene *LINC01133*, utilizamos um sistema lentiviral de superexpressão induzível para células de mamífero – o *Lenti-X™ Tet-One™ Inducible Expression Systems* (Clontech® Laboratories, Inc.), o qual é baseado no plasmídeo pLVX-TetOne-Puro Vector. Este plasmídeo expressa a proteína transativadora Tet-On®3G, contém um gene de interesse sobre controle rígido do promotor TRE3G (PTRE3GS) e possui um cassete de resistência a puromicina que permiti selecionar

as células que incorporaram o plasmídeo. Usando este sistema, a célula alvo irá expressar altos níveis do gene de interesse, mas apenas quando cultivada na presença de Doxiciclina (Dox) (Sigma-Aldrich). Os vetores lentivirais foram construídos com base na sequência completa do cDNA correspondente ao *LINC01133*, o qual foi clonado nesse sistema de maneira sítio-dirigida para que apenas a fita de interesse seja expressa. A superexpressão do *LINC01133* foi feita em células MCF10A, linhagem na qual esse gene apresenta baixa expressão.

#### 3.3.5.1. Síntese da região codificadora do *LINC01133*

A síntese da região codificadora do *LINC01133* foi feita pela empresa GenScript USA Inc. O gene foi clonado no vetor pUC57-Simple utilizando a enzima de restrição *EcoRV*. Nas extremidades do cDNA correspondente ao *LINC01133* foram adicionadas as sequências das enzimas de restrição *EcoRI* (GAATTC) e *AgeI* (ACCGGT), importantes para a inserção do gene de interesse nos respectivos sítios do vetor lentiviral de superexpressão.

#### 3.3.5.2. Transformação por eletroporação de bactérias *Escherichia coli* eletrocompetentes com DNA plasmideal

Conforme descrito no Capítulo I (item 2.3.3.4), com as seguintes modificações: bactérias *E. coli* XL10-Gold eletrocompetentes foram transformadas por eletroporação com a construção plasmideal de interesse descrita acima. Uma alíquota de 2µl do plasmídeo foi adicionada a 50µl de bactérias.

#### 3.3.5.3. Preparação de DNA plasmideal em pequena escala

Conforme descrito no Capítulo I (item 2.3.3.5).

#### 3.3.5.4. Análise da sequência de nucleotídeos

Conforme descrito no Capítulo I (item 2.3.3.6).

### 3.3.5.5. Sub-clonagem da construção no vetor lentiviral de quarta geração

Uma das construções contendo a sequência completa de nucleotídeos do gene *LINC01133* foi utilizada para a subclonagem nos respectivos sítios de restrição do vetor lentiviral induzível pLVX-TetOne-puro e esta construção foi então utilizada no ensaio funcional de superexpressão.

### 3.3.6. Geração da linhagem MCF10A superexpressando o gene *LINC01133*

#### 3.3.6.1. Produção de partículas lentivirais em células 293T

Conforme descrito no Capítulo I (item 2.3.4.1), com as seguintes modificações: para a produção dos lentivírus foram utilizadas as seguintes proporções de DNAs plasmideais: 3µg do vetor de transferência, 2µg dos vetores pREV, pTAT e pHGPM2 e 4µg do vetor pVSVG.

#### 3.3.6.2. Titulação das partículas lentivirais resistentes a puromicina

As partículas virais produzidas foram tituladas em células da linhagem 293T conforme descrito por Tiscornia e colaboradores [Tiscornia et al., 2006]. Foram preparadas diluições  $10^{-1}$ ,  $10^{-2}$  e  $10^{-3}$  dos meios condicionados contendo as partículas virais. Para realização da transdução viral, foi misturado 60µL dos meios condicionados originais ou das diluições em 1,440mL de uma suspensão celular contendo  $1 \times 10^5$  células 293T em DMEM suplementado com 10% de SFB (Hyclone/Gibco). Imediatamente após a mistura, as células foram transferidas para poços de uma placa de 6 poços e incubadas por duas semanas até o momento da contagem do número de colônias. No dia seguinte ao plaqueamento, foram adicionados 2µg/mL de puromicina (Sigma-Aldrich) a cada poço da placa. Para contagem das colônias, as mesmas foram coradas com cristal violeta 0,5%.

O título de cada preparação lentiviral foi obtido a partir da contagem do número

de colônias na maior diluição em que foi observado o crescimento das mesmas, utilizando-se a seguinte relação:

$$\text{Título (UFC)} = (P \times N / 100 \times V) \times /DF, \text{ onde:}$$

P = nº de colônias resistentes a puromicina

N = número de células no momento da transdução ( $10^5$ )

V = volume da diluição usado para a transdução

DF = fator de diluição

### 3.3.6.3. Transdução viral da linhagem MCF10A com partículas lentivirais

As células da linhagem MCF10A foram transduzidas com a preparação lentiviral correspondente ao pLVX-TetOne-Puro/LINC01133 e ao controle pLVX-TetOne-Puro. Para a transdução, cada construção lentiviral foi misturada com uma suspensão de  $5 \times 10^3$  células para se obter uma multiplicidade de infecção de 10 (MOI: Multiplicity of Infection). As células foram, então, plaqueadas em placas de 48 poços e espinoculadas por 2 horas a 28°C e 1.200rpm. Após isto, as células foram incubadas à temperatura de 37°C em atmosfera de 5% de CO<sub>2</sub>. O meio de cultura foi trocado após 48h da transdução e ao mesmo adicionado 2µg/mL de puromicina para início da seleção.

### 3.3.7. Ensaio de metabolismo celular através da redução do sal tetrazólico (MTT)

Com o objetivo de avaliar o metabolismo celular da linhagem MCF10A e sua capacidade redutora frente a presença da Dox, o ensaio de MTT foi realizado. Em uma placa de 96 poços (P96) foram plaqueadas  $1 \times 10^4$  células/poço usando 0,2mL/poço de meio de cultura específico para a linhagem celular. Foram testadas as seguintes concentrações de Dox – 0, 100, 250, 500, 750 e 1000 ng/mL, usando-se seis réplicas técnicas para cada concentração. Três experimentos independentes

foram realizados, diferindo entre si no tempo de ação da doxíclicina frente à linhagem celular MCF10A (48, 96 e 240 horas). Após transcorrido o tempo específico para cada experimento, foram adicionados 10µL da solução 10mg/mL de MTT (Sigma-Aldrich) a cada poço das placas de 96 poços e as mesmas foram protegidas com papel alumínio e incubadas à temperatura de 37°C em atmosfera de 5% de CO<sub>2</sub>. Após 2 horas de incubação, o meio de cultura contendo MTT foi removido das placas por inversão e foram adicionados 100 µL de DMSO em cada poço. As placas foram incubadas à temperatura ambiente e protegidas da luz por cerca de 1 hora e, após isto, foram quantificadas no espectrofotômetro *SpectraMax Paradigm* (Molecular Devices), no comprimento de onda de 600nm. Os resultados da ativação da mitose celular foram expressos convertendo-se os dados de absorbância (ABS) para “atividade metabólica”. Para isso, atribui-se a média das ABS das culturas não-tratadas com Dox o valor arbitrário de 1.

### 3.3.8. Confirmação da superexpressão indutível do *LINC01133* nas linhagens transduzidas

Para avaliar o nível de mRNA do *LINC01133* nas linhagens parentais e transduzidas, foram realizadas reações de qRT-PCR. Para confirmação inicial da superexpressão, a linhagem MCF10A que foi transduzida com o plasmídeo pLVX-TetOne-Puro/*LINC01133* foi cultivada em duas condições: A) Controle: linhagem cultivada sem a presença da Dox; B) Indutível: linhagem cultivada na presença de 500ng/mL de Dox, durante 48 horas. Após a confirmação da superexpressão, com o objetivo de determinar a concentração mínima de Dox necessária para se obter altos índices de superexpressão indutível, uma curva de titulação foi realizada, utilizando-se as mesmas concentrações de Dox utilizadas no tópico anterior - 0, 100, 250, 500, 750 e 1000 ng/mL.

### 3.3.9. Construção dos vetores lentivirais do sistema *CRISPR/Cas9* para ensaio de nocauteamento do *LINC01133*

Para o bloqueio total da expressão do gene *LINC01133* na linhagem celular de interesse, foi realizada a deleção completa do locus deste gene, utilizando o sistema lentiviral de nocauteamento *CRISPR/Cas9*, com os plasmídeos lentivirais pL-CRISPR.EF.GFP (Addgene #57818), o qual expressa o gene repórter GFP, e p-lentiCRISPR-V2 (Addgene #52961), o qual expressa o gene de resistência a puromicina). Ambos os plasmídeos co-expressam a enzima Cas9 e o *guide RNA* (gRNA) em *cis*. Esse sistema já se encontra estabelecido no grupo, e baseia-se na edição genômica de regiões específicas ou do *locus* completo do gene de interesse através da atividade nucleásica da Cas9 e de um *guide RNA*. No caso do gene *LINC01133* o nocauteamento foi realizado em células Hs578T, linhagem onde seus transcritos são muito abundantes.

#### 3.3.9.1. Seleção da região alvo e desenho dos *guides RNAs*

Para a deleção completa do locus do gene, as regiões escolhidas foram os 250 nucleotídeos iniciais da extremidade 5' e os 250 nucleotídeos finais da extremidade 3' do gene *LINC01133*. O desenho dos *guides RNAs* foi feito utilizando o programa MIT disponível no site <http://crispr.mit.edu/>. Este programa fornece as melhores sequências de gRNAs presentes nas regiões anteriormente selecionadas, baseando-se no sítio PAM. As sequências com maiores scores foram escolhidas por apresentarem uma menor quantidade de sítios *off-targets*. Diferentes conjuntos de gRNAs foram desenhados para serem testados (Tabela 6). Aos *primers* que compõem os gRNAs foram adicionados pares de bases adicionais em suas extremidades necessárias para criar extremidades coesivas, as quais serão importantes para a ligação dos gRNAs nos plasmídeos descritos no tópico anterior.

**Tabela 6: Sequência dos *primers up* e *down* utilizados no nocauteamento do *LINC01133*.**

Extremidade 5'		<i>Primer Up</i>	<i>Primer Down</i>
<i>Guide RNA 1</i>	CACCGTCAACACACATTGCCTCGGG	AAACCCCGAGGCAATGTGTGTTGAC	
<i>Guide RNA 2</i>	CACCGGTACCATGGCTCTGCCGTGT	AAACACACGGCAGAGCCATGGTACC	
Extremidade 3'		<i>Primer Up</i>	<i>Primer Down</i>
<i>Guide RNA 3</i>	CACCGCAACTCAAACCGTACCACCT	AAACAGGTGGTACGGTTTGAGTTGC	
<i>Guide RNA 4</i>	CACCGATAAGGCTACCTAGGTGGTA	AAACTACCACCTAGGTAGCCTTATC	

### 3.3.9.2. Fosforilação, anelamento e diluição dos *guides RNAs*

Para a reação de anelamento, foram utilizados os seguintes reagentes: 1µL do primer *gRNA up* (100 µM), 1µL do primer *gRNA down* (100 µM), 1µL do *T4 Ligation Buffer 10x* (NEB), 1µL da enzima *Polynucleotide Kinase* (Thermo) e 6 µL de H<sub>2</sub>O. A fosforilação e o anelamento dos *gRNAs* foi feita em um termociclador utilizando as seguintes condições: 37°C por 30min, 95°C por 5min, *ramp down* para 25°C com a taxa de 5°C por min<sup>-1</sup>. Após o anelamento os *gRNAs* duplex foram diluídos na proporção de 1:100 (1µl dos *gRNAs* duplex em 99µl de H<sub>2</sub>O).

### 3.3.9.3. Digestão e desfosforilação dos plasmídeos lentivirais do sistema *CRISPR/Cas9*

Para a reação de digestão, foram utilizados os seguintes reagentes: cerca 10µg de cada um dos plasmídeos lentivirais selecionados (pL-CRISPR.EF.GFP e p-lentiCRISPR-V2), 1,5µL da enzima *BsmBI* (Esp3I) 10U/ µL (Thermo Fisher Scientific), 5µL do *Buffer Tango 10x* (Thermo Fisher Scientific), acrescido de XµL de H<sub>2</sub>O necessários para completar 50µL de volume final da reação. Para cada vetor foi feita uma reação separadamente, e após isso cada uma foi mantida a 37°C overnight no termociclador. Após o termino da reação de digestão, foram adicionados a mesma reação 6µL de *Alkaline Phosphatase Buffer 10x FastAP* (Thermo Fisher Scientific), 10µL *Thermosensitive Alkaline Phosphatase FastAP* e 14µL de H<sub>2</sub>O, obtendo assim

80µL de volume final da reação. As reações foram incubadas a temperatura de 37°C por 30min em banho seco e após transcorrido este tempo a enzima foi inativada por 5min a 75°C. A digestão total dos plasmídeos foi analisada em gel de agarose 1%, corados com Brometo de Etídio (0,5µg/mL) e visualizado sob luz ultravioleta (UV) no fotodocumentador ImageQuant LAS 4000 (GE Healthcare). Os plasmídeos digeridos foram quantificados no espectrofotômetro (*Nanodrop Spectrophotometer ND-100*, Thermo Scientific, Wilmington, USA).

#### 3.3.9.4. Reação de ligação dos gRNAs aos vetores lentivirais

Para a reação de ligação, foram utilizados os seguintes reagentes: 200ng de cada plasmídeo digerido e desfosforilado em reações separadas, 4µL dos gRNAs duplex diluídos, 1µL do T4 DNA Ligase Buffer 10X (NEB), 1µL da T4 DNA Ligase, acrescido de XµL de H<sub>2</sub>O necessários para completar 10µL de volume final da reação. As reações foram incubadas a temperatura de 16°C overnight em termociclador e após transcorrido este tempo a enzima foi inativada por 10min a 65°C. A reação foi precipitada e ressuspensa em 10µL de H<sub>2</sub>O.

#### 3.3.9.5. Transformação por eletroporação de bactérias *Escherichia coli* eletrocompetentes com DNA plasmideal

Conforme descrito no Capítulo I (item 2.3.3.4), com as seguintes modificações: bactérias *E. coli* XL10-Gold eletrocompetentes foram transformadas por eletroporação com as construções plasmideais de interesse como descrito a seguir: uma alíquota de 2µl da reação de ligação foi adicionada a 50µl de bactérias.

#### 3.3.9.6. Preparação de DNA plasmideal em pequena escala

Conforme descrito no Capítulo I (item 2.3.3.5), com as seguintes modificações: as sequências integras do gRNAs foram extraídas dos vetores pL-CRISPR-EFS\_GFP e pL-lentiCRISPR-V2\_Puro através da digestão com as enzimas *NotI* / *BamHI* e *NotI*

/ *EcoRI* (respectivamente) e, em seguida, visualizados em gel de agarose 1% para confirmação da clonagem. Não foi necessário purificar os fragmentos do gel de agarose.

#### 3.3.9.7. Análise da sequência de nucleotídeos

Conforme descrito no Capítulo I (item 2.3.3.6).

#### 3.3.10. Geração da linhagem Hs578T nocauteada do gene *LINC01133*

##### 3.3.10.1. Produção de partículas lentivirais em células 293T

Conforme descrito no Capítulo I (item 2.3.4.1), com as seguintes modificações: para a produção dos lentivírus foram utilizadas as seguintes proporções de DNAs plasmideais: 3µg do vetor de transferência, 2µg dos vetores pREV, pTAT e pHGPM2 e 4µg do vetor pVSVG.

##### 3.3.10.2. Titulação das partículas lentivirais GFP positivas

As partículas virais produzidas foram tituladas em células da linhagem 293T conforme descrito por Tiscornia e colaboradores [Tiscornia *et al.*, 2006]. Foram preparadas diluições  $10^{-1}$  e  $10^{-2}$  dos meios condicionados contendo as partículas virais. Para realização da transdução viral, foi misturado 20µL dos meios condicionados originais ou das diluições e 480µL de uma suspensão celular contendo  $1 \times 10^5$  células 293T em DMEM suplementado com 10% de SFB. Imediatamente após a mistura, as células foram transferidas para poços de uma placa de 24 poços e incubadas por 48-72h, até o momento da contagem das células fluorescentes. Para contagem, estas células foram dissociadas e a porcentagem de células GFP positivas foi analisada por citometria de fluxo.

O título de cada preparação viral foi obtido a partir da contagem da proporção de células fluorescentes, utilizando-se a seguinte relação:

Título (UFC) = (P x N / 100 x V) x /DF, onde:

P = % de células GFP<sup>+</sup>

N = número de células no momento da transdução (10<sup>5</sup>)

V = volume da diluição usado para a transdução

DF = fator de diluição

### 3.3.10.3. Titulação das partículas lentivirais resistentes a puromicina

As partículas virais produzidas foram tituladas em células da linhagem 293T conforme descrito por Tiscornia e colaboradores [Tiscornia *et al.*, 2006]. Foram preparadas diluições 10<sup>-1</sup>, 10<sup>-2</sup> e 10<sup>-3</sup> dos meios condicionados contendo as partículas virais. Para realização da transdução viral, foi misturado 60µL dos meios condicionados originais ou das diluições e 1,440mL de uma suspensão celular contendo 1x10<sup>5</sup> células 293T em DMEM suplementado com 10% de SFB. Imediatamente após a mistura, as células foram transferidas para poços de uma placa de 6 poços e incubadas por duas semanas até o momento da contagem do número de colônias. No dia seguinte ao plaqueamento, foram adicionados 2µg/mL de puromicina a cada poço da placa. Para contagem das colônias, as mesmas foram coradas com cristal violeta 0,5%.

O título de cada preparação lentiviral foi obtido a partir da contagem do número de colônias na maior diluição em que foi observado o crescimento das mesmas, utilizando-se a seguinte relação:

Título (UFC) = (P x N / 100 x V) x /DF, onde:

P = nº de colônias resistentes a puromicina

N = número de células no momento da transdução (10<sup>5</sup>)

V = volume da diluição usado para a transdução

DF = fator de diluição

#### 3.3.10.4. Transdução viral da linhagem Hs578T com partículas lentivirais

As células da linhagem Hs578T foram co-transduzidas em uma proporção de 1:1 com diferentes combinações das preparações lentivirais geradas direcionadas às extremidades 5' e 3' do locus do gene *LINC01133* (uma construção GFP positiva + uma construção resistente à puromicina). Para a co-transdução, cada construção lentiviral foi misturada com uma suspensão de  $5 \times 10^3$  células para se obter uma multiplicidade de infecção de 10 (MOI: Multiplicity of Infection). As células foram, então, inoculadas em placas de 48 poços e espinoculadas por 2h à 28°C e 1.200rpm. O meio de cultura foi trocado após 48h da transdução e ao mesmo adicionado 2µg/mL de puromicina para início da seleção.

#### 3.3.11. Confirmação do nocauteamento do *LINC01133* nas linhagens transduzidas

Para verificar a eficiência de edição genômica das diferentes combinações de gRNAs (1 + 3; 1 + 4; 2 + 3 e 2 + 4), as populações de células transduzidas foram expandidas para extração do RNA total e do DNA genômico. A extração de RNA total e as reações de qRT-PCR foram realizadas seguindo o mesmo procedimento descrito no Capítulo I (item 2.3.2), com as modificações descritas neste capítulo (item 3.3.3). O DNA genômico foi extraído com o *Illustra Tissue & Cells GenomicPrep Mini Spin Kit* (GE Healthcare) e a deleção do *LINC01133* foi confirmada por PCR genômico (95°C por 4 min; 30 ciclos de 95°C de 45 seg, 60°C por 45 seg e 72°C por 2 min; 72°C por 5 min), utilizando os primers listados abaixo (Tabela 7), desenhados com o auxílio do programa computacional *Primers Express 3.0* (Thermo Scientific), os quais diferem entre si na localização genômica onde os mesmos foram desenhados (*Inner* e *Outer* – primers desenhados, respectivamente, na região interna e externa do locus do gene *LINC01133*). Para verificar a estabilidade da população de células Hs578T KO *LINC01133* 1 + 3, foram coletadas amostras de passagens sequenciais (0, 5, 10, 15,

20 e 25) para verificação do nível de expressão do lncRNA LINC01133 e da porcentagem de células GFP positivas, através reações de qRT-PCR e *sorting* no FACS (BD FACSAria II, BD Bioscience, EUA), respectivamente. Células Hs578T parental (WT) foram usadas como controle.

**Tabela 7: Sequência de *primers* utilizados no PCR genômico para confirmar o nocauteamento do *LINC01133*.**

<b><i>Primers</i></b>	<b>Sequência dos <i>Primers</i> (5' → 3')</b>
LINC01133_KO-Inner_F	AGTAGTGGAAGTGGGCATCCT
LINC01133_KO-Inner_R	GAGCTTGGGAAAGCAAAGAGT
LINC01133_KO-Outer_F	GCTTTCCTCCCTTCCTTGTC
LINC01133_KO-Outer_R	CACTTCGCTTTGTGTCTCCA
CHD7_Control_F	CTATTGAAGATACGTGTGTACCTCTGCCCTTATAGT
CHD7_Control_R	ACTGCACAATACTTAATGACCAAGATACCTTTTGAC

### 3.3.12. Caracterização das linhagens celulares geradas

#### 3.3.12.1. Curva de crescimento celular

Conforme descrito no Capítulo I (item 2.3.7.1).

#### 3.3.12.2. Ensaio de tumorigênese *in vitro*

O potencial tumorigênico, medido através do crescimento de colônias, das linhagens parentais e das linhagens superexpressando e nocauteada do *LINC01133* foi analisado utilizando-se diferentes abordagens *in vitro*.

##### 3.3.12.2.1. Ensaio de crescimento de colônias em substrato sólido

Conforme descrito no Capítulo I (item 2.3.7.2.1).

##### 3.3.12.2.2. Ensaio de crescimento de colônias em suspensão de agarose

Conforme descrito no Capítulo I (item 2.3.7.2.2).

### 3.3.12.3. Ensaios de migração celular *in vitro*

O potencial migratório das linhagens parentais e das linhagens superexpressando e nocauteada do *LINC01133* foi analisado utilizando-se diferentes abordagens *in vitro*.

#### 3.3.12.3.1. Ensaio de migração celular por *scratching* (*wound healing*)

Conforme descrito no Capítulo I (item 2.3.7.3.1).

#### 3.3.12.3.2. Ensaio de migração celular por *Transwell*<sup>TM</sup>

Conforme descrito no Capítulo I (item 2.3.7.3.2), com as seguintes modificações: foram plaqueadas  $50 \times 10^3$  células na superfície superior da membrana porosa do inserto (tamanho do poro:  $8 \mu\text{m}$ ), em  $300 \mu\text{L}$  de meio DMEM/F12 0% SFE (para o conjunto de células superexpressando o *LINC01133*) ou DMEM 0% SFB (para o conjunto de células nocauteadas do *LINC01133*), e no poço abaixo do inserto, adicionou-se  $700 \mu\text{L}$  de meio DMEM/F12 5% SFE ou DMEM 10% SFB (mesmo padrão citado acima), o qual atua como quimioatratante para as células. As células foram mantidas por 16h-18h a  $37^\circ\text{C}$  em atmosfera de 5% (para o conjunto de células superexpressando o *LINC01133*) ou 2% (para o conjunto de células nocauteadas do *LINC01133*) de  $\text{CO}_2$ , seguidas por fixação em metanol gelado por 20min e coradas com hematoxilina de Mayer por 20min.

#### 3.3.12.4. Ensaio de invasão celular *in vitro*

Conforme descrito no Capítulo I (item 2.3.7.4), com as seguintes modificações: foram plaqueadas  $100 \times 10^3$  células na superfície superior da membrana porosa do inserto (tamanho do poro:  $8 \mu\text{m}$ ), em  $300 \mu\text{L}$  de meio DMEM/F12 0% SFE (para o conjunto de células superexpressando o *LINC01133*) ou DMEM 0% SFB (para o conjunto de células nocauteadas do *LINC01133*), e no poço abaixo do inserto,

adicionou-se 700 µL de meio DMEM/F12 5% SFE ou DMEM 10% SFB (mesmo padrão citado acima), o qual atua como quimioatraente para as células. As células foram mantidas por 16h-18h a 37°C em atmosfera de 5% (para o conjunto de células superexpressando o *LINC01133*) ou 2% (para o conjunto de células nocauteadas do *LINC01133*) de CO<sup>2</sup>, seguidas por fixação em metanol gelado por 20min e coradas com hematoxilina de Mayer por 20min.

### 3.3.12.5. Imunofluorescência

Para imunofluorescência, as células foram semeadas em lamínulas de vidro de 13 mm de diâmetro e mantidas em condições de cultura usuais até a subconfluência (menos de 80%). Cada amostra foi então fixada em formaldeído 4% por 20min, lavadas em PBSA e, então, permeabilizada com tampão IF (PBS + 0,2% Triton X-100 +0,05% Tween 20) + 0,5% Triton X-100 por 10min a 4°C e bloqueada em Albumina de Soro Bovino (BSA) 2% por 1 hora a temperatura ambiente. O anticorpo primário rabbit anti-Vimentina (AB92457, 1/75) (Abcam, Cambridge, UK) e Alexa Fluor 594-Faloidina (A12381, 1/100) (Thermo Fischer Scientific, Massachusetts, EUA) foram incubados por 1h30 a temperatura ambiente em câmara úmida. A seguir, as lamínulas foram lavadas cuidadosamente três vezes com tampão IF por 10 min. Os respectivos anticorpos secundários AlexaFluor 568 goat anti-rabbit IgG ou AlexaFluor 568 goat anti-mouse IgG (1/200) foram incubados por 1h a 4°C. As lâminas foram montadas utilizando o meio de montagem VECTASHIELD com DAPI (H-1200, Vector Laboratories, CA) e as imagens foram adquiridas utilizando o *EVOS FL Fluorescence Cell Image System* (Life Technologies) em aumento de 20x.

### 3.3.12.6. Análise da expressão de genes envolvidos na malignidade tumoral

Para entender melhor os possíveis efeitos da superexpressão e nocauteamento do *LINC01133* nos processos de angiogênese, migração e invasão celular, foi investigada a expressão de diferentes marcadores relacionados a malignidade tumoral por qRT-PCR, conforme descrito no Capítulo 1 (item 2.3.2). Os seguintes genes foram utilizados na análise: 1) Gene *XIAP* (anti-apoptótico); 2) Genes *TGF- $\beta$* , *SPARC*, *CXCL12*, *SATB1*, *E-CAD*, *N-CAD* (associados aos processos de tumorigênese e transição epitélio-mesenquimal); 3) Genes *VEGF-A* e *COL3A1* (associados à angiogênese tumoral). A sequência dos primers foram descritas no Capítulo I (item 2.3.7.5).

### 3.3.13. Análises Estatísticas

Conforme descrito no Capítulo I (item 2.3.8).

### 3.4. RESULTADOS

#### 3.4.1. Identificação *in silico* de lncRNAs associados ao câncer de mama

Em busca de possíveis novos lncRNAs associados ao câncer de mama, cerca de dez alvos foram selecionados *in silico* para posterior análise de expressão gênica no painel de linhagens celulares de câncer de mama. Todos estes 10 alvos (Tabela 8) estavam dentro dos critérios estabelecidos conforme descrito em Materiais e Métodos.

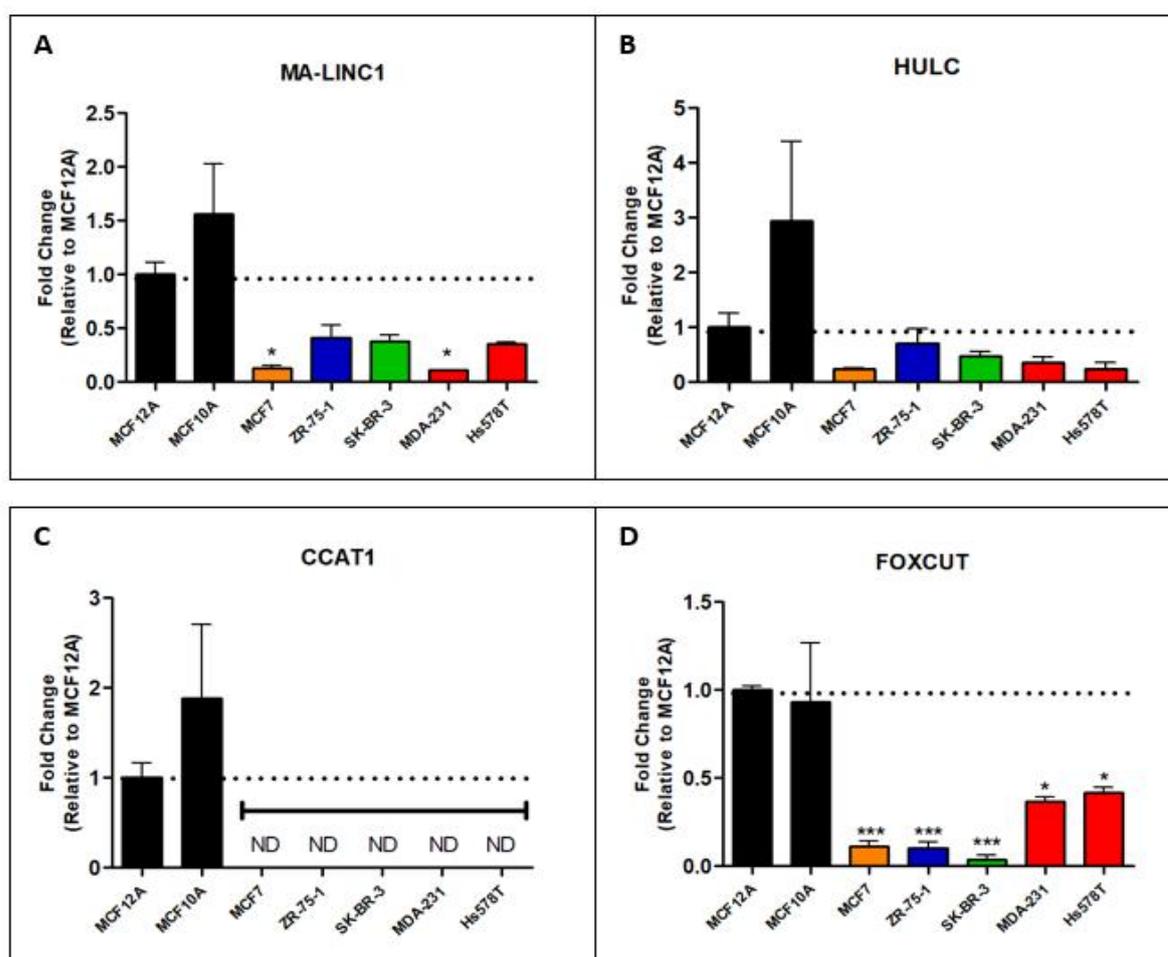
**Tabela 8: Lista dos 10 alvos escolhidos para posterior análise através de qRT-PCR.**

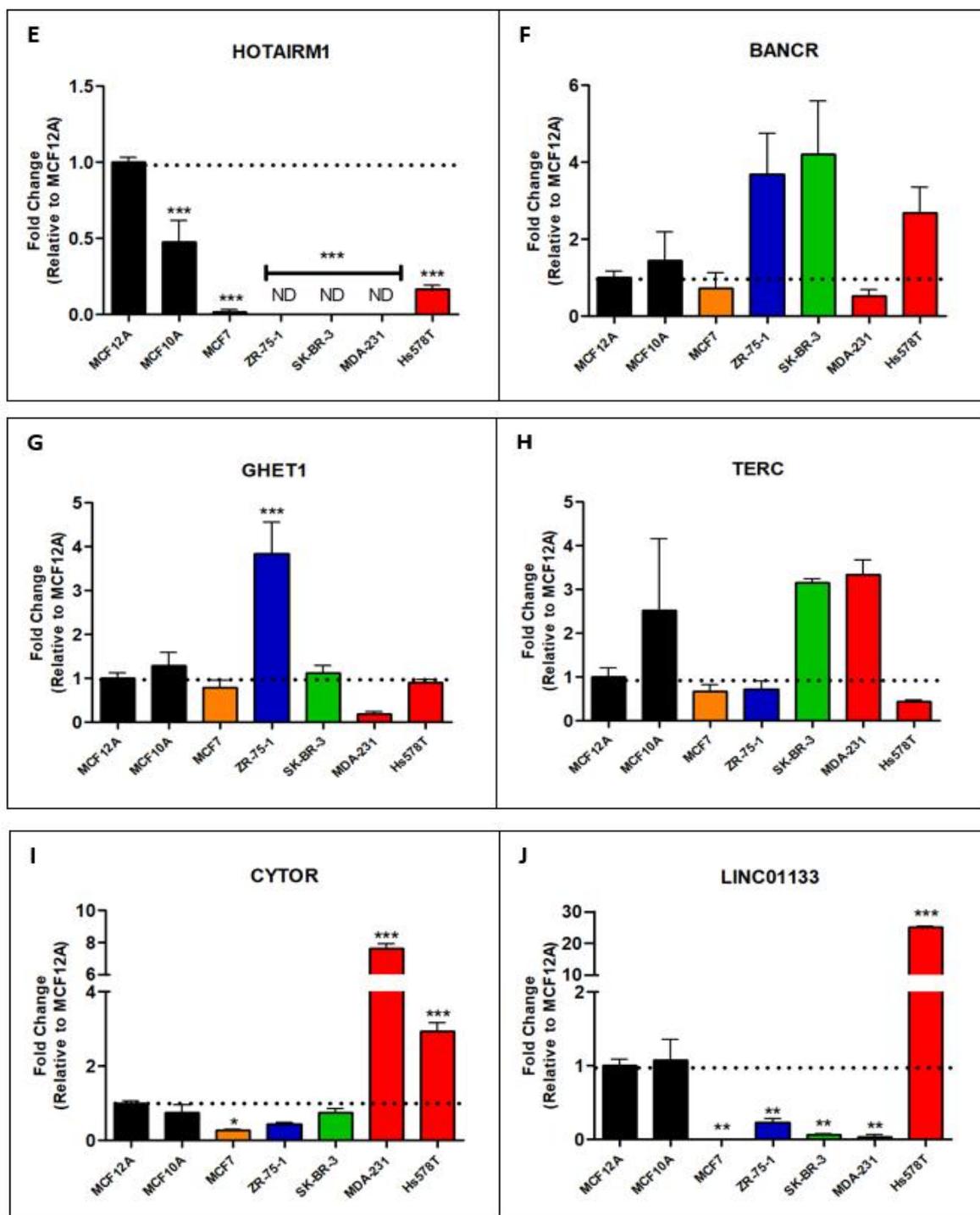
<b>LNCRNAs</b>	<b>TAMANHO (pb)</b>	<b>CLASSIFICAÇÃO</b>
<b>MA-LINC1</b>	3.138	Intergênico
<b>HULC</b>	484	Intergênico
<b>CCAT1</b>	2.795	Intergênico
<b>FOXCUT</b>	1.610	Intergênico
<b>HOTAIRM1</b>	1.084	Intergênico
<b>BANCR</b>	675	Intergênico
<b>GHET1</b>	1.898	Intergênico
<b>TERC</b>	451	Intergênico
<b>CYTOR</b>	732	Intergênico
<b>LINC01133</b>	1.138	Intergênico

#### 3.4.2. Análise da expressão gênica dos lncRNAs selecionados no modelo de progressão tumoral do carcinoma mamário

Em busca de novos marcadores que possam estar relacionados com a grau de malignidade no câncer mama, foram analisados, através da técnica de qRT-PCR, a expressão relativa do mRNA correspondente aos 10 lncRNAs listados na Tabela 8, nas linhagens celulares humanas do painel de mama.

A partir da análise do qRT-PCR, foi observado que os lncRNAs BANCR e TERC apresentaram padrões de expressão bastante variáveis entre as linhagens celulares analisadas, sendo, portanto, eliminados durante a seleção. O lncRNA GHET1 apresentou elevada expressão na linhagem tumorigênica de mama Luminal B (ZR-75-1) quando comparado com as demais linhagens. Os lncRNAs MA-LINC1, HULC, CCAT1, FOXCUT e HOTAIRM1 apresentaram expressão de nível considerável nas linhagens de mama não-tumorigênicas MCF10A e MCF12A, tornando-os alvos promissores para estudo. Por fim, os lncRNAs CYTOR e LINC01133 apresentaram expressão de nível considerável e distinto nas linhagens celulares que são modelos, *in vitro*, de tumores de mama triplo negativo (MDA-MB-231 e Hs578T), sendo considerados, também, alvos promissores a serem estudados (Figura 23).





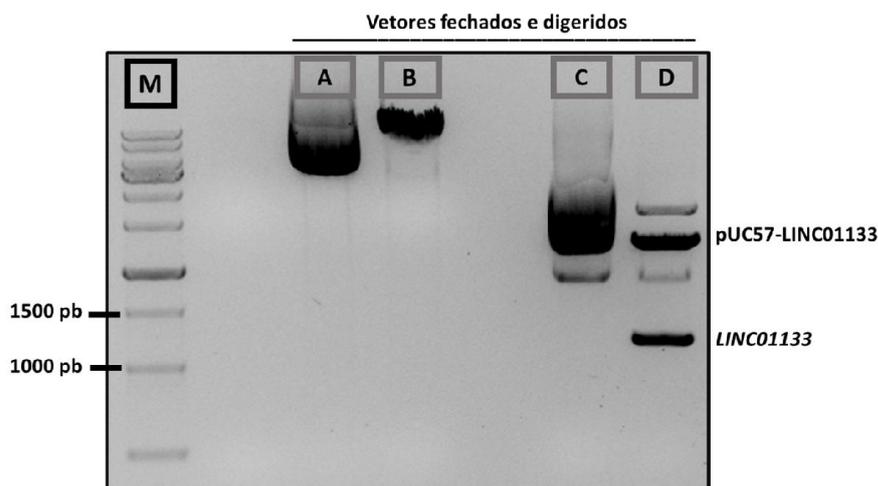
**Figura 23: Análise da expressão dos lncRNAs selecionados no painel de linhagens celulares humanas de câncer de mama.** O RNA total foi extraído conforme descrito em Materiais e Métodos e os níveis de expressão dos diferentes lncRNAs foram analisados por qRT-PCR utilizando-se *primers* específicos: A) MA-LINC1, B) HULC, C) CCAT1, D) FOXCUT, E) HOTAIRM1, F) BANCRC, G) GHET1, H) TERC, I) CYTOR, J) LINC01133. O gráfico representa a razão de fold change relativo à expressão na linhagem MCF12A ± SEM. O gene endógeno utilizado para normalização foi o *HPRT1*. As análises estatísticas foram feitas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett (\*\*\*)  $P < 0.0001$ ). Foram realizados três experimentos independentes, em triplicatas. Barras em preto – Linhagens não-tumorigênicas; Barras em laranja – Linhagem Luminal A; Barra em Azul – Linhagem Luminal B; Barra em Verde – Linhagem HER2+; Barras em vermelho – Linhagens Triplo Negativas.

### 3.4.3. Seleção de RNA longo não codificador (lncRNA) para análise funcional

Devido à ausência de tratamentos eficazes e por apresentar o prognóstico mais sombrio entre os subtipos do câncer de mama, os tumores de mama triplo negativo foram importantes para nos direcionar na seleção do lncRNA a ser estudado. Diante disso, o RNA longo não codificador LINC01133 foi selecionado por apresentar uma alta expressão na linhagem tumorigênica Hs578T (uma das linhagens do nosso painel que representam os tumores de mama triplo negativos) e baixa expressão nas linhagens não-tumorigênicas MCF12A e MCF10A.

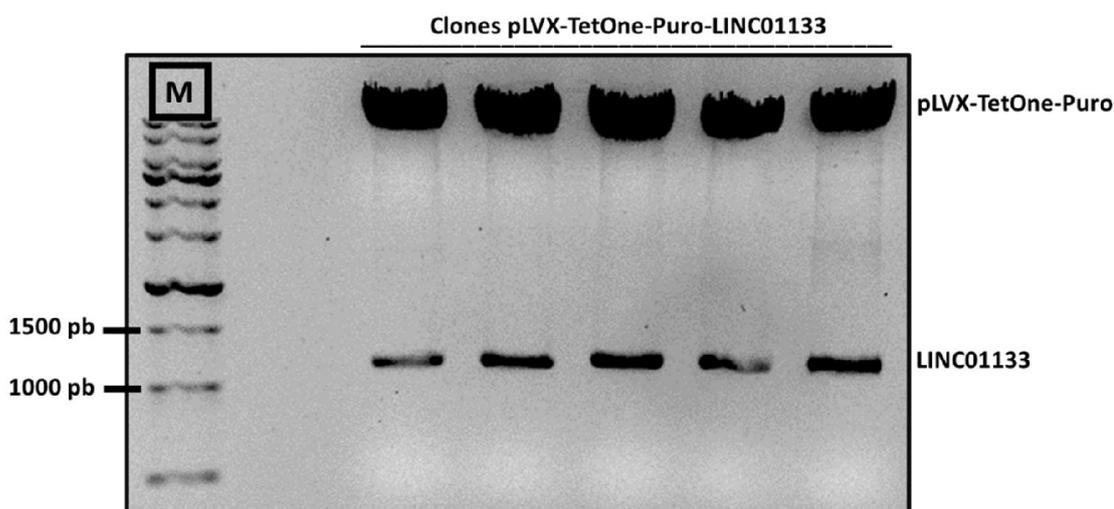
### 3.4.4. Construção do vetor lentiviral pLVX-TetOne-Puro-LINC01133 para ensaio de superexpressão do *LINC01133*

Como citado em Materiais e Métodos, o vetor pUC57-Simple contendo o gene *LINC01133* (pUC57-LINC01133) foi obtido da empresa *GenScript USA Inc.* De posse deste plasmídeo, o mesmo foi digerido com as enzimas *EcoRI* e *AgeI* e os produtos desta digestão foram fracionados em gel de agarose 1% para separação do fragmento de interesse (*LINC01133*) do vetor. O vetor pLVX-TetOne-Puro, o qual foi utilizado, posteriormente, na superexpressão indutível do *LINC01133*, também foi digerido utilizando-se as mesmas enzimas citadas anteriormente e fracionado no mesmo gel de agarose (Figura 24).



**Figura 24: Digestão dos plasmídeos pLVX-TetOne-Puro e pUC57-LINC01133 com as enzimas de restrição *EcoRI* e *AgeI*.** A) Vetor pLVX-TetOne-Puro, B) Vetor pLVX-TetOne-Puro Digerido, C) Vetor pUC57-LINC01133, D) Vetor pUC57-LINC01133 Digerido. Ambos os plasmídeos foram digeridos com as enzimas *EcoRI* e *AgeI* e, em seguida, visualizados em gel de agarose 1%. A presença do fragmento do *LINC01133* foi confirmada conforme mostrado na figura. M – Marcador de peso molecular.

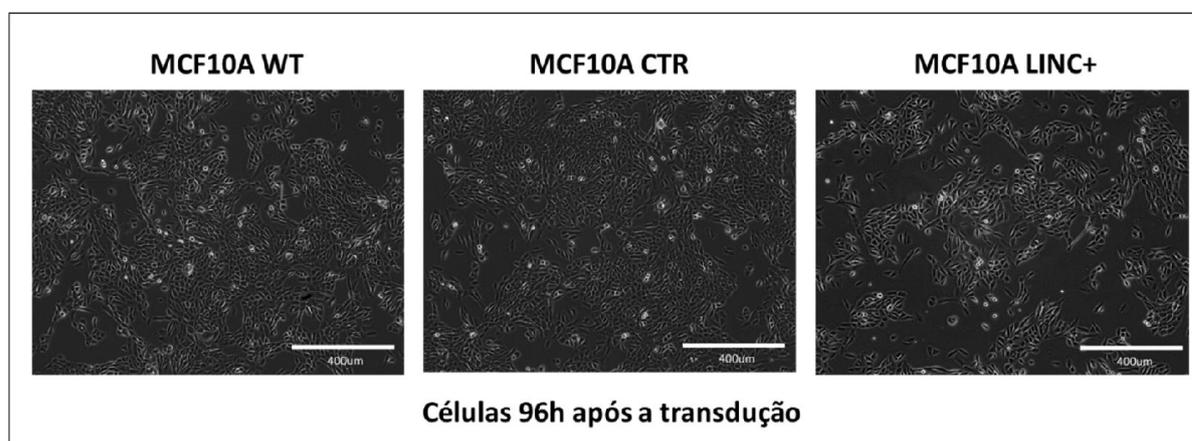
Conforme mostrado na figura anterior, a presença do fragmento do *LINC01133* no vetor pUC57-LINC01133 foi confirmada e o plasmídeo foi submetido a sequenciamento e a análises de Bioinformática, com o objetivo de confirmar a integridade da sequência do gene *LINC01133*. Após a confirmação, o fragmento do *LINC01133* foi purificado do gel e subclonado nos respectivos sítios de restrição do vetor lentiviral indutível pLVX-TetOne-Puro visando à superexpressão induzida deste gene. O DNA plasmideal obtido a partir de cinco clones correspondentes à construção pLVX-TetOne-Puro-LINC01133 foi digerido com as enzimas *EcoRI* e *AgeI* e, em seguida, fracionados em gel de agarose 1%, tendo sido detectada a presença do fragmento de interesse em todos estes clones (Figura 25).



**Figura 25: Digestão da construção pLVX-TetOne-Puro-LINC01133 para confirmação dos transfectantes.** O DNA plasmideal obtido a partir de cinco clones foi digerido com as enzimas *EcoRI* e *AgeI* e, em seguida, visualizados em gel de agarose 1% para detecção dos clones bacterianos transformantes contendo o gene de interesse. M: Marcador de peso molecular.

### 3.4.5. Geração da linhagem MCF10A superexpressando o gene *LINC01133*

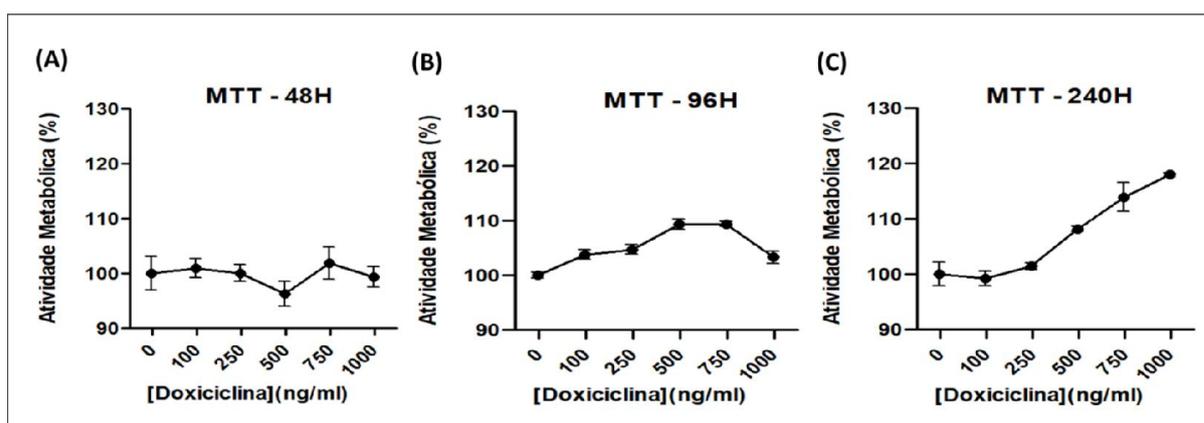
De posse de um dos clones contendo o plasmídeo recombinante, as partículas virais recombinantes foram geradas por meio de transfecção da linhagem 293T, conforme descrito em Materiais e Métodos. Além disto, células da linhagem 293T foram transduzidas para realizar o cálculo do título das partículas lentivirais. As células foram cultivadas durante duas semanas, na presença de 2µg/mL de puromicina. Observou-se que cerca de 45 e 70 colônias das células transduzidas com o vetor controle e com o vetor de superexpressão do *LINC01133* eram resistentes à puromicina, respectivamente. A partir destes números foi realizado o cálculo do título lentiviral. Para transduzir a linhagem não-tumorigênica MCF10A (10A WT) foi utilizado multiplicidade de infecção (MOI) de 10 das partículas lentivirais contendo o vetor controle pLVX-TetOne-Puro (10A CTR) e o pLVX-TetOne-Puro-LINC01133 (10A LINC+), visando, assim, obter uma linhagem controle e uma linhagem superexpressando o *LINC01133*. As células foram, então, inoculadas em placas de 48 poços e espinoculadas por 2h a 28°C e 1.200rpm. O meio de cultura foi trocado após 48h da transdução e 2µg/mL de puromicina foram adicionados para início da seleção. As células transduzidas com as preparações lentivirais citadas acima se mostraram resistentes à presença da puromicina, indicando o sucesso da transdução (Figura 26).



**Figura 26: Células MCF10A transduzidas com as construções lentivirais.** Os vetores pLVX-TetOne-Puro e o pLVX-TetOne-Puro-LINC01133 foram transduzidos em células MCF10A. Após 48h de transdução, o meio de cultura foi trocado e 2µg/mL de puomicina foram adicionados para início da seleção. Após 48h de adição de puomicina, as células se mostraram resistentes indicando o sucesso da transdução. Micrografias obtidas utilizando-se o microscópio *EVOS FL Cell Imaging Systems* (Thermo Scientific), no aumento de 100X. Barra: 400µm.

### 3.4.6. Ensaio de metabolismo celular através da redução do sal tetrazólico (MTT)

No intuito de avaliar a viabilidade celular da linhagem MCF10A parental (10A WT) frente à presença da Dox, o ensaio de MTT foi realizado. As células foram testadas frente as seguintes concentrações de Dox: 0, 100, 250, 500, 750 e 1.000 ng/mL, utilizando-se seis réplicas técnicas para cada concentração. Três experimentos independentes foram realizados, diferindo entre si no tempo de ação da Dox frente à linhagem celular MCF10A (48, 96 e 240 horas) (Figura 27).



**Figura 27: Ensaio de MTT utilizando a linhagem MCF10A cultivada na presença de diferentes concentrações de Dox, em diferentes intervalos de tempo.** (A) Ensaio de MTT no intervalo de 48h; (B) Ensaio de MTT no intervalo de 96h; (C) Ensaio de MTT no intervalo de 240h. Os dados são apresentados como a média  $\pm$  SEM de um experimento independente, realizado em triplicata.

Conforme observado, no intervalo de 48h, nenhuma concentração de Dox interferiu na atividade metabólica das células. Enquanto isto, nos intervalos de 96 e 240h, é possível observar que a partir da concentração de 500ng/mL de Dox, as células aumentam sutilmente sua atividade metabólica, indicando uma possível interferência de Dox no metabolismo destas células. Os dados deste experimento foram importantes, futuramente, para definir a concentração ideal de Dox, a qual foi

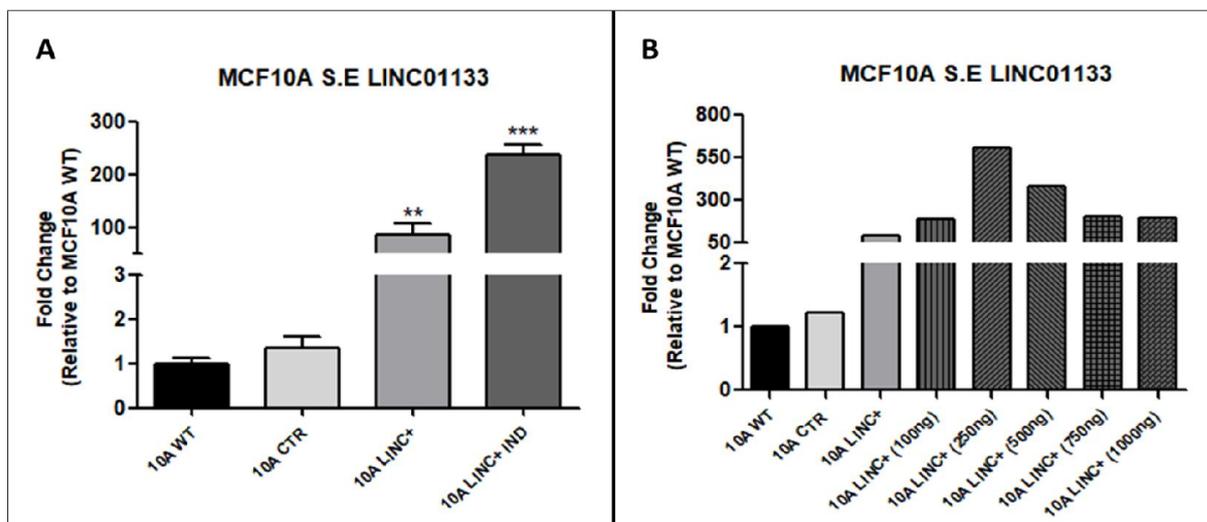
utilizada nos ensaios de superexpressão indutível do *LINC01133* na linhagem MCF10A.

### 3.4.7. Confirmação da superexpressão indutível do *LINC01133* nas linhagens transduzidas

O nível de expressão do *LINC01133* nas linhagens transduzidas foi confirmado através de reações de qRT-PCR, conforme descrito em Materiais e Métodos. Para confirmação inicial da superexpressão, a linhagem MCF10A que foi transduzida com o plasmídeo pLVX-TetOne-Puro/*LINC01133* foi cultivada em duas condições: A) Controle: linhagem cultivada na ausência da Dox (10A LINC+); B) Indutível: linhagem cultivada na presença de 500ng/mL de Dox (10A LINC+ IND), durante 48 horas. Conforme esperado, as linhagens 10A WT e 10A CTR apresentaram baixos níveis de expressão do *LINC01133*. A linhagem 10A LINC+, a qual contém o plasmídeo pLVX-TetOne-Puro/*LINC01133* e foi cultivada na ausência da Dox, apresentou um pequeno vazamento de expressão do *LINC01133*. O normal seria que estas células não apresentassem expressão do *LINC01133*, mas o vazamento apresentado é comum e já foi observado em outros estudos que utilizaram este sistema de vetores. Por fim, a linhagem 10A LINC+ IND, a qual foi cultivada inicialmente na presença de 500ng/mL de Dox, apresentou altos níveis de expressão, indicando que o objetivo foi atingido e a linhagem celular foi gerada com êxito (Figura 28 A).

Após a confirmação da superexpressão do lncRNA *LINC01133*, no intuito de determinar a concentração mínima de Dox necessária para se obter altos índices de superexpressão induzível, uma curva de titulação foi realizada, utilizando-se as concentrações de 0, 100, 250, 500, 750 e 1.000 ng/mL. Os resultados deste experimento, em conjunto com aqueles do ensaio de MTT, permitiram determinar a concentração ideal de 250ng/mL de Dox, a qual deve ser utilizada para induzir a

superexpressão do *LINC01133* nos ensaios de caracterização funcional. Esta concentração foi escolhida, pois foi a que induziu maiores níveis de expressão (Figura 28 B) e não alterou significativamente a atividade metabólica das células (Figura 27).



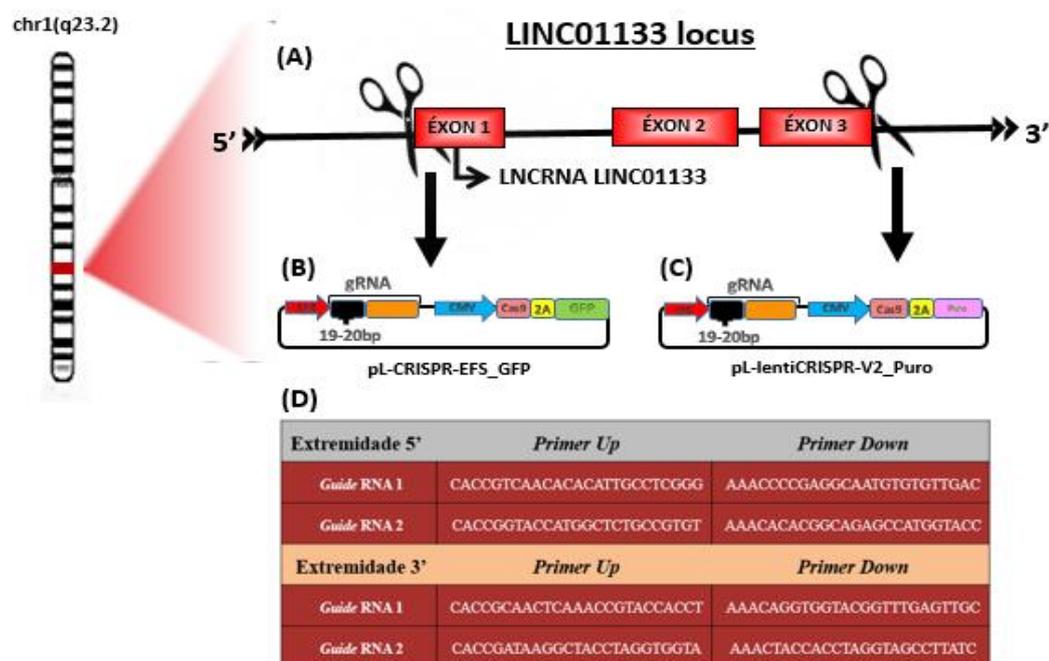
**Figura 28: Análise da expressão do *LINC01133*.** (A) Análise da expressão do *LINC01133* nas linhagens MCF10A geradas. Os níveis de expressão do *LINC01133* foram analisados por qRT-PCR, utilizando-se *primers* específicos, nas linhagens MCF10A parental (10A WT), naquelas transduzidas com o vetor vazio (10A CTR) e nas transduzidas com o vetor de superexpressão do *LINC01133*, cultivadas na ausência (10A LINC+) e na presença de Dox (10A LINC+ IND). O gráfico representa a razão de *fold change* relativo à expressão na linhagem MCF10A WT  $\pm$  SEM. O gene endógeno utilizado para normalização foi o *HPRT1*. As análises estatísticas foram feitas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett (\*\* $P < 0.0001$ ). Foram realizados três experimentos independentes, em triplicatas. (B) Análise da expressão do *LINC01133* frente à diferentes concentrações de Dox. Os níveis de expressão do *LINC01133* foram analisados nas linhagens MCF10A parental (10A WT), naquelas transduzidas com o vetor vazio (10A CTR) e nas transduzidas com o vetor de superexpressão do *LINC01133*, cultivadas na ausência (10A LINC+) e na presença de diferentes concentrações de Dox (10A LINC+ IND). O gráfico representa a razão de *fold change* relativo à expressão na linhagem MCF10A WT. O gene endógeno utilizado para normalização foi o *HPRT1*. S.E: superexpressão.

### 3.4.8. Construção dos vetores lentivirais do sistema *CRISPR/Cas9* para ensaio de nocauteamento do *LINC01133*

#### 3.4.8.1. Seleção da região alvo e desenho dos *guides* RNAs (gRNAs)

Os *guides* RNAs foram desenhados para provocar a deleção completa do locus do gene *LINC01133*. Por este lncRNA ser classificado como intergênico, ou seja, o gene responsável pela codificação do mesmo está localizado entre genes codificadores de proteína, a sua deleção não irá interferir na expressão de outros genes codificadores de proteínas. Dois gRNAs foram projetados para cada uma das

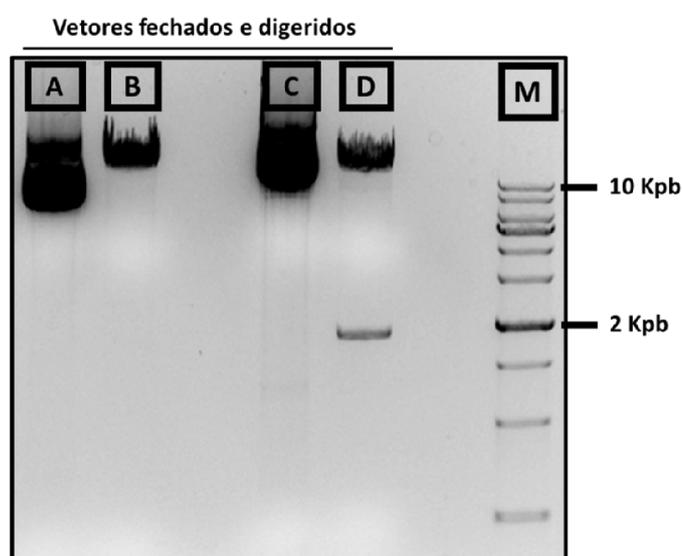
extremidades do *locus* deste gene (Figura 29 A e 29 D). Para a extremidade 5', os gRNAs foram clonados no plasmídeo pL-CRISPR-EFS\_GFP, enquanto para a extremidade 3' os gRNAs foram clonados no plasmídeo pL-lentiCRISPR-V2\_Puro. O vetor pL-CRISPR.EFS possui um gene eGFP como marcador de seleção (Figura 29 B), enquanto o vetor pL-lentiCRISPR-v2 possui um gene de resistência à puromicina (Figura 29 C). Cada vetor possui um *locus* Cas9 associado ao marcador de seleção e dois sítios de restrição para a enzima *BsmBI*, que são necessários para a inserção das sequências de gRNA. Combinando dois dos quatro gRNAs (sendo cada um direcionado a cada uma das extremidades), geramos quatro populações eGFP<sup>+</sup> e resistentes à puromicina. Portanto, as células de interesse devem apresentar os dois marcadores de seleção.



**Figura 29: Esquema do modelo de edição gênica para nocautear o *locus* do gene *LINC01133* baseado no sistema lentiviral CRISPR/Cas9.** (A) Estratégia geral de nocauteamento, (B) Cassete de expressão do plasmídeo pL-CRISPR-EFS\_GFP, (C) Cassete de expressão do plasmídeo pL-lentiCRISPR-V2\_Puro, (D) Tabela com as sequências direta e reversa de quatro gRNAs, projetadas utilizando-se a plataforma *CRISPR Optimized Design* [Feng Zhang Lab, MIT, 2015]. As duas sequências que exibiram o menor número de sítios *off-targets* foram escolhidas para desenhar dois gRNAs para cada uma das extremidades do locus *LINC01133*.

### 3.4.8.2. Digestão dos plasmídeos lentivirais e ligação dos gRNAs

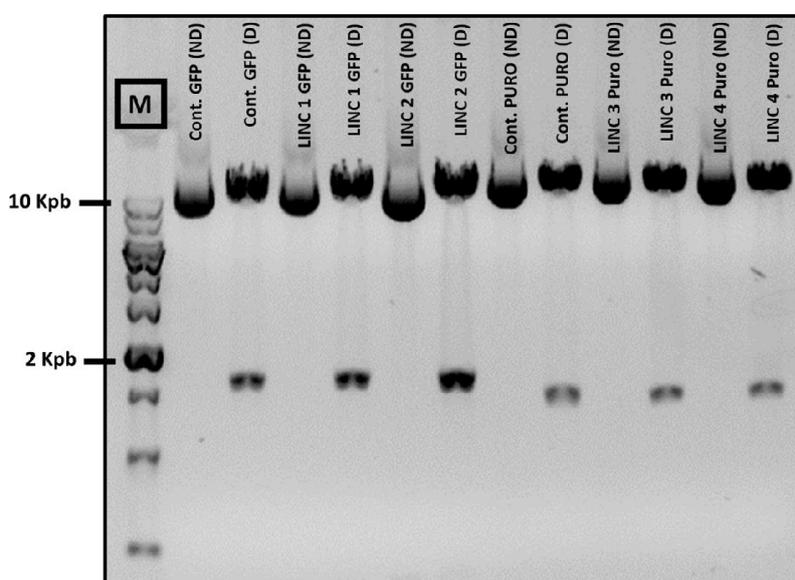
Os plasmídeos lentivirais usados para o nocaute foram digeridos com a enzima de restrição *BsmBI*, gerando dois fragmentos, uma vez que cada um possui dois sítios de restrição para a *BsmBI*. A digestão do pL-CRISPR-EFS\_GFP gera um fragmento de ~11,7 Kpb e um fragmento de ~20 pb (que não é visível no gel de agarose), confirmado através da visualização de apenas uma banda de DNA. A digestão do pL-lentiCRISPR-V2\_Puro gera fragmentos de ~12,8 Kpb e ~2,0 Kpb, confirmados através de eletroforese em gel de agarose (Figura 30).



**Figura 30: Digestão dos plasmídeos lentivirais utilizados para nocauteamento do *LINC01133*.** Os plasmídeos foram digeridos com a enzima *BsmBI* e, em seguida, visualizados em gel de agarose 1%. (A) pL-CRISPR-EFS\_GFP; (B) pL-CRISPR-EFS\_GFP digerido com a enzima de restrição *BsmBI*; (C) pL-lentiCRISPR-V2\_Puro; (D) pL-lentiCRISPR-V2\_Puro digerido com a enzima de restrição *BsmBI*. M: Marcador de peso molecular.

Quatro plasmídeos recombinantes foram construídos, após a ligação dos gRNAs aos vetores lentivirais, para se obter o nocauteamento (K.O) completo do gene *LINC01133*. Dois plasmídeos recombinantes, *LINC01133* 1#5' e o *LINC01133* 2#5', possuem a marca de seleção eGFP e foram desenhados no intuito de clivarem a extremidade 5' do locus do *LINC01133*, enquanto outros dois plasmídeos recombinantes, *LINC01133* 3#3' e o *LINC01133* 4#3', possuem um gene de

resistência à puromicina e foram desenhados no intuito de clivarem a extremidade 3' do mesmo locus gênico. As construções feitas com o plasmídeo pL-CRISPR-EFS\_GFP foram digeridas com as enzimas *NotI* e *BamHI*, enquanto as construções feitas com o plasmídeo pL-lentiCRISPR-V2\_Puro foram digeridas com as enzimas *NotI* e *EcoRI* para confirmação da clonagem (Figura 31). A sequência correta dos clones contendo cada um dos gRNAs foi confirmada por sequenciamento de DNA e análise de Bioinformática.

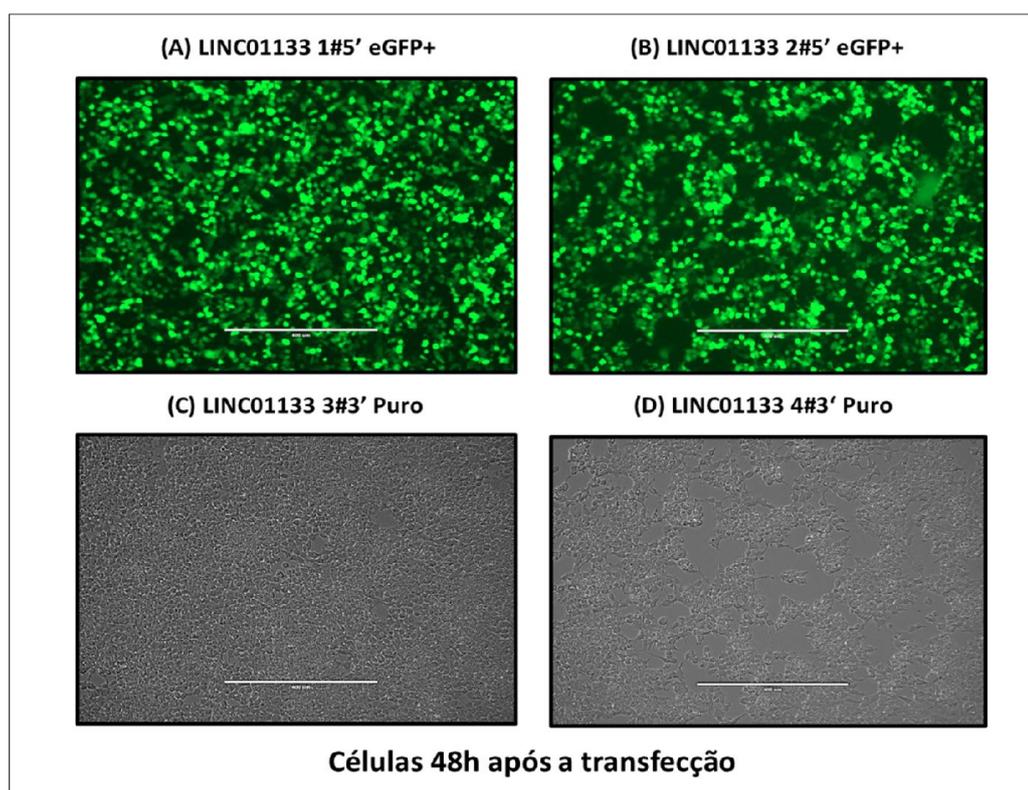


**Figura 31: Confirmação da clonagem dos gRNAs nos plasmídeos lentivirais utilizados para nocauteamento do *LINC01133*.** Os quatro plasmídeos recombinantes construídos, em conjunto com os plasmídeos controles, foram digeridos com enzimas de restrição específicas e, em seguida, visualizados em gel de agarose 1%. A presença do fragmento na faixa de 2Kpb indica o sucesso da clonagem. M: marcador de peso molecular; ND: plasmídeo não-digerido; D: plasmídeo digerido.

#### 3.4.9. Geração da linhagem Hs578T nocauteada para o gene *LINC01133*

De posse dos clones contendo os plasmídeos recombinantes, as partículas virais recombinantes foram geradas por meio de transfecção da linhagem 293T, conforme descrito em Materiais e Métodos. Além disto, células da linhagem 293T foram transduzidas para realizar o cálculo do título das partículas lentivirais (Figura 32). Para calcular o título das partículas lentivirais eGFP+, as células foram cultivadas por 48h e a porcentagem de células eGFP+ foi analisada por citometria de fluxo.

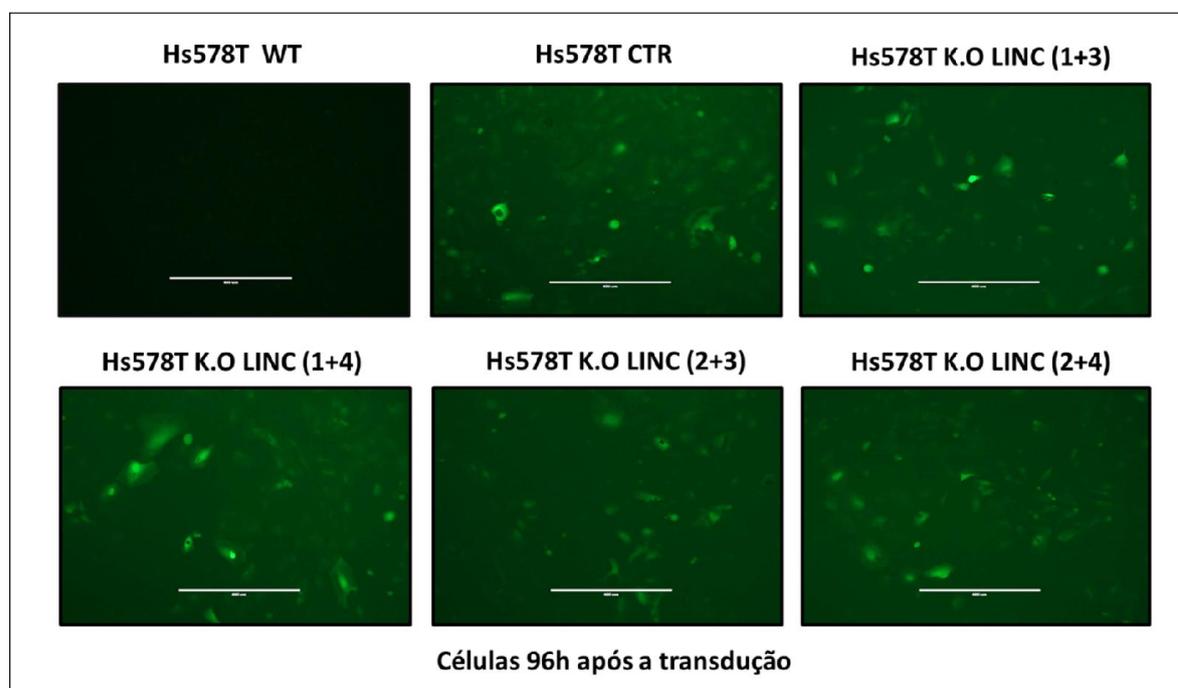
Observou-se que cerca de 12,4%, 12,8% e 41,3% das células transduzidas com o vetor controle pL-CRISPR-EFS\_GFP e com os vetores recombinantes LINC01133 1-5' e LINC01133 2-5' eram eGFP+, respectivamente. A partir destas porcentagens, foi realizado o cálculo do título lentiviral. Para calcular o título das partículas lentivirais resistentes à puromicina, as células foram cultivadas durante duas semanas, na presença de 2µg/mL de puromicina. Observou-se que cerca de 45, 40 e 86 colônias das células transduzidas com o vetor controle pL-lentiCRISPR-V2\_Puro e com os vetores recombinantes LINC01133 3-3' e LINC01133 4-3' eram resistentes à puromicina, respectivamente. A partir destes números foi realizado o cálculo do título lentiviral.



**Figura 32: Células 293T transfectadas com os vetores recombinantes para produção de partículas lentivirais.** As células 293T foram transfectadas com os plasmídeos eGFP+ contendo os gRNAs LINC01133 1-5' (A) e LINC01133 2-5' (B), e com os plasmídeos resistentes à puromicina contendo os gRNAs LINC01133 3-3' (C) e LINC01133 4-3' (D). Micrografias obtidas utilizando-se o microscópio *EVOS FL Cell Imaging Systems* (Thermo Scientific), no aumento de 100X. Barra: 400µm.

As células da linhagem Hs578T foram co-transduzidas com diferentes combinações das preparações lentivirais acima (uma construção GFP positiva + uma

construção resistente à puromicina), gerando quatro populações de células LINC01133\_K.O designadas de acordo com os gRNAs utilizados: 1 + 3, 1 + 4, 2 + 3 e 2 + 4. Para a co-transdução, cada construção lentiviral foi misturada com uma suspensão de  $5 \times 10^3$  células para se obter uma multiplicidade de infecção (MOI) de 10. As células foram, então, inoculadas em placas de 48 poços e espinoculadas por 2 h à 28°C e 1.200rpm. O meio de cultura foi trocado após 48h da transdução e 2µg/mL de puromicina foram adicionados ao mesmo para início da seleção. As células co-transduzidas com as preparações lentivirais citadas acima se mostraram resistentes à presença da puromicina e GFP+, indicando o sucesso da transdução (Figura 33).

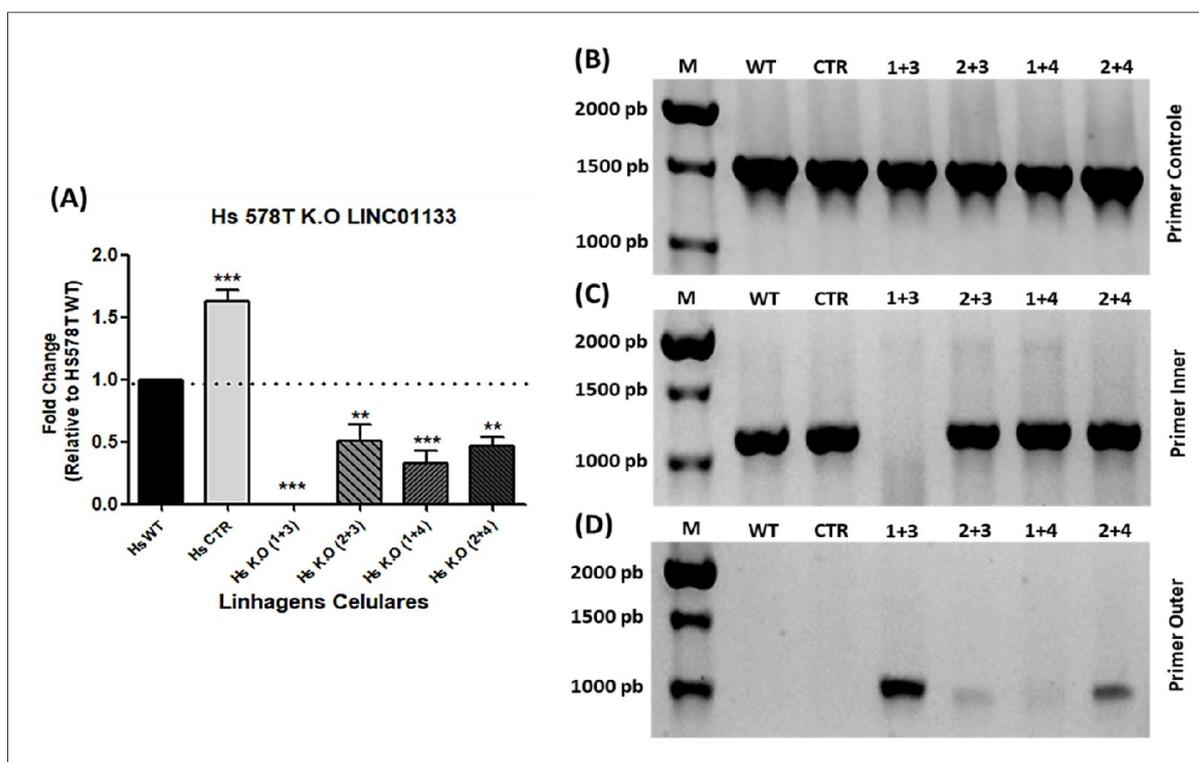


**Figura 33:** Células da linhagem Hs578T transduzidas com as partículas lentivirais do sistema CRISPR/Cas9 para nocauteamento do gene *LINC01133*. Células Hs578T contendo as construções lentivirais LINC01133 (eGFP+) + LINC01133 (resistentes à puromicina), após 48h de seleção com puromicina. Micrografias obtidas utilizando-se o microscópio *EVOS FL Cell Imaging Systems* (Thermo Scientific), no aumento de 100X. Barra: 400µm.

#### 3.4.10. Confirmação do nocauteamento do *LINC01133* nas linhagens transduzidas

Para verificar a eficiência de edição genômica das diferentes combinações de gRNAs (1+3, 1+4, 2+3 e 2+4), as populações de células transduzidas foram expandidas para extração de RNA e de DNA genômico, conforme descrito em

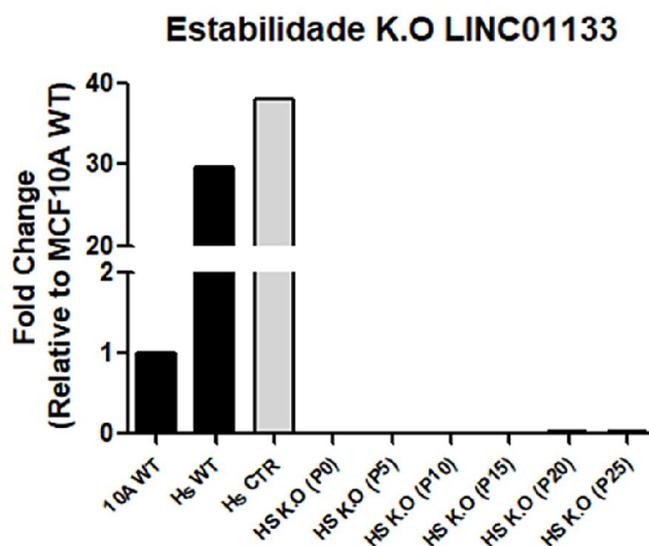
Materiais e Métodos. Conforme esperado, as linhagens Hs578T WT e Hs578T CTR apresentaram níveis elevados de expressão do *LINC01133*. Observamos também que, embora todas as populações K.O apresentaram redução significativa na expressão do lncRNA *LINC01133*, as diferentes combinações de gRNAs apresentaram eficiências distintas na edição genômica. A combinação de gRNAs 1+3 apresentou redução completa na expressão do *LINC01133* quando comparada com as células Hs WT (Figura 34 A). Resultado semelhante foi observado no PCR genômico, tendo a edição genômica completa sido confirmada apenas nesta população de células (Figura 34 C e D).



**Figura 34: Confirmação do nocauteamento do *LINC01133* nas linhagens geradas.** (A) Análise da expressão do *LINC01133* nas linhagens geradas. Os níveis de expressão do *LINC01133* foram analisados por qRT-PCR, utilizando-se *primers* específicos, na linhagem Hs578T parental (Hs WT), controle (Hs CTR) e naquelas co-transduzidas com diferentes combinações de gRNAs (Hs K.O 1+3, 1+4, 2+3 e 2+4). O gráfico representa a razão de *fold change* relativo à expressão na linhagem Hs578T WT  $\pm$  SEM. O gene endógeno utilizado para normalização foi o *HPRT1*. As análises estatísticas foram feitas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Dunnett (\*\*\*)  $P < 0.0001$ ). Foram realizados três experimentos independentes, em triplicatas. (B) PCR genômico do *LINC01133* nas linhagens geradas. As amostras foram visualizadas em gel de agarose 1%, utilizando um primer controle (CHD7) e os *primers inner* e *outer*, os quais diferem entre si na localização genômica a partir do qual os mesmos foram desenhados (*Inner* e *Outer* – primers desenhados, respectivamente, na região interna e externa do locus do gene *LINC01133*). A ausência de amplificação do *primer inner* e a presença de amplificação do *primer outer* nas linhagens

transduzidas com a combinação de gRNAs 1+3 confirmam a deleção gênica do *LINC01133*. M: marcador de peso molecular.

Além disso, para investigar a estabilidade do nocaute nesta população de células, coletamos amostras de passagens sequenciais para verificar o nível de expressão do *LINC01133* e a porcentagem de células GFP+. Os níveis de expressão do lncRNA mostraram-se constantemente reduzidos/ausentes em todas as passagens analisadas (Figura 35), enquanto que o número de células que expressam GFP permaneceu acima de 97% ao longo das passagens (dados não mostrados). Baseados nestes resultados, selecionamos a população de células Hs578T KO *LINC01133* 1+3 (Hs KO LINC) para investigar os efeitos fenotípicos da deleção do *LINC01133* nesta linhagem celular.



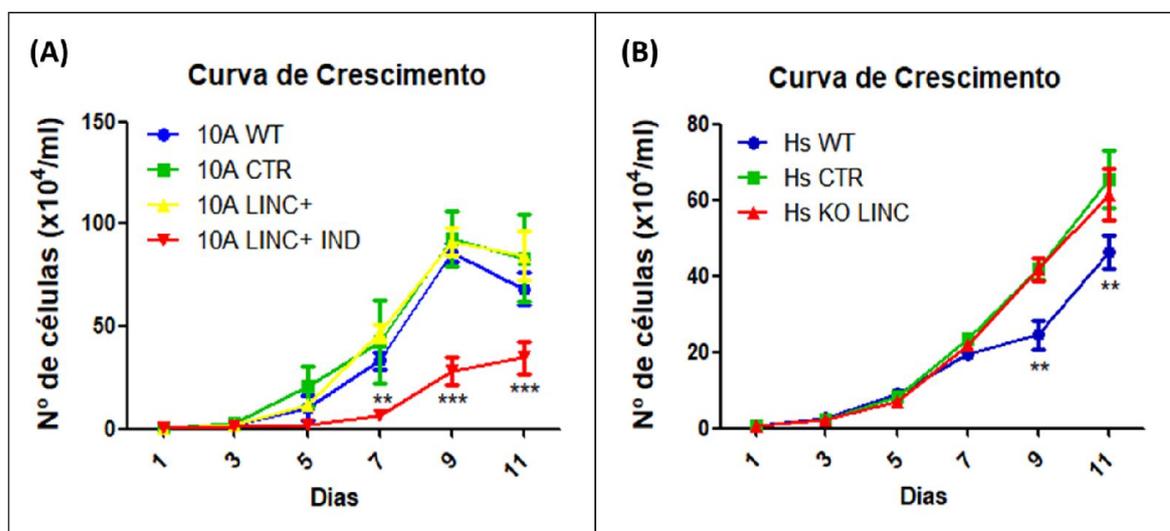
**Figura 35: Análise da estabilidade do nocauteamento do *LINC01133* em diferentes passagens celulares na linhagem HS KO LINC.** Os níveis de expressão do *LINC01133* foram analisados por qRT-PCR, utilizando-se *primers* específicos, nas linhagens Hs578T parental (Hs WT), controle (Hs CTR) e nocauteada do *LINC01133* (Hs KO LINC) em diferentes passagens celulares (P0, P5, P10, P15, P20, P25). O gráfico representa a razão de *fold change* relativo à expressão na linhagem MCF10A WT. O gene endógeno utilizado para normalização foi o *HPRT1*.

### 3.4.11. Caracterização das linhagens celulares geradas

#### 3.4.11.1. Curva de crescimento celular

Para investigar a influência do lncRNA LINC01133 na taxa de proliferação celular foram realizados ensaios *in vitro* de curva de crescimento celular. Foram realizados três experimentos independentes, em triplicatas, com as linhagens parentais MCF10A e Hs578T, com as linhagens controles e com aquelas transduzidas com os vetores de superexpressão e nocauteamento do gene *LINC01133*.

Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, observou-se uma redução considerável, após 7, 9 e 11 dias, na taxa de proliferação das células superexpressando o LINC (10A LINC+ IND) quando comparada com as linhagens 10A WT, 10A CTR e 10A LINC+ (Figura 36 A).



**Figura 36: Análise da taxa de proliferação das linhagens superexpressando e nocauteadas para o lncRNA LINC01133.** A) Curva de crescimento referente às células MCF10A superexpressando o *LINC01133*. B) Curva de crescimento referente às células Hs578T nocauteadas para o *LINC01133*. Os ensaios de curva de crescimento foram realizados inoculando-se  $1 \times 10^4$  células em placas de 35mm de diâmetro. Após 1, 3, 5, 7, 9 e 11 dias de cultivo, estas células foram coletadas e contadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância Two-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0.01$ , \*\*\* $P < 0.0001$ ).

Por outro lado, em relação ao conjunto de células relacionados ao nocauteamento do LINC01133, observou-se aumento considerável, após 9 e 11 dias,

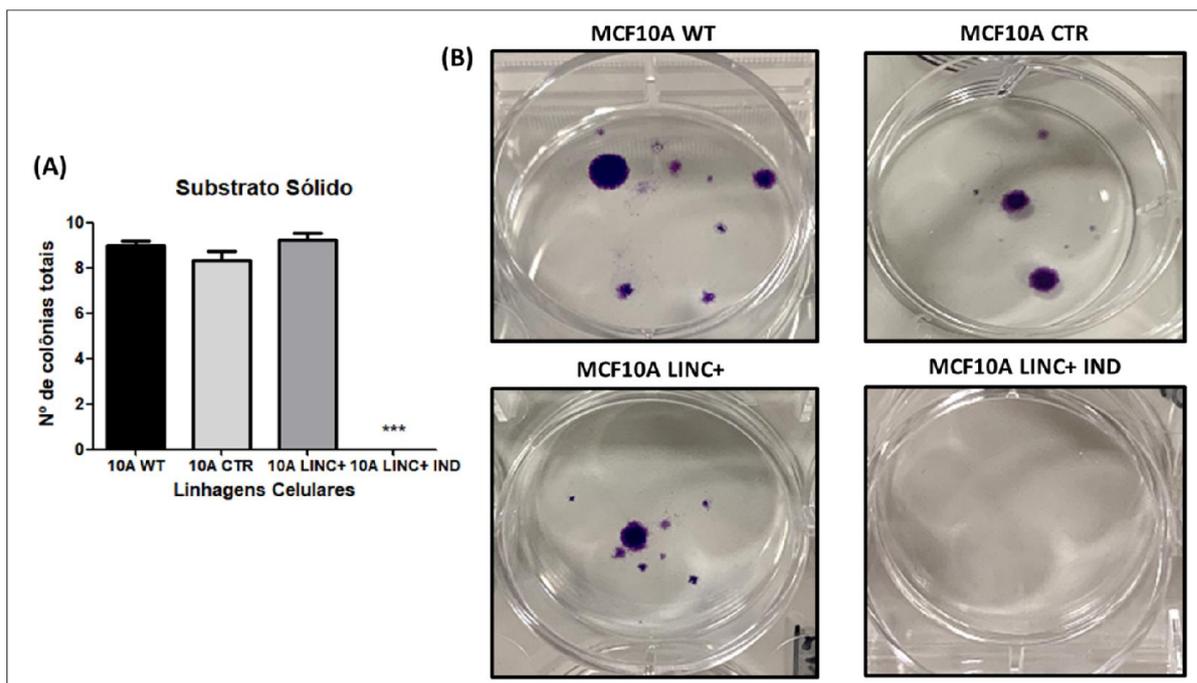
na taxa de proliferação na linhagem nocauteada para o *LINC01133* (Hs KO LINC), embora, também tenha sido observado aumento na taxa de proliferação da linhagem controle (Hs CTR) (Figura 36 B). Conclui-se, portanto, que o *LINC01133* pode ser um importante regulador da proliferação celular nas linhagens analisada.

#### 3.4.11.2. Ensaio de tumorigênese *in vitro*

Para investigar a influência do lncRNA *LINC01133* no processo de tumorigênese, a capacidade das células de formar colônias foi avaliada através de duas abordagens *in vitro*: ensaio de crescimento de colônias em substrato sólido e em suspensão de agarose. Foram realizados três experimentos independentes, em triplicatas, com as linhagens parentais MCF10A e Hs578T, com as linhagens controles e com aquelas transduzidas com os vetores de superexpressão e nocauteamento do gene *LINC01133*.

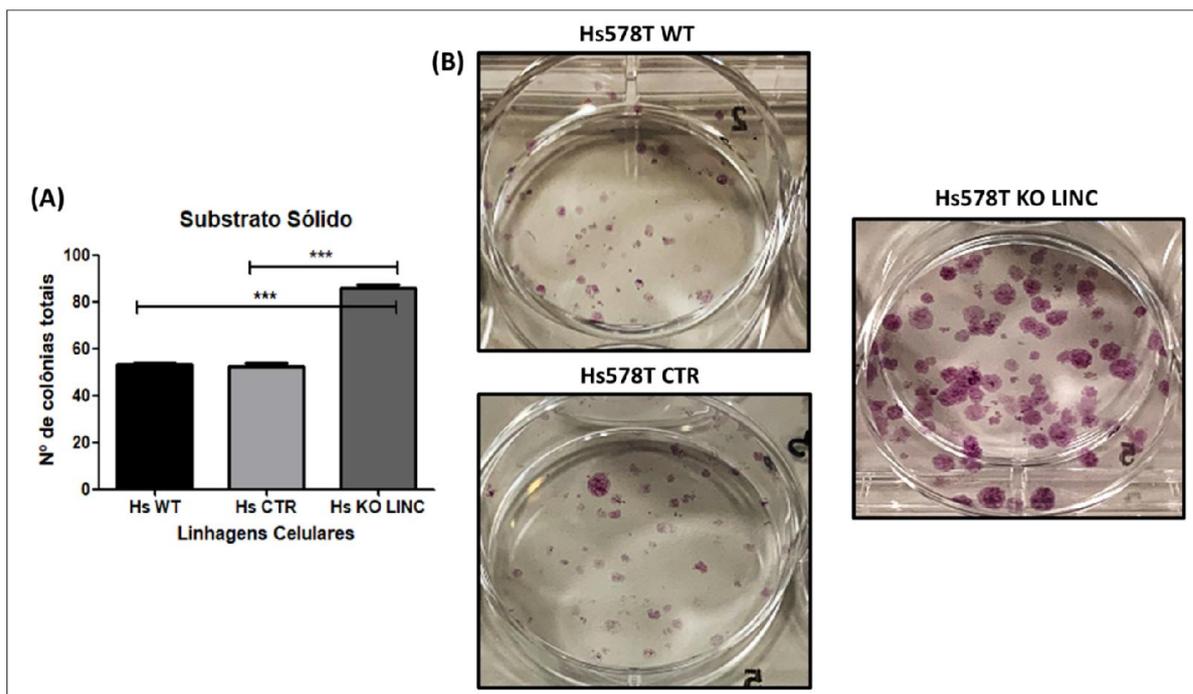
##### 3.4.11.2.1. Ensaio de crescimento de colônias em substrato sólido

Para investigar a influência do *LINC01133* na tumorigênese, medido através da capacidade das células de formar colônias dependente de ancoragem, foram realizados ensaios *in vitro* de crescimento de colônias em substrato sólido. Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, não foi observada formação de colônias na linhagem 10A LINC+ IND quando comparada com as linhagens 10A WT, 10A CTR e 10A LINC+ (Figura 37 A e B), indicando a inibição da capacidade de formação de colônias dependente de ancoragem nestas células.



**Figura 37: Análise do potencial tumorigênico de células MCF10A superexpressando o lncRNA LINC01133 através do ensaio de crescimento de colônias em substrato sólido.** A) Quantificação do número total de colônias ao final do experimento. As colônias foram quantificadas e fotografadas manualmente. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,05$ ). B) Imagem representativa referente à cada linhagem celular ao final do experimento de 12 dias.

Enquanto isto, no conjunto de células relacionadas ao nocauteamento do *LINC01133*, observou-se um aumento significativo na capacidade de formação de colônias dependente de ancoragem na linhagem Hs KO LINC, quando comparada com as demais linhagens. Além disso, um outro fator observado foi o aumento no tamanho das colônias, visivelmente maiores do que nas linhagens Hs WT e CTR (Figura 38 A e B).

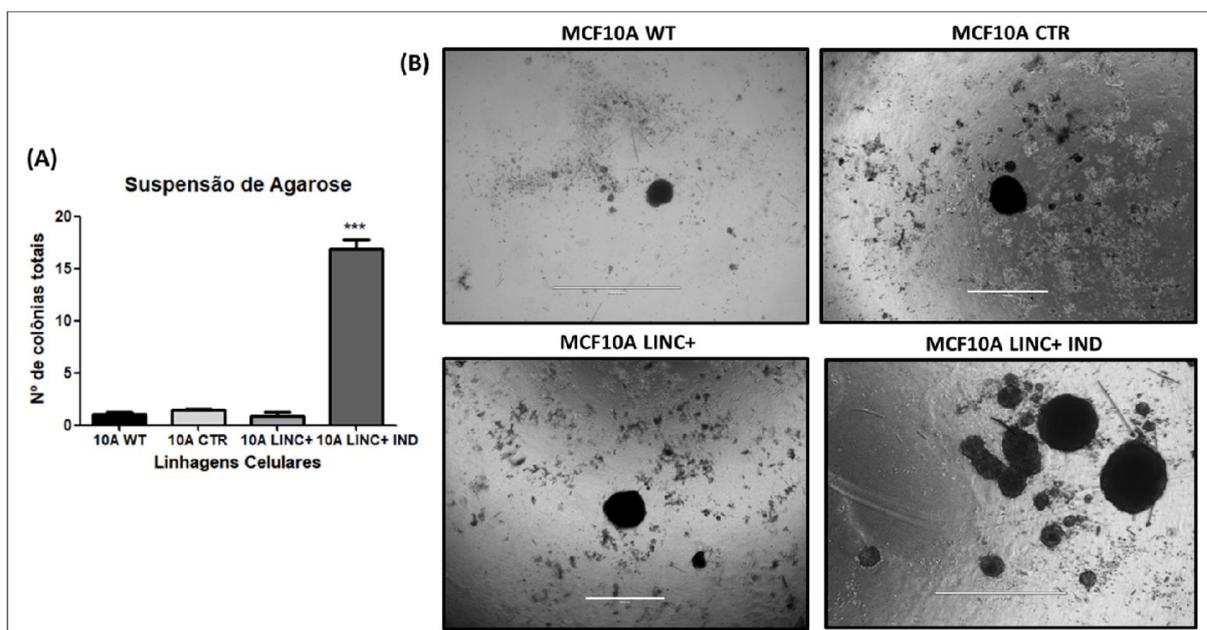


**Figura 38: Análise do potencial tumorigênico de células Hs578T nocauteadas para o lncRNA *LINC01133* através do ensaio de crescimento de colônias em substrato sólido.** A) Quantificação do número total de colônias ao final do experimento. As colônias foram quantificadas e fotografadas manualmente. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni ( $***P < 0,05$ ). B) Imagem representativa referente à cada linhagem celular ao final do experimento de 12 dias.

#### 3.4.11.2.2. Ensaio de crescimento de colônias em suspensão de agarose

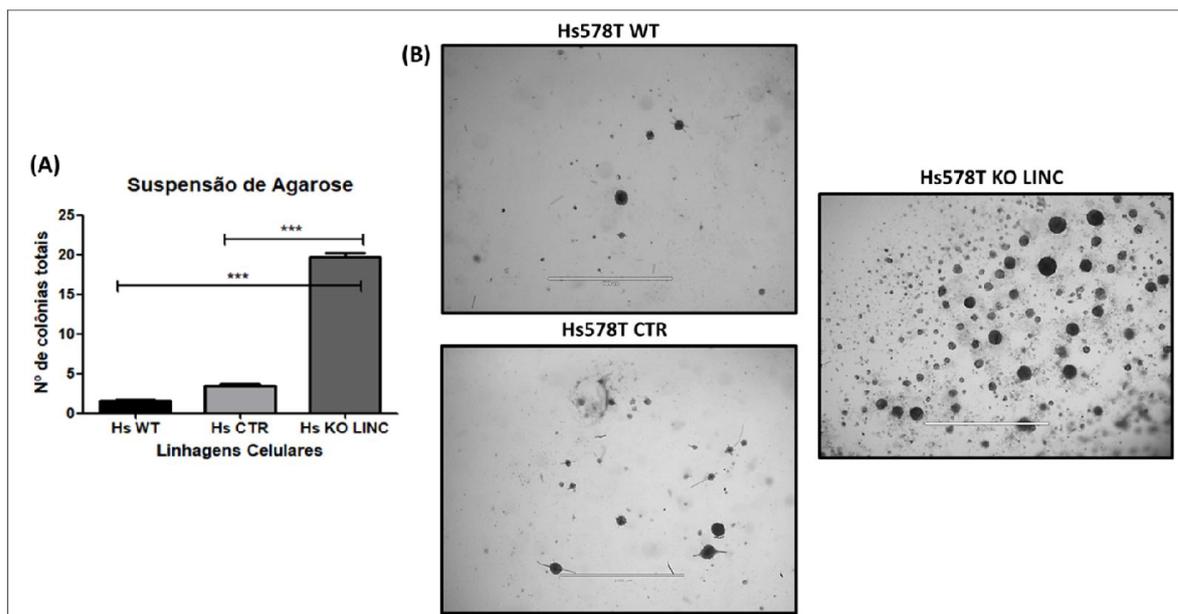
Para investigar a influência do gene *LINC01133* na tumorigênese, medido através da capacidade das células de formar colônias independente de ancoragem, foram realizados ensaios *in vitro* de crescimento de colônias em suspensão de agarose. Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, houve crescimento de colônias independente de ancoragem em todas as linhagens analisadas, mas as linhagens 10A WT, 10A CTR e 10A LINC+ apresentaram um número significativamente menor de colônias (formam geralmente 1 única colônia) quando comparadas com a linhagem 10A LINC+ IND (indução com 250ng/mL de Dox) (Figura 39 A). Um outro fato interessante a ser descrito é que as colônias formadas nesta linhagem são maiores e mais bem definidas (Figura 39 B),

indicando que a superexpressão induzida do *LINC01133* nestas células aumentou de forma significativa o seu potencial tumorigênico.



**Figura 39: Análise do potencial tumorigênico de células MCF10A superexpressando o lncRNA LINC01133 através do ensaio de crescimento de colônias em suspensão de agarose.** A) Quantificação do número total de colônias ao final do experimento. Foram consideradas colônias apenas aquelas que apresentaram diâmetro maior do que 250µm. Os dados são apresentados como a média ± SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\*\*) $P < 0,05$ ). B) Imagem representativa referente à cada linhagem celular ao final do experimento de 21 dias. As colônias foram fotografadas utilizando-se o microscópio EVOS FL Fluorescence Cell Image System no aumento de 40x. Barra: 2.000µm.

Resultado semelhante foi observado no conjunto de células relacionadas ao nocauteamento do *LINC01133*, onde a linhagem nocauteada (Hs KO LINC) apresentou uma capacidade significativamente maior de formar colônias independente de ancoragem (Figura 40 A e B) quando comparada com as linhagens parental (Hs WT) e controle (Hs CTR). A análise desses dados em conjunto pode inferir que, por meio de algum mecanismo, o *LINC01133* regula o potencial tumorigênico e metastático destas linhagens celulares.



**Figura 40: Análise do potencial tumorigênico de células Hs578T nocauteadas para o lincRNA LINC01133 através do ensaio de crescimento de colônias em suspensão de agarose.** A) Quantificação do número total de colônias ao final do experimento. Foram consideradas colônias apenas aquelas que apresentaram diâmetro maior do que 250µm. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\*\*)  $P < 0,05$ ). B) Imagem representativa referente à cada linhagem celular ao final do experimento de 21 dias. As colônias foram fotografadas utilizando-se o microscópio EVOS FL Fluorescence Cell Image System no aumento de 40x. Barra: 2.000µm.

### 3.4.11.3. Ensaio de migração celular *in vitro*

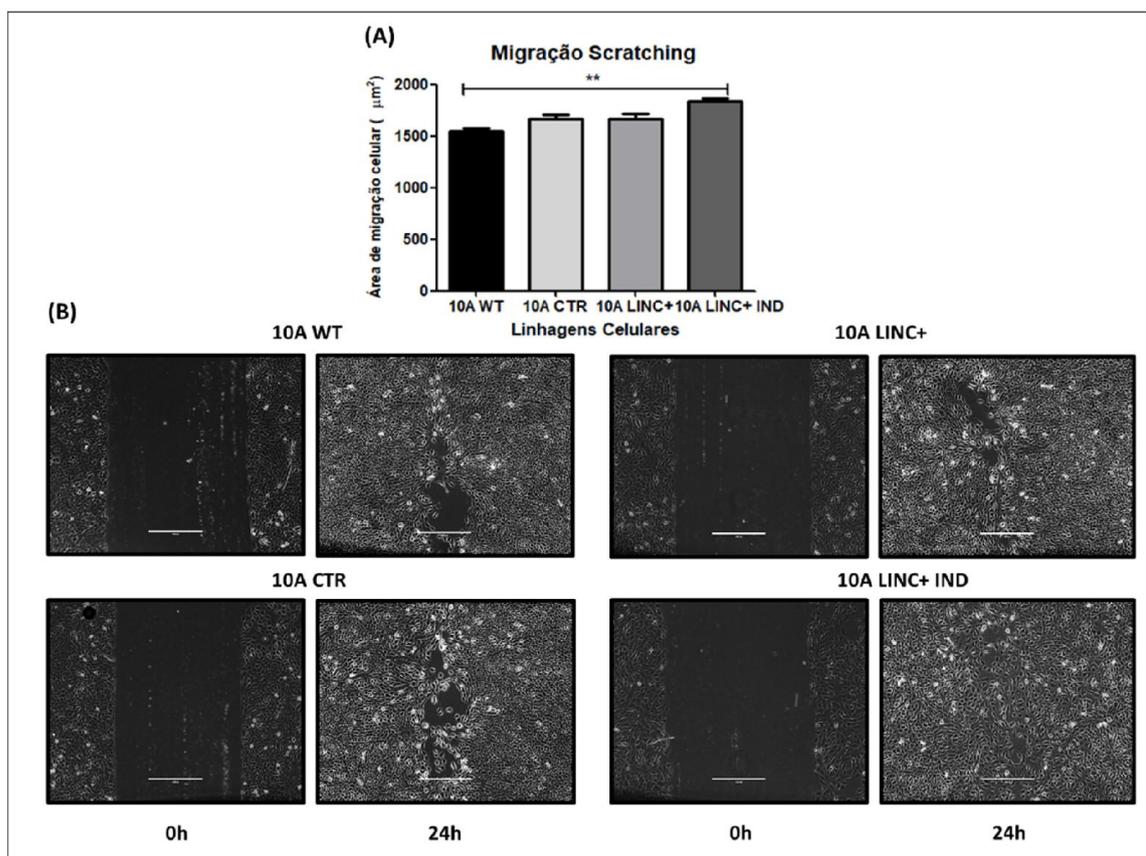
Para investigar a influência do lincRNA LINC01133 no processo de migração celular, a capacidade migratória destas células foi avaliada através de duas abordagens *in vitro*, descritas em Materiais e Métodos: migração por *scratching* e por *Transwell™*. Foram realizados três experimentos independentes, em triplicatas, com as linhagens parentais MCF10A e Hs578T, com as linhagens controles e com aquelas transduzidas com os vetores de superexpressão e nocauteamento do gene *LINC01133*.

#### 3.4.11.3.1 Ensaio de migração celular por *scratching* (*wound healing*)

Nesta abordagem, um risco contínuo foi feito com o auxílio de uma ponteira de 200 µL numa monocamada confluyente das células. Imagens do mesmo campo de

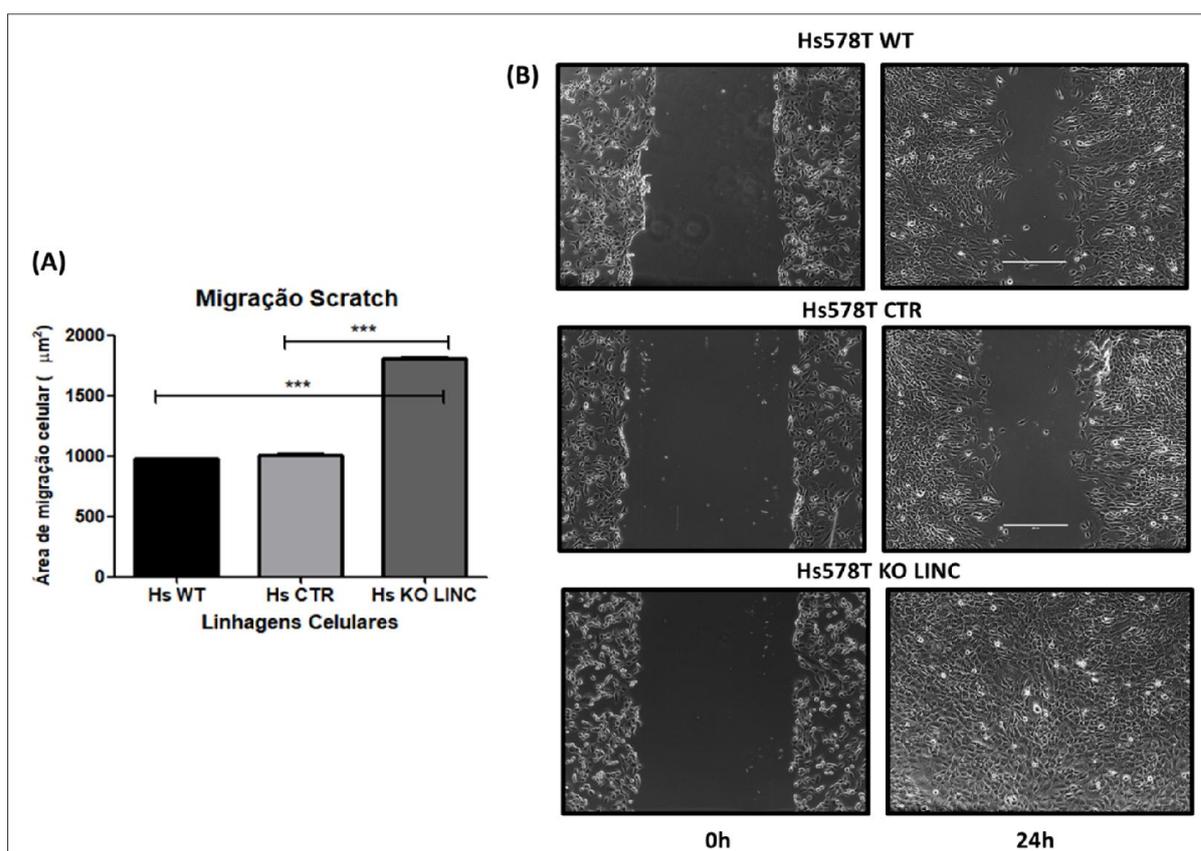
cada poço foram obtidas no tempo de 0h e 24h e a diferença na fração da área da placa ocupada pelas células foi medida conforme descrito em Materiais e Métodos.

Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, pode-se observar que a linhagem 10A LINC+ IND apresentou um leve aumento no potencial migratório, após 24 horas do *scratching*, apenas quando comparada com a linhagem 10A WT. No entanto, nenhuma diferença significativa foi observada quando à comparamos com a linhagem 10A CTR e 10A LINC+ (Figura 41 A e B). Diante deste fato, não podemos afirmar se o aumento no potencial migratório destas células se deve à presença do vetor ou do excesso do lncRNA *LINC01133*.



**Figura 41: Análise do potencial migratório em células MCF10A superexpressando o lncRNA *LINC01133*.** A) O gráfico representa a quantificação da fração de área do risco preenchida pelas células após 0h e 24h do *scratching*. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,05$ ). B) Imagens representativas referente à cada linhagem celular ao final do experimento, nos tempos de 0h e 24h. As células foram fotografadas utilizando-se o microscópio EVOS FL Fluorescence Cell Imagem System no aumento de 100x. Barra: 400µm.

Já em relação ao conjunto de células relacionadas ao nocauteamento do *LINC01133*, pode-se observar que a linhagem Hs578T nocauteada para o *LINC01133* apresentou um potencial migratório significativamente maior, após 24h do *scratching*, quando comparada com a linhagem parental e a linhagem controle (Figura 42 A e B). Conclui-se, portanto, que a ausência do *LINC01133* influencia no potencial migratório das células Hs578T.

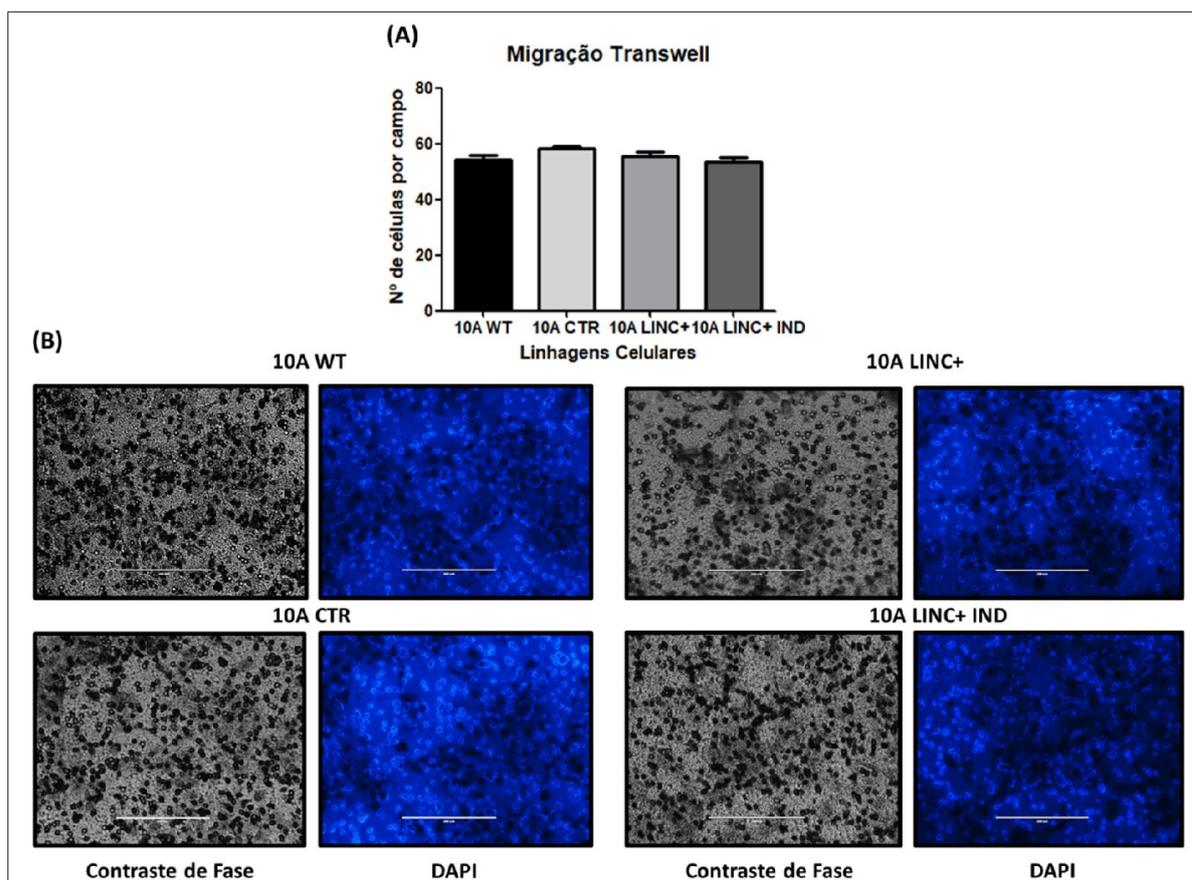


**Figura 42: Análise do potencial migratório em células Hs578T nocauteadas para o lncRNA *LINC01133*.** A) O gráfico representa a quantificação da fração de área do risco preenchida pelas células após 0 e 24h do *scratching*. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em triplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,05$ ). B) Imagens representativas referente à cada linhagem celular ao final do experimento, nos tempos de 0h e 24h. As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 100 $\mu$ m. Barra: 400 $\mu$ m.

#### 3.4.11.3.2. Ensaio de migração celular por *Transwell*<sup>TM</sup>

Nesta abordagem, as células foram plaqueadas sobre insertos do tipo *Transwell*<sup>TM</sup> em placas de 24 poços e mantidas em cultura por 16h-18h, conforme

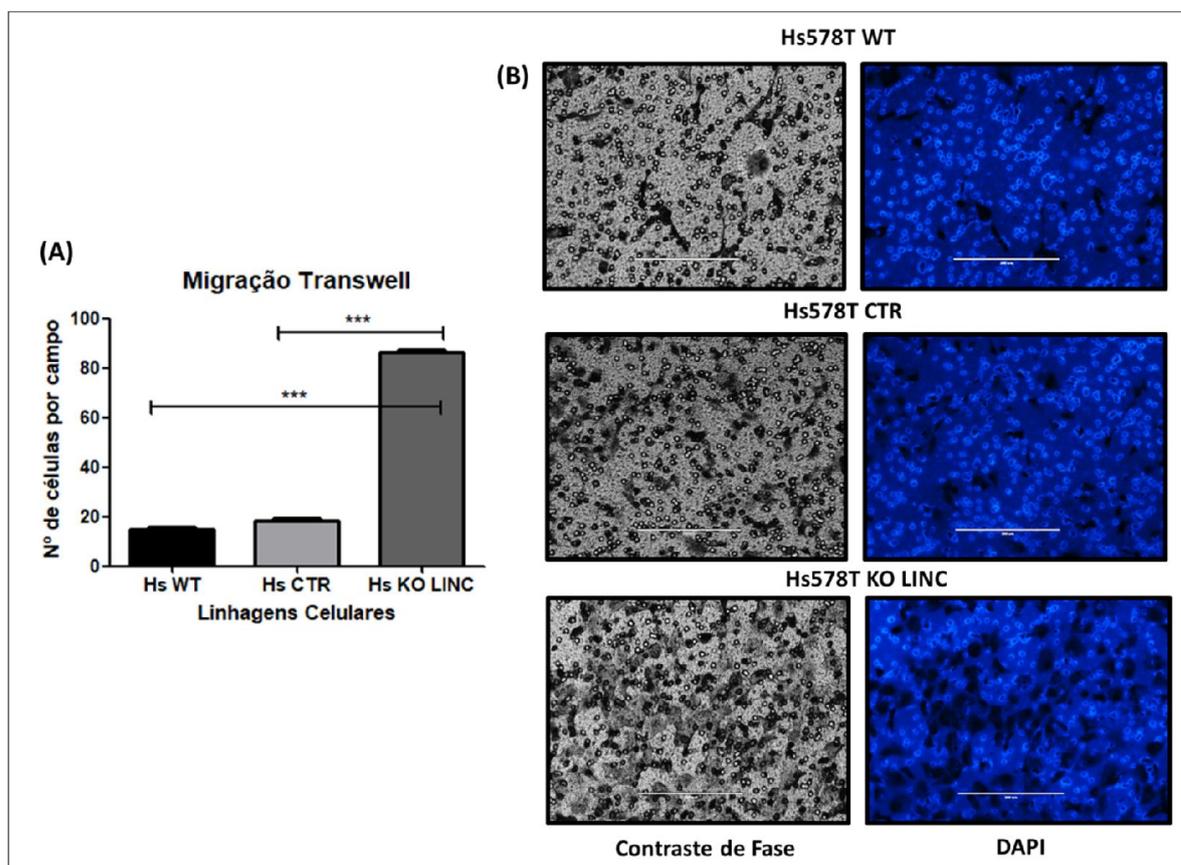
descrito em Materiais e Métodos. Nenhuma diferença significativa foi observada no conjunto de células relacionadas à superexpressão do *LINC01133* (Figura 43 A e B). A partir destes resultados, em conjunto com os resultados obtido nos ensaios de migração por *scratching*, conclui-se, portanto, que a presença do LINC01133 não possui influencia consideravelmente no potencial migratório da linhagem MCF10A.



**Figura 43: Análise do potencial migratório em células MCF10A superexpressando o lincRNA LINC01133.** A) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em duplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni. B) Imagens representativas referente à cada linhagem celular ao final do experimento, em campo microscopia de fase e de fluorescência (DAPI). As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 200x. Barra: 200 $\mu$ m.

Corroborando os resultados obtidos no ensaio de migração por *scratching*, observou-se, no conjunto de células relacionadas ao nocauteamento do *LINC01133*, um aumento significativo e expressivo do potencial migratório na linhagem Hs578T nocauteadas para o *LINC01133*, quando comparada com a Hs WT e Hs CTR (Figura

44 A e B). Conclui-se, portanto, que a ausência do LINC01133 influencia no potencial migratório das células Hs578T.



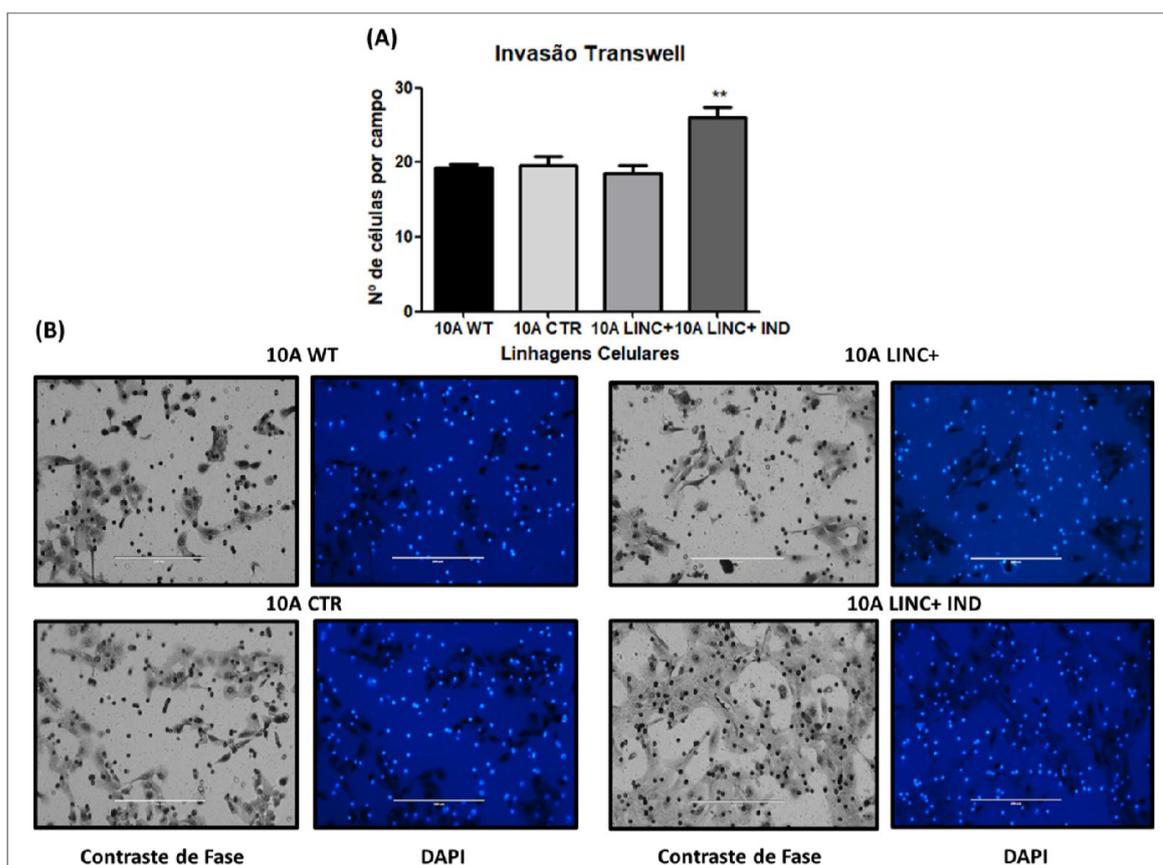
**Figura 44: Análise do potencial migratório em células Hs578T nocauteadas para o lncRNA LINC01133.** A) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em duplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni ( $**P < 0,05$ ). B) Imagens representativas referente à cada linhagem celular ao final do experimento, em microscopia de fase e de fluorescência (DAPI). As células foram fotografadas utilizando-se o microscópio EVOS FL Fluorescence Cell Image System no aumento de 200x. Barra: 200 $\mu$ m.

#### 3.4.11.4. Ensaio de invasão celular *in vitro*

Para investigar a influência do lncRNA LINC01133 no processo de invasão celular foram realizados ensaios de invasão celular, *in vitro*, conforme descrito em Materiais e Métodos. Foram realizados três experimentos independentes, em duplicatas, com as linhagens parentais MCF10A e Hs578T, com as linhagens

controles e com aquelas transduzidas com os vetores de superexpressão e nocauteamento do gene *LINC01133*.

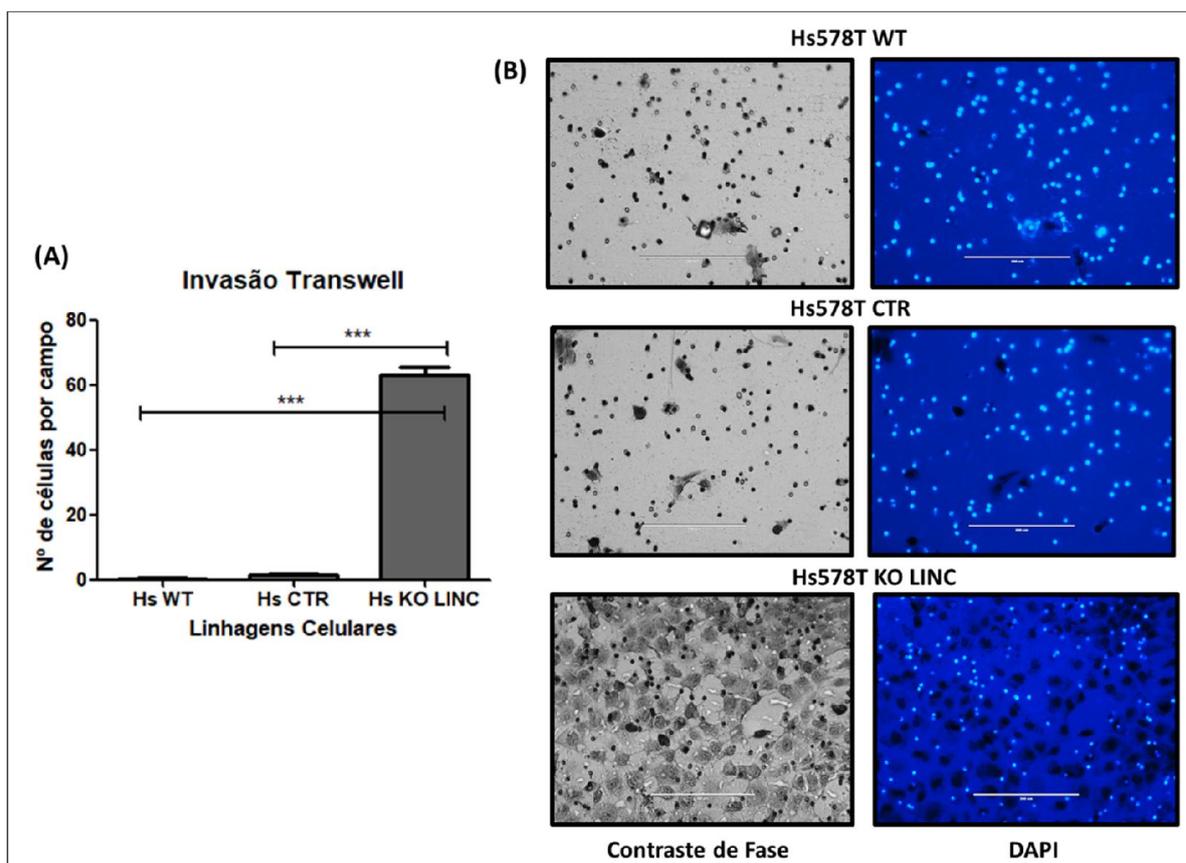
Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, pode-se observar que a linhagem 10A LINC+ IND apresentou aumento significativo no potencial invasivo quando comparada com as demais linhagens (Figura 45 A e B).



**Figura 45: Análise do potencial invasivo em células MCF10A superexpressando o IncRNA *LINC01133*.** A) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em duplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,05$ ). B) Imagens representativas referente à cada linhagem celular ao final do experimento, em microscopia de fase e de fluorescência (DAPI). As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Imagem System* no aumento de 200x. Barra: 200 $\mu$ m.

Já em relação ao conjunto de células relacionadas ao nocauteamento do *LINC01133*, um aumento significativo e expressivo do potencial invasivo foi observado na linhagem Hs578T nocauteadas para o *LINC01133*, quando comparada com a

linhagem parental e a linhagem controle (Figura 46 A e B). Conclui-se, portanto, que a ausência do *LINC01133* influencia no potencial invasivo das células Hs578T.

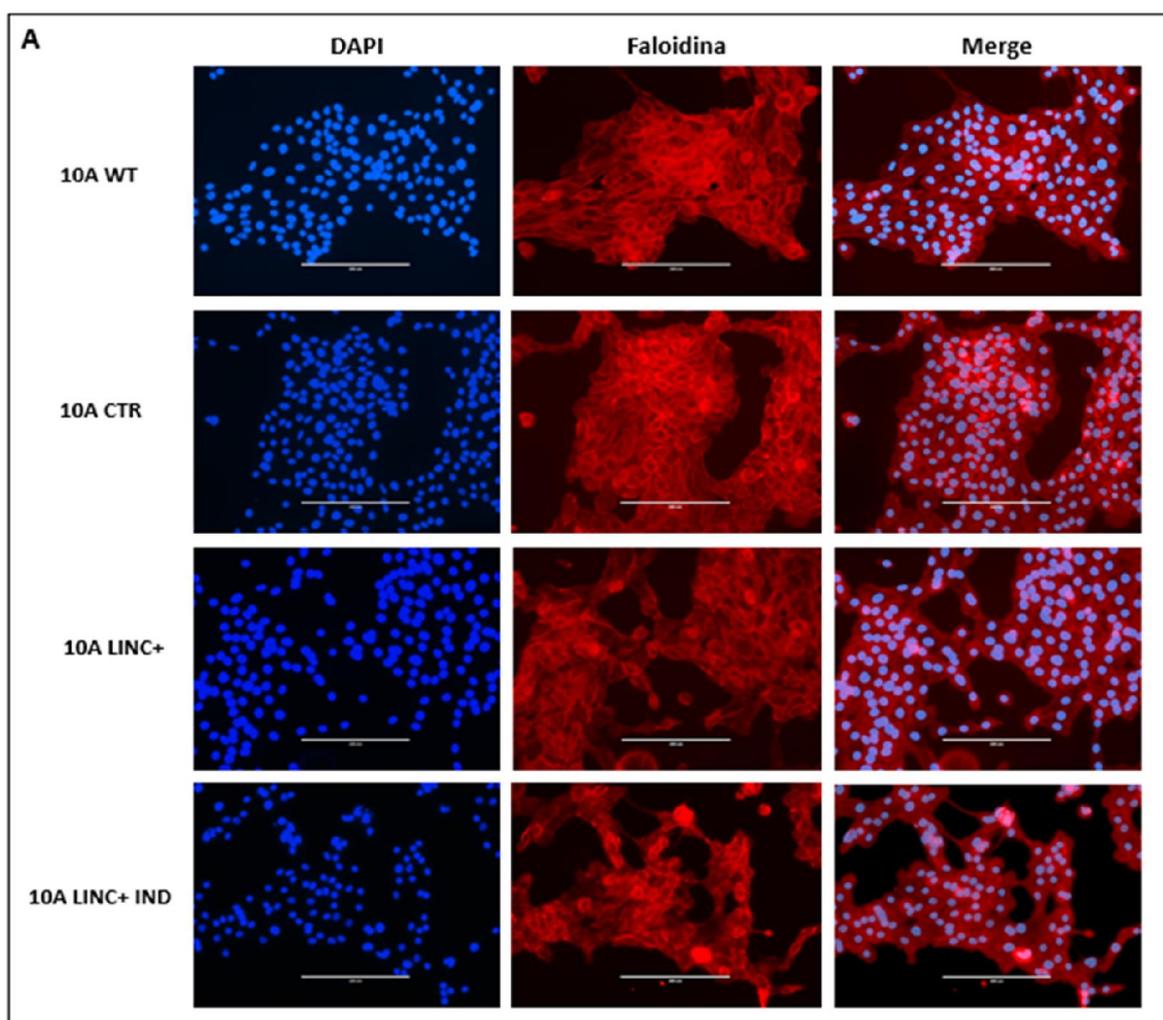


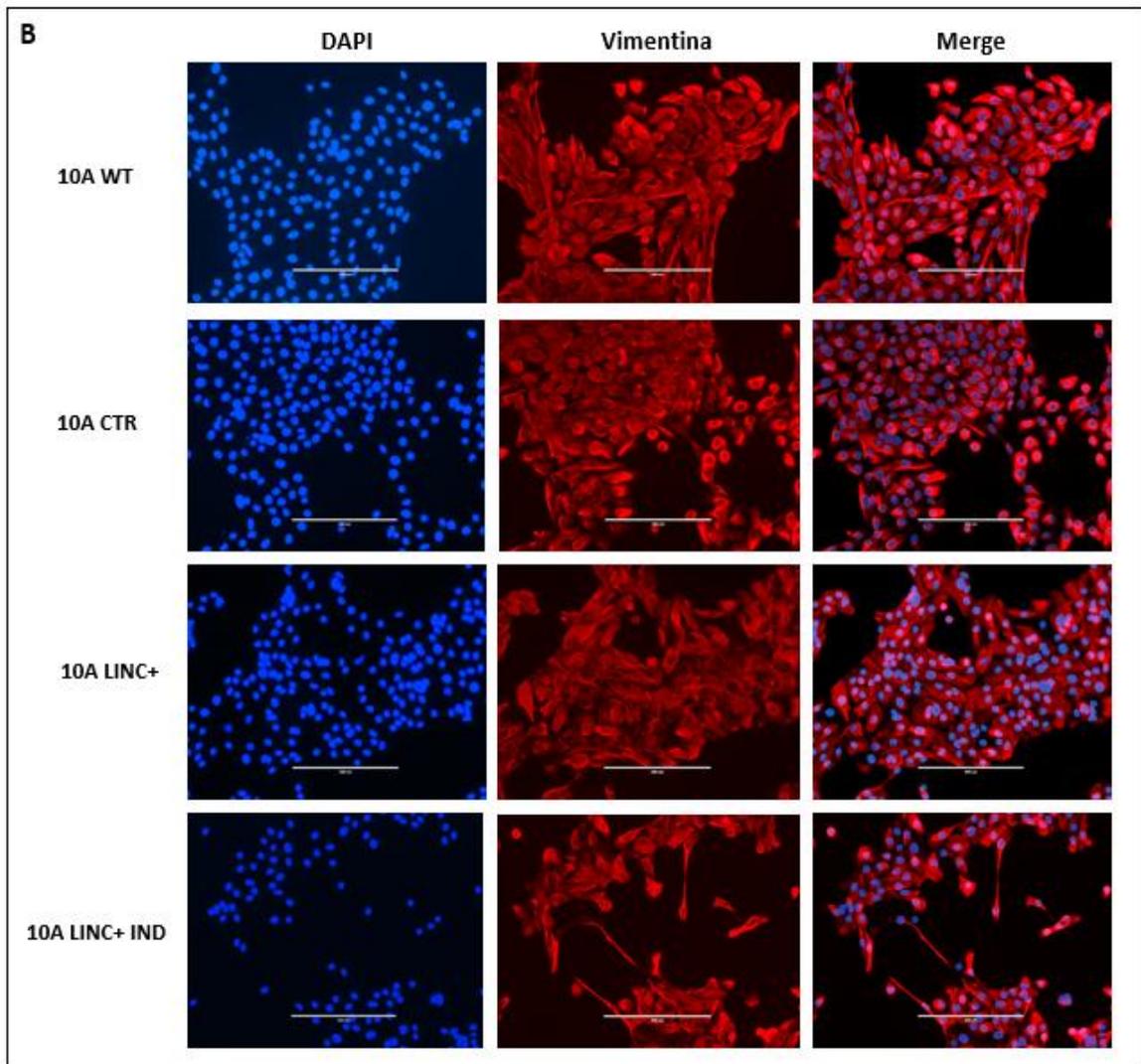
**Figura 46: Análise do potencial invasivo em células Hs578T nocauteadas do lncRNA LINC01133.** A) O gráfico representa a quantificação do número de células por campo nas linhagens analisadas. Os dados são apresentados como a média  $\pm$  SEM de três experimentos independentes, realizados em duplicata. As análises estatísticas foram realizadas utilizando-se o teste de variância One-Way ANOVA, seguido de teste de comparação múltipla Bonferroni (\*\* $P < 0,05$ ). B) Imagens representativas referente à cada linhagem celular ao final do experimento, em microscopia de fase e de fluorescência (DAPI). As células foram fotografadas utilizando-se o microscópio *EVOS FL Fluorescence Cell Image System* no aumento de 200x. Barra: 200 $\mu$ m.

#### 3.4.11.5. Imunofluorescência

Para investigar a influência do lncRNA *LINC01133* na morfologia celular e no citoesqueleto foram realizadas reações de imunofluorescência utilizando os marcadores faloidina (para microfilamentos) e vimentina (para filamentos intermediários). As células da linhagem MCF10A são epiteliais e costumam crescer em ilhas, de maneira bastante organizada, característica dos tecidos epiteliais. Entretanto, na linhagem 10A LINC+ IND observamos uma clara mudança no formato

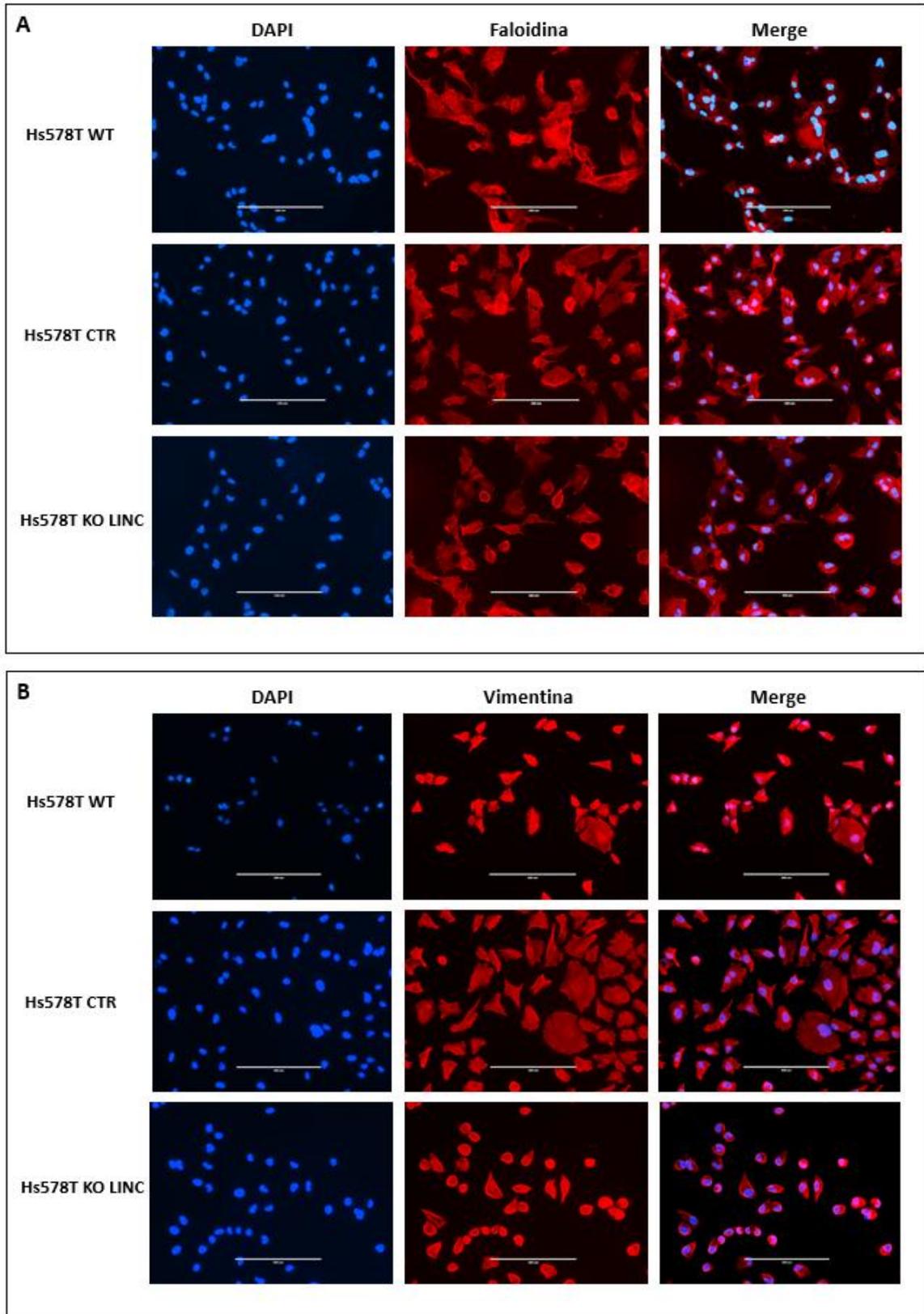
epitelial das células, possivelmente provocada por alterações no citoesqueleto (Figura 47 A). Os limites entre as células perdem em parte a sua definição e estruturas longas e finas, semelhantes a pseudópodes (Figura 47 B), são vistos com grande frequência. Essas características podem estar relacionadas ao leve aumento da capacidade de migração e invasão observada nestas células.





**Figura 47: Imunofluorescência em células MCF10A superexpressando o IncRNA *LINC01133*.** A) Conjuntos de células marcadas com faloidina (em vermelho). B) Conjunto de células marcadas com vimentina (em vermelho). As imagens foram adquiridas utilizando o *EVOS FL Fluorescence Cell Imagem System* (Life Technologies) em aumento de 200x. Barra: 200 $\mu$ m.

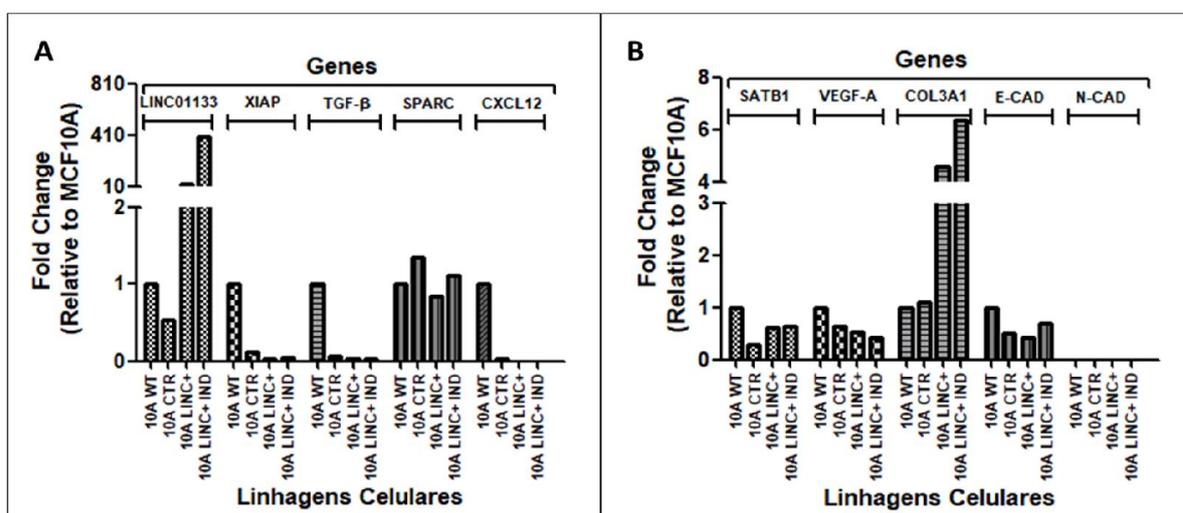
As células da linhagem HS578T são tumorigênicas e, apesar de serem epiteliais, perdem essa característica. As células apresentam formato estrelado semelhante a fibroblastos e emitem pseudópodes bastante característicos. Entretanto, na linhagem Hs KO LINC observamos a perda, em parte, destas características. As células passam a apresentar formato ovalado e perdem consideravelmente a capacidade de formar pseudópodes em grande quantidade (Figura 48 A e B). Essas características provavelmente estão relacionadas ao aumento de características de malignidade observada nestas células, vistos nos resultados dos experimentos anteriores.



**Figura 48: Imunofluorescência em células Hs578T nocauteadas para o lncRNA *LINC01133*.** A) Conjuntos de células marcadas com faloidina (em vermelho). B) Conjunto de células marcadas com vimentina (em vermelho). As imagens foram adquiridas utilizando o *EVOS FL Fluorescence Cell Imagem System* (Life Technologies) em aumento de 200x. Barra: 200µm

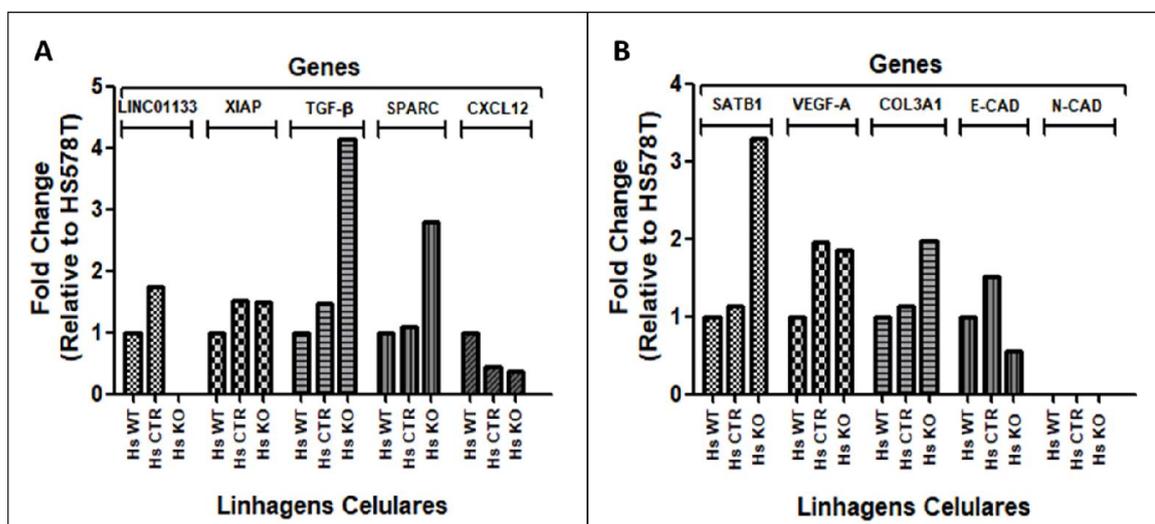
### 3.4.11.6. Análise da expressão de genes envolvidos na malignidade tumoral

Para compreender os possíveis efeitos do lncRNA LINC01133 nos processos de apoptose, angiogênese, migração e invasão celular, investigamos, por qRT-PCR, a expressão de diferentes envolvidos na malignidade tumoral. Em relação ao conjunto de células relacionadas à superexpressão do *LINC01133*, foi observado que a expressão da maior parte dos genes (*XIAP*, *TGF- $\beta$* , *SATB1*, *CXCL12*, *VEGF-A* e *E-CAD*) foi reduzida nas linhagens 10A LINC+ e 10A LINC+ IND, embora também tenha sido observado na linhagem 10A CTR, quando comparadas com a linhagem 10A WT. O gene *SPARC* apresentou padrão de expressão divergente entre as linhagens analisadas, enquanto que *N-CAD* não teve sua expressão detectada. Diante disto, a análise de todos estes genes precisa ser reavaliada. O único gene que apresentou um padrão de expressão destacado foi o *COL3A1*, o qual sua expressão aumentou nas linhagens 10A LINC+ e 10A LINC+ IND, quando comparadas com as linhagens 10A WT e 10A CTR (Figura 49).



**Figura 49: Análise da expressão de genes envolvidos na malignidade tumoral em células MCF10A superexpressando o lncRNA LINC01133.** (A) Análise da expressão dos genes *LINC01133*, *XIAP*, *TGF- $\beta$* , *SPARC* e *CXCL12*. B) Análise da expressão dos genes: *SATB1*, *VEGF-A*, *COL3A1*, *E-CAD* e *N-CAD*. Os genes foram analisados por qRT-PCR e o gene endógeno utilizado para normalização foi o *HPRT1*. Os dados apresentados são referentes a apenas um experimento independente realizado em triplicatas e foram utilizados para nos dar uma visão inicial e geral de possíveis marcadores/vias alterados com a superexpressão do *LINC01133* nas linhagens geradas.

Já em relação ao conjunto de células relacionadas ao nocauteamento do *LINC01133*, foi observado que a expressão do gene anti-apoptótico *XIAP* permaneceu praticamente inalterada nas linhagens analisadas. Para os genes associados aos processos de tumorigênese e transição epitélio-mesenquimal *TGF- $\beta$* , *SPARC* e *SATB1* foi observado o aumento da expressão na linhagem nocauteada para o lncRNA *LINC01133* (Hs KO), enquanto que para o gene *E-CAD* foi observada a diminuição da expressão nesta mesma linhagem, quando comparado com as linhagens parental (Hs WT) e controle (Hs CTR). O gene *CXCL12* apresentou sua expressão reduzida nas linhagens Hs CTR e Hs KO, enquanto que o *N-CAD* não teve sua expressão detectada, o qual necessita ser reavaliado, visto que é importante na transição epitélio-mesenquimal. O gene *VEGF-A* apresentou aumento da expressão nas linhagens Hs CTR e Hs KO, enquanto *COL3A1* apresentou aumento de expressão na linhagem Hs KO (Figura 50).



**Figura 50: Análise da expressão de genes envolvidos na malignidade tumoral em células Hs578T nocauteadas para o lncRNA *LINC01133*.** (A) Análise da expressão dos genes *LINC01133*, *XIAP*, *TGF- $\beta$* , *SPARC* e *CXCL12*. (B) Análise da expressão dos genes: *SATB1*, *VEGF-A*, *COL3A1*, *E-CAD* e *N-CAD*. Os genes foram analisados por qRT-PCR e o gene endógeno utilizado para normalização foi o *HPRT1*. Os dados apresentados são referentes a apenas um experimento independente realizado em triplicatas.

### 3.5. DISCUSSÃO

Com o desenvolvimento das tecnologias de sequenciamento do genoma e de transcriptoma foi estimado que cerca de 85% do genoma é ativamente transcrito, sendo a maior parte referente à transcrição de RNAs não codificadores de proteínas [Djebali *et al.*, 2012]. RNAs longos não codificadores pertencem a uma grande classe de RNAs não codificadores que apresenta comprimento maior que 200 nucleotídeos, sendo frequentemente expressos em padrão espacial, temporal e tecido específico [Derrien *et al.*, 2012]. Numerosos estudos têm mostrado que essas moléculas estão envolvidas em uma ampla gama de processos celulares/biológicos incluindo a iniciação e progressão do ciclo celular, diferenciação, reprogramação da pluripotência de células-tronco, modulação da apoptose, entre outros [Mercer *et al.*, 2009; Fatica *et al.*, 2014; Payer & Lee, 2014]. No entanto, a desregulação destas moléculas tem sido relatada em várias doenças, incluindo o câncer de mama triplo negativo (TNBC), desempenhando importantes papéis na Biologia da doença através da regulação da expressão gênica [Tian *et al.*, 2018; Zhang *et al.*, 2020]. Muitos deles atuam promovendo o desenvolvimento do câncer e de metástases e a resistência a drogas, exibindo, portanto, importante relevância clínica [Prensner & Chinnaiyan, 2011]. A expressão tecido específica dos lncRNAs pode servir como um biomarcador clinicamente útil para diferenciar tecidos normais de tumorais, assim como, também, para a classificação dos subtipos do câncer de mama. Diante disto, os lncRNAs podem ser utilizados no diagnóstico, prognóstico e atuar, também, como potenciais alvos terapêuticos [Lv *et al.*, 2016; Pecero *et al.*, 2019].

Em busca de identificar e caracterizar novos marcadores não proteicos que estejam relacionados com a progressão do câncer de mama triplo negativo (TNBC), no presente trabalho, analisou-se, inicialmente, a expressão de 10 RNAs longos não

codificadores de proteínas (lncRNAs) frente à um painel de linhagens celulares humanas de câncer de mama que apresentam diferentes graus de malignidade e potencial metastático, sendo duas linhagens não-tumorigênicas (MCF12A e MCF10A), uma linhagem referente ao subtipo molecular luminal A (MCF-7), uma linhagem referente ao subtipo molecular luminal B (ZR-75-1), uma linhagem referente ao subtipo molecular HER2+ (SK-BR-3) e duas linhagens referentes ao subtipo molecular *basal-like*/triplo negativo (MDA-MB-231 e Hs578T).

Os resultados da análise de expressão gênica dos lncRNAs selecionados mostraram alguns dados interessantes. Alguns lncRNAs, como o BANCR e o TERC, apresentaram padrões de expressão variáveis entre as linhagens analisadas, evidenciado que não possuem relação com a progressão do carcinoma mamário. O lncRNA GHET1 apresentou elevada expressão na linhagem tumorigênica luminal B ZR-75-1 e o seu papel necessita ser melhor investigado, visto que não existem estudos de caracterização do mesmo nesta linhagem. Um estudo recente relatou, *in vitro* e *in vivo*, o seu papel na linhagem tumorigênica luminal A MCF-7, onde o seu silenciamento suprimiu a proliferação, migração e invasão celular e aumentou a apoptose celular ao manter as células na fase G1 do ciclo celular [Han *et al.*, 2019]. Em outro estudo recente, foi observado que a expressão do GHET1 pode ser induzida por hipóxia levando a ativação excessiva da via de sinalização Hippo/YAP e, com isto, promovendo a progressão do TNBC. Neste estudo foi realizado o silenciamento do GHET1 nas linhagens tumorigênicas triplo negativas HCC1937 e MDA-MB-468, tendo sido observado redução significativa na proliferação, viabilidade e invasão destas células, além da diminuição dos níveis de consumo de glicose e de produção de lactato sob condição de hipóxia [Wang & Liu, 2021].

Os lncRNAs MA-LINC1, HULC, CCAT1, FOXCUT e HOTAIRM1 apresentaram elevada expressão nas linhagens não-tumorigênicas MCF10A e MCF12A quando comparados com as demais linhagens celulares, sugerindo um potencial papel de supressores tumorais. Estudos que investigam o papel destes lncRNAs na progressão do câncer de mama são extremamente escassos, tornando-os alvos promissores para futuros estudos. O MA-LINC1 (RNA longo intergênico não codificador associado à mitose 1) é descrito por ser transcricionalmente regulado pelo fator de transcrição E2F1, exibindo um papel importante na regulação do ciclo celular. O alto nível de expressão do *MA-LINC1* em pacientes com câncer de mama está relacionado à diminuição da sobrevivência, e seu silenciamento sensibiliza as células cancerígenas para o efeito apoptótico da droga Paclitaxel [Bida *et al.*, 2015]. No entanto, o papel e o mecanismo de ação do lncRNA MA-LINC1 na progressão do TNBC permanecem desconhecidos, necessitando ser melhor investigados. O lncRNA altamente superexpresso em câncer de fígado, HULC, foi o primeiro lncRNA identificado como sendo altamente expresso no carcinoma hepatocelular, atuando como um lncRNA oncogênico [Panzitt *et al.*, 2007]. Sua elevada expressão já foi relatada em outros tipos de câncer tais como no câncer de estômago, no câncer colorretal e no câncer de mama [Shi *et al.*, 2016; Xian *et al.*, 2018; Dong *et al.*, 2019]. Divergindo dos nossos resultados, dois estudos relataram que HULC é superexpresso em pacientes e linhagens celulares de TNBC estando associado com o status de malignidade e um mau prognóstico. O silenciamento do HULC nas linhagens triplo negativas MDA-MB-231 e BT549 inibiu a metástase através da redução da expressão de MMP-2 e MMP-9 [Shi *et al.*, 2016; Wang *et al.*, 2019]. Baseado nestes dados, a expressão deste lncRNA necessita ser reavaliada em nosso painel de linhagens celulares.

O lncRNA CCAT1 (transcrito 1 associado ao câncer de cólon) apresenta dados divergentes relatados na literatura em relação à sua expressão no câncer de mama. Em um estudo, realizado com 74 pares de tecidos normais e tumorais de mama derivados da população iraniana de Isfahan, foi observado baixos níveis de expressão do CCAT1 nos tecidos tumorais quando comparados com os tecidos normais [Azadeh *et al.*, 2021], corroborando o observado em nossa análise de expressão gênica, na qual o CCAT1 exibiu expressão apenas nas linhagens não-tumorigênicas MCF10A e MCF12A. Em contrapartida, em outros dois estudos o oposto foi observado. Em um deles, realizado com 94 pares de tecidos normais e tumorais da mama, foi observado níveis de expressão do CCAT1 significativamente maiores nos tecidos tumorais quando comparados com tecidos normais, sendo considerado um potencial marcador de prognóstico na progressão do câncer de mama [Zhang *et al.*, 2015]. No outro, foi relatado a ocorrência de maiores níveis de expressão do CCAT1 em tecidos/linhagens celulares de TNBC quando comparados com tecidos/linhagens normais da mama. Neste trabalho, foi descoberto que o CCAT1 tem como importantes alvos o microRNA miR-218 e a proteína dedo de zinco ZFX e que seu silenciamento inibiu a progressão do TNBC *in vitro* e *in vivo* [Han *et al.*, 2019]. Diante disto, o papel do CCAT1 na progressão do câncer de mama necessita ser melhor investigado, devido aos escassos e divergentes relatos na literatura.

O lncRNA FOXCUT (transcrito 1 a montante do promotor FOXC1) foi inicialmente identificado através de análises de Bioinformática e descrito por apresentar maior expressão em TNBC do que nos outros subtipos de câncer de mama. A expressão do *FOXCUT* e do seu gene adjacente *FOXC1* estão positivamente correlacionadas, atuando em conjunto na forma de um par lncRNA-mRNA e promovendo proliferação e metástases através da ativação da via PI3K/Akt

[Zhang *et al.*, 2020]. O silenciamento do FOXCUT nas linhagens celulares triplo negativas MDA-MB-231 e MDA-MB-468 inibiu a proliferação e migração *in vitro*, indicando um possível envolvimento na progressão do TNBC [Liu *et al.*, 2015]. Este é o único estudo que investigou o papel deste lncRNA no câncer de mama e os dados observados divergem da nossa análise de expressão gênica, na qual observamos uma maior expressão nas linhagens não-tumorigênicas MCF10A e MCF12A.

Por fim, o lncRNA HOTAIRM1, segue o mesmo padrão do FOXCUT e tendo sido relatado como sendo superexpresso em TNBC, em um estudo de Bioinformática que interligou a expressão de diversos lncRNAs com os diferentes subtipos de câncer de mama através de dados obtidos do projeto “*The Cancer Genome Atlas*” desenvolvido pelo *National Institutes of Health* (NIH). Neste estudo, foi descrito que a expressão do *HOTAIRM1* e do seu gene adjacente *HOXA1* é altamente correlacionada, mas que não exibia envolvimento com o câncer [Su *et al.*, 2014]. No entanto, estudos recentes têm mostrado o envolvimento deste lncRNA na progressão de diversos tipos de câncer, tais como: glioblastoma multiforme [Li *et al.*, 2018], câncer de endométrio [Li *et al.*, 2019], câncer de tireoide [Li *et al.*, 2021] e câncer de mama [Kim *et al.*, 2020]. Especificamente no câncer de mama, a expressão do *HOTAIRM1* está correlacionada com o *HOXA1*, estando ambos superexpressos em linhagens MCF-7 resistentes ao tamoxifeno. Estes autores mostraram que HOTAIRM1 interage com a metiltransferase EZH2 e impede que o complexo repressivo polycomb 2 (PRC2) ligue-se ao promotor do *HOXA1*, regulando sua expressão. O silenciamento do HOTAIRM1 restaurou a sensibilidade ao tamoxifeno, podendo vir a ser um potencial alvo terapêutico para pacientes de câncer de mama do tipo luminal A que adquirem resistência ao tamoxifeno [Kim *et al.*, 2020].

Em contrapartida aos lncRNAs listados até aqui, o CYTOR (LINC00152) exibiu elevados níveis de expressão nas linhagens tumorigênicas triplo negativas MDA-MB-231 e Hs578T, quando comparados com as demais linhagens celulares, corroborando os dados da literatura e sugerindo um potencial papel oncogênico para este lncRNA. CYTOR é um lncRNA intergênico localizado a mais de 100 kb do gene codificador mais próximo *PLGLB2* (Plasminogen Like B2). Diversos estudos têm descrito a associação da expressão do *CYTOR* com o desenvolvimento tumoral. A superexpressão do *CYTOR* já foi descrita em diversos tipos de câncer, tais como: estômago, pulmão e fígado [Lyer *et al.*, 2015], além de ser descrita em todos os subtipos do câncer de mama, promovendo proliferação e metástase das células cancerígenas [Van Grembergen *et al.*, 2016]. Este lncRNA atua através da competição com microRNAs endógenos, induzindo a tradução de RNAs mensageiros alvos destes microRNAs, ou através do silenciamento transcricional de genes supressores tumorais, levando à proliferação celular e à transição epitélio-mesenquimal (EMT) [Seo *et al.*, 2019]. O silenciamento do CYTOR na linhagem triplo negativa MDA-MB-231 reduziu significativamente a proliferação e migração celular, devido a modificações nas vias do EGFR e do mTOR, além de ter ocasionado mudanças importantes na morfologia celular, através da reorganização dos filamentos de actina do citoesqueleto, tornando-o um potencial alvo para terapias futuras [Van Grembergen *et al.*, 2016].

Dentre os 10 lncRNAs analisados, o LINC01133 mostrou-se o mais interessante e promissor. Em nossa análise de expressão gênica, observou-se elevado nível de expressão do *LINC01133* na linhagem tumorigênica triplo negativa Hs578T e baixos níveis de expressão nas linhagens não-tumorigênicas MCF10A e MCF12A, corroborando os dados presentes na literatura, sendo que, em um estudo

foi observado o mesmo padrão de expressão do *LINC01133* em um painel com diversas linhagens celulares de câncer de mama, incluindo a linhagem Hs578T [Tu *et al.*, 2019]. Localizado no cromossomo 1q23.2, o *LINC01133*, foi inicialmente descrito como um lncRNA oncogênico no carcinoma de células escamosas do pulmão [Zhang *et al.*, 2015] e, posteriormente, a desregulação da sua expressão foi relatada nos mais diversos tipos de tumores, atuando, em alguns, como um lncRNA oncogênico e, em outros, como um lncRNA supressor tumoral, de maneira tecido-dependente [Huang *et al.*, 2018; Kong *et al.*, 2018]. Atuando como um lncRNA oncogênico, o *LINC01133* apresentou elevada expressão em tecidos e linhagens celulares de osteosarcoma quando comparado com tecidos e linhagens celulares não-tumorigênicas. O silenciamento do *LINC01133*, em linhagens celulares de osteosarcoma, inibiu significativamente a proliferação, migração e invasão celular. Foi descrito que nesses tumores, o *LINC01133* funciona como esponja, ligando-se ao microRNA supressor tumoral miR-442a, impedindo, assim, sua ação [Zeng *et al.*, 2018]. Elevado nível de expressão do *LINC01133* foi descrito, também, no adenocarcinoma ductal pancreático. A superexpressão do *LINC01133* nestes tumores estava associada com a metilação e o silenciamento do gene *DKK1*, sendo responsável por codificar o inibidor da sinalização Wnt DKK1, resultando no aumento da expressão de genes da via de sinalização Wnt, tais como MMP-7,  $\beta$ -catenina e Wnt-5a, e promovendo, desta forma, a proliferação, migração e invasão celular destes tumores [Weng *et al.*, 2019]. Atuando como um lncRNA supressor tumoral, foi descrito baixos níveis de expressão do *LINC01133* em amostras de pacientes com carcinoma de células escamosas orais, quando comparadas, com amostras de pacientes normais. A superexpressão do *LINC01133* em linhagens celulares deste carcinoma inibiu a proliferação e invasão, enquanto o silenciamento exibiu efeitos opostos. Nestes tumores, o *LINC01133* teria

como alvo o fator de crescimento/diferenciação 15 (GDF15), uma citocina pertencente à superfamília do TGF- $\beta$ , atuando através de um *loop* de regulação por *feedback* com este fator [Kong *et al.*, 2018]. No câncer colorretal, foi relatado baixos níveis do LINC01133 quando comparado com tecidos normais [Zhang *et al.*, 2017]. Em um outro estudo, foi descrito que o LINC01133 regula a EMT em linhagens celulares de câncer colorretal, através da interação com o SRSF6, um fator de *splicing* alternativo responsável por promover a EMT e a metástase. O aumento do TGF- $\beta$  inibiria a expressão do *LINC01133*, levando à indução da EMT pelo SRSF6. Por outro lado, a diminuição do TGF- $\beta$  resulta no aumento da expressão do *LINC01133*, o qual interage com o SRSF6 e inibe a EMT. Estes dados indicam que o LINC01133 pode servir como biomarcador de prognóstico para estes tumores [Kong *et al.*, 2016].

O papel do LINC01133 também já foi descrito no câncer de mama. No primeiro trabalho, desenvolvido no ano de 2019 por Song e colaboradores, foi descoberto que a expressão do *LINC001133* foi significativamente reduzida em tecidos de câncer de mama, em comparação com tecidos normais, estando associada à progressão e mau prognóstico em pacientes com câncer de mama. O silenciamento do LINC01133 na linhagem luminal A MCF-7 e na linhagem triplo negativa MDA-MB-231 levou ao aumento da viabilidade, da migração e da invasão celular. Em contrapartida, a superexpressão do *LINC01133* reduziu a viabilidade, a migração e a invasão destas células, revelando que o LINC01133 pode regular negativamente estes processos, atuando como supressor tumoral. *In vivo*, utilizando células da linhagem MDA-MB-231 silenciadas para o LINC01133 e aplicadas em camundongos, foi relatado lesões metastáticas maiores e em maior quantidade quando comparado com o grupo controle. Por fim, foi descrito que o potencial alvo de LINC01133 é o gene SOX4, inibindo sua expressão [Song *et al.*, 2019]. Em outro trabalho, desenvolvido também

no ano de 2019 por Tu e colaboradores, o papel do LINC01133 foi avaliado levando em consideração o microambiente tumoral do câncer de mama triplo negativo. Inicialmente, a expressão do *LINC01133* foi analisada em um painel com diversas linhagens de TNBC. Um dado interessante a se notar é que a expressão do *LINC01133* foi divergente entre as linhagens analisadas, apresentando elevados níveis de expressão nas linhagens HCC70, Hs578T e MDA-MB-468, níveis de expressão levemente elevados nas linhagens HCC1937, BT-20 e BT549 e baixos níveis nas linhagens HCC1143 e MDA-MB-231, quando comparado com a linhagem não-tumorigênica MCF10A. Isto pode indicar que o LINC01133 apresenta papéis distintos dependendo da linhagem de TNBC a ser estudada. Além disto, os níveis de expressão do *LINC01133* foram significativamente maiores em linhagens de TNBC quando comparado com linhagens dos outros subtipos tumorais. Foi relatado, também, que células tronco mesenquimais presentes no microambiente tumoral induzem fortemente a expressão do LINC01133 em células de TNBC vizinhas. Neste estudo, a superexpressão do *LINC01133* nas linhagens MCF10A e MDA-MB-231 não foi capaz de induzir a EMT, mas estimulou a expressão dos marcadores de células-tronco tumorais ALDH1+ e CD44+/CD24- [Tu et al., 2019].

A partir dos nossos resultados e dos dados presentes na literatura, o LINC01133 pode ser considerado como um marcador tumoral promissor e, diante disto, foi escolhido como alvo de estudo. Por ser um lncRNA descrito recentemente, ainda são escassos os trabalhos que investigam o seu papel funcional durante a progressão do câncer de mama, especificamente para o subtipo triplo negativo (TNBC), necessitando ser melhor investigado. Para tentar descobrir o potencial papel do lncRNA LINC01133 no modelo celular de câncer de mama triplo negativo, experimentos de ganho e perda de função foram realizados, *in vitro*. Para tanto,

superexpressamos o *LINC01133* na linhagem não-tumorigênica MCF10A, utilizando um sistema lentiviral de superexpressão indutível por doxiciclina, e nocauteamos o *LINC01133* na linhagem tumorigênica triplo negativa Hs578T, utilizando o sistema CRISPR/ Cas9. A superexpressão e o nocauteamento foram confirmados com êxito.

Os primeiros resultados referentes à caracterização funcional do conjunto de células relacionadas à superexpressão do *LINC01133* mostraram alterações importantes nas taxas de proliferação e no crescimento dependente de ancoragem destas células. Foi observado que houve diminuição significativa na taxa de proliferação e inibição do crescimento de colônias dependente de ancoragem na linhagem 10A LINC+ IND. Alguns possíveis mecanismos podem ser descritos para justificar estas alterações. O primeiro baseia-se na redução da expressão da proteína anti-apoptótica XIAP, a qual possui um papel crítico na promoção da sobrevivência celular e na manutenção da homeostase celular (balanço entre proliferação e morte celular), através da regulação da apoptose. Este processo é iniciado por uma ampla gama de sinais responsáveis por ativar receptores específicos de morte celular presentes na membrana (via extrínseca) e/ou vias intracelulares nas mitocôndrias controladas por membros da família Bcl2 (via intrínseca), as quais convergem na ativação de caspases efetoras específicas promotoras das principais formas de apoptose [Obexer & Ausserlechner, 2014]. XIAP pertence a um extenso grupo de proteínas inibidoras da apoptose e sua expressão tem sido descrita em tecidos normais e cancerígenos [Vischioni *et al.*, 2006; Mohamed *et al.*; 2017]. Esta proteína exibe três domínios importantes responsáveis pela ligação e inibição das caspases 3, 7 e 9, inibindo, desta forma, a apoptose [Shrikhande *et al.*, 2006]. Diante disto, o reduzido nível de expressão de XIAP observado na linhagem 10A LINC+ IND, não seria suficiente para inibir a ação das caspases, levando a uma maior ativação das

vias da apoptose e, conseqüentemente, interferindo na proliferação e na sobrevivência destas células.

O segundo baseia-se no fator de transcrição dedo de zinco KLF4. MCF10A é uma linhagem de células epiteliais mamárias espontaneamente imortalizadas, não-tumorigênicas, derivada de tecido fibrocístico humano. É considerada uma célula epitelial normal da mama por ser quase diploide, apresentar cariótipo estável e importantes características do epitélio normal da mama, tais como ausência de tumorigenicidade, incapacidade de crescer independente de ancoragem e capacidade de formar mamosferas em cultura. Por não expressar os receptores hormonais de estrogênio e progesterona, além do receptor HER2, é considerada uma linhagem não-tumorigênica triplo negativa [Soule *et al.*, 1990; Debnath *et al.*, 2003]. Na literatura já foi descrito que cerca de 70% das linhagens triplo negativas expressam baixos níveis de KLF4 [Nagata *et al.*, 2017]. No entanto, foi relatado que a superexpressão do *LINC01133* na linhagem MCF10A induziu fortemente a expressão do KLF4 [Tu *et al.*, 2020]. KLF4 é descrito por induzir a parada do crescimento, inibindo a proliferação celular através da regulação de genes-chave do ciclo celular, bloqueando, desta forma, a progressão G1/S do ciclo [Shi & Ai, 2013]. Além disso, o silenciamento do KLF4 na linhagem MCF10A levou ao aumento significativo na proliferação destas células, através do aumento da expressão do EGFR [Roberts *et al.*, 2020]. Complementando estes dados, nossos resultados indicaram a redução na expressão da E-caderina, podendo ser um fator adicional para a inibição do crescimento dependente de ancoragem observado nestas células, visto que num estudo que utilizou uma linhagem MCF10A deficiente de E-caderina foi observada redução na adesão célula-substrato, além da diminuição da expressão de genes envolvidos neste processo, tais como *ITGA1*, *COL12A1*, entre outros [Chen *et al.*, 2014].

Outros parâmetros analisados, referentes à caracterização funcional do conjunto de células relacionadas à superexpressão do *LINC01133*, foi o crescimento de colônias independente de ancoragem e a capacidade de migração e invasão destas células. Foi observado que houve um aumento significativo no crescimento de colônias independente de ancoragem, além de um leve aumento na capacidade de migração (visto apenas no ensaio de migração por *scratching*) e invasão destas células. Células normais, como a linhagem MCF10A, são impedidas de crescer independente de ancoragem, devido à *anoikis*, um tipo particular de morte apoptótica [Taddei *et al.*, 2012]. Por outro lado, células transformadas exibem a capacidade de crescer e se dividir sem se ligar a um substrato, característica essa observada na linhagem 10A LINC+ IND e também descrita no trabalho desenvolvido por Tu e colaboradores, os quais também realizaram a superexpressão do *LINC01133* na linhagem MCF10A [Tu *et al.*, 2019]. O mecanismo que leva à transformação destas células ainda não foi desvendado, embora já tenha sido sugerido que estas células podem sofrer transformação espontânea, mas sem apresentar potencial tumorigênico [Shi & Ai, 2013].

No presente trabalho, hipotetizamos que a transformação observada nestas células pode estar relacionada ao aumento considerável da expressão do *COL3A1*, gene responsável por codificar o colágeno tipo III. Componente importante do microambiente tumoral, a matriz extracelular desempenha funções importantes no desenvolvimento tumoral. Os colágenos são as principais proteínas estruturais da matriz extracelular e sua interação com as células é responsável por controlar diversas atividades celulares, incluindo a migração e invasão [Lu *et al.*, 2012; Yeh *et al.*, 2012; Zhang *et al.*, 2013]. O aumento na expressão e deposição de colágeno é descrito por estar associado ao desenvolvimento tumoral e progressão, através da formação de

redes no tecido tumoral. Diversos trabalhos relataram que o aumento na deposição de colágeno aumenta a rigidez e densidade do tecido mamário, sendo considerados importantes fatores de risco para o desenvolvimento do câncer de mama [Provenzano *et al.*, 2006; Provenzano *et al.*, 2008]. Elevados níveis de expressão do *COL3A1* têm sido relatados em diversos tipos de câncer, incluindo o câncer de mama triplo negativo, estando associados ao desenvolvimento tumoral [Liu *et al.*, 2010; Xiong *et al.*, 2014; Engqvist *et al.*, 2019; Shi *et al.*, 2020]. Diante disto, o elevado nível de expressão do *COL3A1*, associado à leve redução da expressão da E-caderina, justificaria o aumento do crescimento de colônias independente de ancoragem, assim como, também, o leve aumento da capacidade migratória e invasiva observada na linhagem 10A LINC+ IND.

No câncer de mama, os lncRNAs atuam por diversos mecanismos, incluindo o recrutamento de enzimas modificadoras da cromatina e modulando componentes epigenéticos [Hu *et al.*, 2012]. Na literatura foi demonstrado que a hipermetilação de ilhas CpGs do gene *CDH1*, responsável por codificar a E-caderina, resultou na redução de sua expressão, estando associado com a EMT em diferentes carcinomas humanos, incluindo o câncer de mama [Nas *et al.*, 200; Chang *et al.*, 2002; Lombaerts *et al.*, 2006]. Além disso, Tu e colaboradores relataram que a superexpressão do *LINC01133* na linhagem MCF10A estimulou a expressão do marcador ALDH1+ e de população de células CD44+/CD24- [Tu *et al.*, 2019]. Células CD44+ expressam altos níveis da proteína FOXO1, devido à sua desmetilação, estando também associado com a repressão da expressão da E-caderina e indução da EMT [Novak *et al.*, 2012]. A superexpressão de lncRNAs e seu papel na indução da EMT tem sido descritos no câncer de mama. A superexpressão do linc-ROR induziu o processo de EMT em células epiteliais mamárias humanas [Hou *et al.*, 2014], enquanto que a

superexpressão do LINC00617 induziu a EMT através da ativação da transcrição do SOX2 em linhagens celulares de câncer de mama [Li *et al.*, 2017]. Por fim, a superexpressão do lncRNA-HIT levou ao aumento da EMT induzida por TGF- $\beta$  em linhagens celulares de câncer de mama [Richards *et al.*, 2015]. Estes dados podem embasar o que foi observado em nossos resultados, indicando que a superexpressão do *LINC01133* na linhagem MCF10A pode contribuir, via redução da E-caderina, para o leve aumento da capacidade de migração e invasão. No entanto, isto não seria suficiente para induzir a EMT na linhagem 10A LINC+ IND, conforme já descrito na literatura [Tu *et al.*, 2019].

Por fim, o último parâmetro analisado no conjunto de células relacionadas à superexpressão do *LINC01133*, foi a alteração na morfologia celular, visto através de reações de imunofluorescência, utilizando os marcadores faloidina e vimentina. Alterações morfológicas importantes foram observadas na linhagem 10A LINC+ IND, tais como a mudança no formato epitelial das células, perda parcial dos limites celulares e a presença de estruturas longas e finas, semelhantes a pseudópodes. A redução da expressão da E-caderina observada nestas células possivelmente contribui para estas alterações. Em um estudo que utilizou uma linhagem MCF10A deficiente de E-caderina foram observadas alterações importantes no arranjo citoesquelético de actina e de tubulina, apresentando características semelhantes aquelas encontradas em nossos resultados [Chen *et al.*, 2014]. Linhagens celulares de câncer de pâncreas que superexpressavam o *LINC01133*, após indução com periostina, também apresentaram diminuição na expressão de E-caderina e os mesmos padrões de alterações morfológicas observado em nossos resultados [Liu *et al.*, 2021].

Nossos dados, referentes ao conjunto de células relacionadas ao nocauteamento do *LINC01133*, mostraram aumento significativo, na linhagem Hs KO LINC, em todos os parâmetros analisados: na taxa de proliferação, no crescimento de colônias dependente e independente de ancoragem e na capacidade migratória e invasiva destas células. Estes resultados indicam que o lncRNA *LINC01133* exerce um importante papel supressor tumoral no câncer de mama triplo negativo, estando de acordo com os dados já descritos na literatura. O silenciamento do *LINC01133* na linhagem triplo negativa MDA-MB-231 promoveu significativamente a proliferação, a migração e invasão, enquanto que a sua superexpressão exibiu efeitos opostos, revelando que o *LINC01133* regula negativamente estes processos. A expressão do *LINC01133* na linhagem triplo negativa MDA-MB-231 é importante para reprimir a expressão do fator de transcrição *SOX4* através do recrutamento de *EZH2* para o promotor do *SOX4* [Song *et al.*, 2019]. Este mesmo papel pode ser exercido pelo *LINC01133* na linhagem triplo negativa Hs578T, a qual foi utilizada em nosso estudo, embora esta linhagem apresente níveis mais elevados de expressão de *LINC01133*, quando comparada, com a linhagem MDA-MB-231. A diferença na expressão do *LINC01133* nestas células pode ser justificada pelas diferentes características observadas entre elas. A linhagem MDA-MB-231 é derivada de efusão pleural, sendo classificada patologicamente como adenocarcinoma não-invasivo e pouco diferenciada (grau histológico III). Por outro lado, a linhagem Hs578T é derivada de tumor primário, sendo classificada patologicamente como carcinosarcoma ductal invasivo e também é pouco diferenciada (grau histológico III). Devido a estas características, estas células são mais agressivas, apresentando maior potencial tumorigênico e metastático quando comparada com a linhagem MDA-MB-231 [Chavez *et al.*, 2010].

O aumento na taxa proliferativa e no crescimento de colônias dependente de ancoragem da linhagem Hs KO LINC estão possivelmente relacionadas a expressão do KLF4, um importante regulador da proliferação celular e de *stemness* (capacidade de uma célula para perpetuar sua linhagem, de originar células diferenciadas e de interagir com o microambiente) no câncer de mama triplo negativo [Yu *et al.*, 2011]. Células-tronco tumorais, descritas por serem abundantes no TNBC, exibem propriedades de *stemness*, sendo responsável por sustentar a progressão tumoral, através do aumento da capacidade proliferativa e metastática [Aponte & Caicedo, 2017]. Tem sido descrito que o LINC01133 é um importante ativador da expressão do *KLF4* em células de câncer de mama triplo negativo, regulando negativamente estes processos. Cerca de 68% dos TNBCs exibem baixa expressão do *KLF4*, enquanto 32% expressam altos níveis [Nagata *et al.*, 2017], os quais são importantes para regular a capacidade proliferativa, migratória e invasiva destes tumores. A diferença na expressão do KLF4 no TNBC provavelmente é devida aos diferentes níveis de agressividade observado nestes tumores [Roberts *et al.*, 2020]. O silenciamento do LINC01133, utilizando oligonucleotídeos antisense (ASO, da sigla em inglês), nas linhagens celulares triplo negativas MDA-MB-468 e Hs578T, resultou na inibição de aproximadamente 40% da expressão do KLF4 endógeno [Tu *et al.*, 2019]. Além disso, foi descrito que o KLF4 reprime a expressão do EGFR, receptor do fator de crescimento epidermal [Roberts *et al.*, 2020], responsável por ativar importantes vias de sinalização promotoras da proliferação, migração, invasão e sobrevivência em células do câncer de mama, tais como PI3K-AKT-mTOR, Ras-Raf-MEK-ERK e Src-STAT3 [Wee & Wang, 2017]. Diante disto, o nocauteamento do *LINC01133* reduziria, ainda mais, a expressão do KLF4, levando à ativação da

expressão do EGFR, sendo responsável pelo aumento das taxas proliferativas observadas na linhagem Hs KO LINC.

Por outro lado, o aumento no crescimento de colônias independente de ancoragem (potencial tumorigênico) e da capacidade de migração e invasão observados na linhagem Hs KO LINC estão possivelmente relacionadas com a expressão do fator de transcrição SOX4. Conforme descrito anteriormente, o LINC01133 é um importante regulador negativo da transcrição do SOX4, via recrutamento do EZH2 para o promotor do SOX4 [Song *et al.*, 2019]. O fator de transcrição SOX4 é estruturalmente caracterizado, por apresentar um domínio altamente conservado que se liga diretamente ao sulco menor da hélice de DNA e que está envolvido, através de sua atividade transcricional, em vários processos do desenvolvimento embrionário [Cheung *et al.*, 2000; Schepers *et al.*, 2002]. No entanto, diversos estudos têm descrito que SOX4 é altamente expresso nos mais diversos tipos de tumores, atuando como um oncogene e promovendo fenótipos malignos [Moreno, 2020]. Pesquisas recentes têm relatado o envolvimento do SOX4 na promoção da tumorigênese e da EMT no câncer de mama, *in vitro* e *in vivo*. Um dos mecanismos descritos de atuação do SOX4 é através de sua ligação ao promotor do CXCR7, induzindo fortemente sua expressão (Zhang *et al.*, 2012; Zhang *et al.*, 2020). A ativação do CXCR7 é responsável por promover a tumorigênese, migração, invasão, angiogênese, entre outros processos [Xin *et al.*, 2017]. Além disso, no TNBC, a expressão do SOX4 é diretamente induzida pelo TGF- $\beta$  [Bagati *et al.*, 2021], tendo sido observado aumento de sua expressão na linhagem Hs KO LINC. Baseado nisto, o nocauteamento do LINC01133 nesta linhagem, levaria ao aumento da expressão do SOX4, por meio de dois mecanismos: I) pela ausência da repressão, via EZH2, da transcrição do SOX4 e II) pela indução da transcrição do SOX4 pelo TGF- $\beta$ . Além

disto, estudos tem descrito que a expressão de *SOX4* pode reduzir a transcrição da E-caderina indiretamente, através da ativação da expressão de *ZEB1*, um importante indutor da EMT [Zhang *et al.*, 2012]. Isto justificaria a redução da expressão da E-caderina observada na linhagem Hs KO LINC. Assim, o conjunto destes fatores justificaria o expressivo aumento no potencial tumorigênico e metastático observado na linhagem Hs KO LINC.

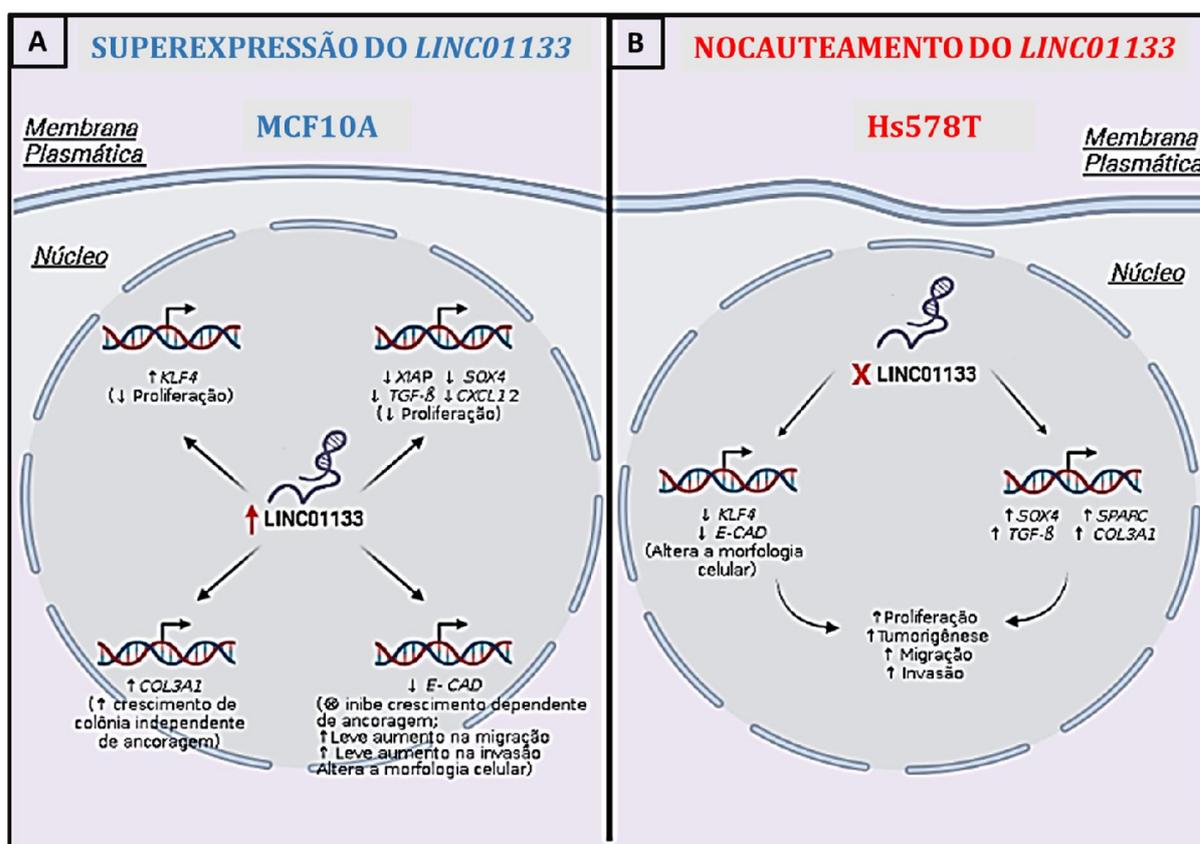
Nossos resultados, indicaram aumento na expressão de *SPARC*, *SATB1* e *COL3A1*, sendo fatores adicionais para justificar o aumento da malignidade observada na linhagem Hs KO LINC. *SPARC* (proteína ácida secretada e rica em cisteína) é uma glicoproteína de ligação à albumina, sendo secretada pelas células, modulando suas interações com a matriz extracelular. Esta proteína desempenha papéis importante na proliferação e migração celular e elevados níveis de sua expressão tem sido relacionado com uma maior agressividade de TNBC [Zhu *et al.*, 2016]. *SATB1* (proteína especial 1 de ligação à sequências ricas em AT) pertence a uma classe de reguladores transcricionais que atuam como “plataformas de aterrissagem” para diversas enzimas remodeladoras da cromatina. Estima-se que *SATB1* é responsável por regular mais de 1.000 genes e sua expressão tem sido descrita por contribuir para a progressão do câncer de mama [Zheng, 2008]. O papel do *COL3A1* já foi descrito anteriormente e, assim como o *SPARC* e o *SATB1*, elevados níveis de sua expressão estão associados ao desenvolvimento do TNBC [Shi *et al.*, 2020].

Por fim, o último parâmetro analisado no conjunto de células relacionadas ao nocauteamento do *LINC01133*, foi a alteração na morfologia celular, visto através de reações de imunofluorescência utilizando os marcadores faloidina e vimentina. Alterações morfológicas importantes foram observadas na linhagem Hs KO LINC, tais como a perda do formato estrelado, além da perda parcial dos pseudópodes,

características marcantes da linhagem Hs578T. Assim como descrito na linhagem 10A LINC+ IND, estas mudanças podem estar relacionadas com a redução da expressão da E-caderina, bem como ao aumento das características de malignidade observado nestas células. Hipoteticamente, a alteração da morfologia observada nestas células poderia favorecer os processos de migração e invasão. O silenciamento do LINC01133 em linhagens celulares de câncer colorretal inibiu significativamente a E-caderina e ocasionou importantes mudanças na morfologia das células [Kong et al., 2016]. Em outro estudo, o silenciamento do LINC01133 em linhagens celulares de endometriose também alterou a morfologia das células, as quais tornaram-se mais largas e achatadas, com um aumento no número de fibras de estresse de actina [Yotova et al., 2021].

Assim, em suma, identificamos o lncRNA intergênico LINC01133 e analisamos seu papel na progressão do câncer de mama triplo negativo. Um conjunto de mecanismos hipotéticos de atuação deste lncRNA neste subtipo tumoral foram descritos e correlacionados com as alterações fenotípicas observadas em nossos resultados. Estes mecanismos poderão ser confirmados ou rejeitados através da análise do transcriptoma destas células pela tecnologia do RNA-Seq (análises em andamento), devendo fornecer informações importantes sobre as bases biológicas deste subtipo tumoral, além de contribuir para o desenvolvimento de novas estratégias terapêuticas direcionadas. Os dados descritos aqui sugerem que o lncRNA LINC01133 exerce um importante papel supressor tumoral, regulando o potencial tumorigênico e metastático da linhagem celular triplo negativa Hs578T. É importante ressaltar que, pela primeira vez, o nocauteamento deste lncRNA, via sistema CRISPR/Cas9, foi realizado em linhagens celulares humanas de câncer, em nosso caso, especificamente em linhagens celulares de câncer de mama triplo negativo.

Diante disto, os dados de caracterização funcional obtidos aqui adicionarão informações pioneiras e complementares aos dados já descritos na literatura referentes a este subtipo tumoral. Juntos, nossos resultados sugerem um cenário extremamente promissor para o LINC01133, demonstrando que este lncRNA pode servir como um novo biomarcador de prognóstico, além um potencial alvo terapêutico para o tratamento do câncer de mama triplo negativo.



**Figura 51: Mecanismos hipotéticos ativados ou reprimidos devido à superexpressão e ao nocauteamento do gene não codificador de proteína LINC01133 nas linhagens estudadas.** A) A superexpressão do LINC01133 na linhagem não-tumorigênica MCF10A resultou na redução da taxa de proliferação celular, possivelmente ocasionada pelo aumento da expressão do KLF4 e pela redução da expressão do SOX4, XIAP, TGF-β e CXCL12. O aumento da expressão do COL3A1 e a redução da expressão de E-CAD podem estar relacionados ao aumento do crescimento de colônias independente de ancoragem observado. Além disto, a redução da expressão de E-CAD é responsável, também, por inibir o crescimento de colônias dependente de ancoragem, ocasionar um leve aumento no potencial migratório e invasivo, além de alterar a morfologia celular. B) O nocauteamento do LINC01133 na linhagem triplo negativa Hs578T resultou no aumento de todos os parâmetros analisados associados à malignidade tumoral: taxa de proliferação, crescimento de colônias dependente e independente de ancoragem e capacidade migratória e invasiva. O aumento da expressão de SOX4, TGF-β, SPARC e COL3A1, e a redução da expressão de KLF4 e E-CAD são os possíveis fatores responsáveis por estas modificações. Além disto, a redução da expressão de E-CAD, possivelmente, também é responsável por alterar a morfologia destas células.

### 3.6. CONCLUSÃO

Através da análise da expressão de diversos lncRNAs em um painel de linhagens celulares humanas de câncer de mama, neste presente trabalho, identificamos o lncRNA LINC01133 e analisamos o seu papel na progressão do câncer de mama triplo negativo. Para tanto, superexpressamos, de forma induzível, o *LINC01133* na linhagem não-tumorigênica MCF10A e o nocauteamos, via sistema CRISPR/Cas9, na linhagem triplo negativa Hs578T. Através da caracterização funcional das linhagens geradas, descrevemos seus possíveis papéis na progressão deste subtipo tumoral de câncer de mama. Com base nos dados obtidos, sugerimos que o elevado nível de expressão do LINC01133 observado na linhagem Hs578T é importante na regulação negativa de processos relacionados com a progressão tumoral, tais como o potencial tumorigênico e metastático, atuando com um lncRNA supressor tumoral. Juntos, nossos resultados contribuem para o enriquecimento de informações relacionadas a Biologia do câncer de mama triplo negativo e sugerem que o LINC01133 pode servir com um novo biomarcador de prognóstico e um potencial alvo terapêutico para o tratamento do câncer de mama triplo negativo, auxiliando, desta forma, no desenvolvimento de potenciais protocolos clínicos e terapêuticos.

### 3.7. PRÓXIMOS PASSOS

Os próximos passos abaixo são importantes para complementação dos resultados obtidos neste trabalho:

- Repetir o experimento de curva de crescimento no conjunto de células relacionadas ao nocauteamento do *LINC01133*.
- Repetir a análise da expressão de alguns dos genes envolvidos na malignidade tumoral, os quais não tiveram a expressão detectada ou apresentaram padrões de expressão divergentes entre as linhagens geradas.
- Analisar, por sequenciamento, se o nocauteamento do *LINC01133* na linhagem Hs578T afetou possíveis sítios *off-targets*.
- Analisar, nas linhagens geradas, a expressão das metaloproteinases de matriz (MMP-2 e MMP-9) e seus inibidores.
- Analisar, nas linhagens geradas, a expressão dos genes *KLF4* e *SOX4*, os quais são possíveis alvos regulados pelo *LINC01133*.
- Realizar reações de imunofluorescência utilizando marcadores de transição epitélio-mesenquimal, no intuito de identificar com precisão possíveis alterações morfológicas nas linhagens geradas.

- Realizar ensaios de tumorigênese *in vivo* utilizando as células geradas, avaliando parâmetros histopatológicos nos possíveis tumores gerados.
- Analisar, por *Western-Blot* e RNA-Seq, o mecanismo de ação do LINC01133 na progressão tumoral do câncer de mama triplo negativo.
- Analisar a expressão do *LINC01133* em amostras de pacientes normais e portadoras do câncer de mama triplo negativo.

#### 4. CONCLUSÕES GERAIS

Dentre todos os subtipos do câncer de mama, o triplo negativo é o que apresenta o pior prognóstico devido ao seu comportamento clínico altamente agressivo. Mesmo com o avanço dos estudos e das metodologias de classificação, a taxa de mortalidade observada nestes tumores é extremamente elevada, ocasionada devido à alta incidência de recidivas e metástases, associada à escassez de tratamentos direcionados eficazes [Rusness *et al.*, 2017]. Diante disto, este grupo de tumores tem sido considerado um desafio para a clínica e a busca por novos marcadores, sejam eles proteicos ou não proteicos, que possam ser utilizados no diagnóstico, no prognóstico ou como potenciais alvos terapêuticos é de extrema importância, tanto para auxiliar na identificação de pacientes portadoras deste subtipo tumoral quanto para auxiliar na redução de suas altas taxas de mortalidade.

Na primeira parte deste trabalho, mostramos que a baixa expressão do gene codificador de proteína *CD14*, naturalmente observada na linhagem triplo negativa Hs578T, exibe um papel importante na regulação de processos relacionados com a progressão tumoral, tais como o potencial tumorigênico e metastático. O *CD14* atuaria como um oncogene, contribuindo para a progressão do câncer do estágio primário para o estágio avançado através da ativação da via de sinalização TLR2/4/6-MyD88-IRAK1. Atuando como um co-receptor para padrões moleculares associados a patógenos (PAMPs), como o lipopolissacarídeo bacteriano (LPS), ou para padrões moleculares associados ao dano (DAMPs), como alguns proteoglicanos presentes na matriz extracelular (como, por exemplo, biglican e vesican), o *CD14* levaria a ativação desta via, a qual é responsável por aumentar a expressão de citocinas pró-inflamatórias. O aumento da expressão destas citocinas induziria a quadros de inflamação crônica, favorecendo a progressão tumoral. Diante disto, sugerimos um

papel importante da via de sinalização CD14-TLR2/4/6-MyD88-IRAK1 na progressão do TNBC, tornando-a um potencial alvo terapêutico para o tratamento deste subtipo tumoral.

Na segunda parte deste trabalho, identificamos o lncRNA LINC01133 e analisamos seu papel na progressão do câncer de mama triplo negativo (TNBC). Elevados níveis de expressão do *LINC01133* foram observados na linhagem triplo negativa Hs578T, sendo importantes na regulação negativa de processos relacionados com a progressão tumoral, tais como o potencial tumorigênico e metastático, atuando com um lncRNA supressor tumoral. É importante ressaltar que, pela primeira vez, o nocauteamento deste lncRNA, via sistema CRISPR/Cas9, foi realizado em linhagens celulares humanas de câncer, em nosso caso, especificamente em linhagens celulares de câncer de mama triplo negativo. Nossos resultados mostraram que o LINC01133 desempenha um papel importante na regulação da expressão de genes envolvidos na malignidade tumoral, tais como *KLF4*, *SOX4*, *TGF- $\beta$* , *SPARC*, *COL3A1* e *E-CAD*. Diante disto, sugerimos que o LINC01133 pode servir com um novo biomarcador de prognóstico e um potencial alvo terapêutico para o tratamento do câncer de mama triplo negativo.

Assim, em suma, os dados de caracterização funcional obtidos aqui, em conjunto com os mecanismos hipotéticos de atuação propostos para estas moléculas (os quais necessitam ser melhor investigados), adicionarão informações pioneiras e complementares aos dados já descritos na literatura referentes ao câncer de mama triplo negativo, contribuindo para o enriquecimento de informações relacionadas à Biologia deste subtipo tumoral e , auxiliando, desta forma, no futuro desenvolvimento de potenciais protocolos clínicos e terapêuticos.

## 5. REFERÊNCIAS BIBLIOGRÁFICAS

- ADAMI, H.; HUNTER, D. J.; TRICHOPOULOS, D. (Ed.). **Textbook of cancer epidemiology**. Oxford University Press, USA, 2008.
- ADAMSON, E. D. Oncogenes in development. **Development**, v. 99, n.4, p. 449-471, 1987.
- AJCC. Breast. In: American Joint Committee on Cancer Cancer. **Staging Manual – 8th ed.** New York, NY: Springer, p. 589, 2017.
- AL-HAJJ, M. et al. Therapeutic implications of cancer stem cells. **Current opinion in genetics & development**, v. 14, n. 1, p. 43-47, 2004.
- ALLRED, D. C.; MOHSIN, S. K.; FUQUA, S. A. Histological and biological evolution of human premalignant breast disease. **Endocrine-related cancer**, v.8, n. 1, p. 47-61, 2001.
- ALONZO, T. A. Standards for reporting prognostic tumor marker studies. **Journal of clinical oncology**, v. 23, n. 36, p. 9053-9054, 2005.
- ANAS, A.; VAN DER POLL, T.; DE VOS, A. F. Role of CD14 in lung inflammation and infection. **Yearbook of Intensive Care and Emergency Medicine 2010**, p. 129-140, 2010.
- ANDRES, R. et al. Association of BRCA1 germline mutations in young onset triple-negative breast cancer (TNBC). **Clinical and Translational Oncology**, v. 16, n.3, p. 280-284, 2014.
- ANOTHASINTAWEE, T. et al. Risk factors of breast cancer: a systematic review and meta-analysis. **Asia Pacific Journal of Public Health**, v. 25, n. 5, p. 368-387, 2013.
- APONTE, P. M.; CAICEDO, A. Stemness in cancer: stem cells, cancer stem cells, and their microenvironment. **Stem cells international**, v. 2017, 2017.
- ASSELIN-LABAT, M. et al. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. **Nature cell biology**, v.9, n. 2, p. 201-209, 2007.
- ASSELIN-LABAT, M. et al. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. **Molecular and cellular biology**, v. 31, n. 22, p. 4609-4622, 2011.
- AZADEH, M. et al. Decreased expression level of long non-coding RNA CCAT1, was observed in breast cancer tissue of an Isfahanian population (Iran). **Gene Reports**, v. 23, p. 101154, 2021.
- BACH, K. et al. Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. **Nature communications**, v.8, n. 1, p. 1-11, 2017.
- BADVE, S. et al. Basal-like and triple-negative breast cancers: a critical review with an emphasis on the implications for pathologists and oncologists. **Modern Pathology**, v. 24, n. 2, p. 157-167, 2011.
- BAGATI, A. et al. Integrin  $\alpha\beta 6$ -TGF $\beta$ -SOX4 Pathway Drives Immune Evasion in Triple-Negative Breast Cancer. **Cancer Cell**, v. 39, n. 1, p. 54-67, 2021.
- BARRAS, D.; SAVAS, Sevtap. Biomarkers for colon cancer. 2015.
- BARTON, V. N. et al. Androgen receptor biology in triple negative breast cancer: a case for classification as AR+ or quadruple negative disease. **Hormones and cancer**, v. 6, n. 5, p. 206-213, 2015.
- BASIC, M. et al. Loss of CD14 leads to disturbed epithelial-B cell crosstalk and impairment of the intestinal barrier after *E. coli* Nissle monoassociation. **Scientific reports**, v. 8, n. 1, p. 1-14, 2018.
- BATTULA, V. L. et al. Ganglioside GD2 identifies breast cancer stem cells and promotes tumorigenesis. **The Journal of clinical investigation**, v. 122, n. 6, p. 2066-2078, 2012.
- BERTUCCI, F. et al. Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. **Cancer research**, v. 66, n. 9, p. 4636-4644, 2006.

- BERTUCCI, F. et al. How basal are triple-negative breast cancers? **International journal of Cancer**, v.123, n.1, p. 236-240, 2008.
- BHATTACHARYYA, S. et al. Toll-like receptor 4 signaling augments transforming growth factor- $\beta$  responses: a novel mechanism for maintaining and amplifying fibrosis in scleroderma. **The American journal of pathology**, v. 182, n. 1, p. 192-205, 2013.
- BIANCHINI, G. et al. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. **Nature reviews Clinical oncology**, v. 13, n. 11, p. 674-690, 2016.
- BIDA, O. et al. A novel mitosis-associated lncRNA, MA-linc1, is required for cell cycle progression and sensitizes cancer cells to Paclitaxel. **Oncotarget**, v. 6, n. 29, p. 27880, 2015.
- BOROWICZ, S. et al. The soft agar colony formation assay. **Journal of visualized experiments: JoVE**, n. 92, 2014.
- BOYD, N. F. et al. Mammographic densities and breast cancer risk. **Cancer Epidemiology and Prevention Biomarkers**, v. 7, n. 12, p. 1133-1144, 1998.
- BOYER, B.; VALLÉS, A. M.; EDME, N. Induction and regulation of epithelial–mesenchymal transitions. **Biochemical pharmacology**, v. 60, n. 8, p. 1091-1099, 2000.
- BRABLETZ, T. EMT and MET in metastasis: where are the cancer stem cells?. **Cancer cell**, v. 22, n. 6, p. 699-701, 2012.
- BRAY, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, v. 68, n. 6, p. 394-424, 2018.
- BRAY, F. et al. The ever-increasing importance of cancer as a leading cause of premature death worldwide. **Cancer**, 2021.
- BREAST CANCER ASSOCIATION CONSORTIUM. Breast cancer risk genes—association analysis in more than 113,000 women. **New England Journal of Medicine**, v. 384, n. 5, p. 428-439, 2021.
- BROWN, J. M.; WASSON, M. D.; MARCATO, P. The Missing Lnc: the potential of targeting triple-negative breast cancer and cancer stem cells by inhibiting long non-coding RNAs. **Cells**, v. 9, n. 3, p. 763, 2020.
- BURSTEIN, M. D. et al. Comprehensive genomic analysis identifies novel subtypes and targets of triple-negative breast cancer. **Clinical Cancer Research**, v. 21, n. 7, p. 1688-1698, 2015.
- CABILI, M.N. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. **Genes & development**, v. 25, n. 18, p. 1915-1927, 2011.
- CAMPOROTA, L. Interleukin-12 and the development of atopy. 2001.
- CAPARICA, R. et al. Anthracycline and taxane-based chemotherapy versus docetaxel and cyclophosphamide in the adjuvant treatment of HER2-negative breast cancer patients: a systematic review and meta-analysis of randomized controlled trials. **Breast cancer research and treatment**, v. 174, n. 1, p. 27-37, 2019.
- CHANG, H. W. et al. Loss of E-cadherin expression resulting from promoter hypermethylation in oral tongue carcinoma and its prognostic significance. **Cancer**, v. 94, n. 2, p. 386-392, 2002.
- CHAVEZ, K. J.; GARIMELLA, S. V.; LIPKOWITZ, S. Triple negative breast cancer cell lines: one tool in the search for better treatment of triple negative breast cancer. **Breast disease**, v. 32, n. 1-2, p. 35, 2010.
- CHEAH, M. T. et al. CD14-expressing cancer cells establish the inflammatory and proliferative tumor microenvironment in bladder cancer. **Proceedings of the National Academy of Sciences**, v. 112, n. 15, p. 4725-4730, 2015.

- CHEANG, M. C. et al. Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. **JNCI: Journal of the National Cancer Institute**, v. 101, n. 10, p. 736-750, 2009.
- CHEN, A. et al. E-cadherin loss alters cytoskeletal organization and adhesion in non-malignant breast cells but is insufficient to induce an epithelial-mesenchymal transition. **BMC cancer**, v. 14, n. 1, p. 1-14, 2014.
- CHEUNG, M. et al. Roles of Sox4 in central nervous system development. **Molecular brain research**, v. 79, n. 1-2, p. 180-191, 2000.
- CHEUNG, S. Y. et al. Role of epithelial–mesenchymal transition markers in triple-negative breast cancer. **Breast cancer research and treatment**, v. 152, n. 3, p. 489-498, 2015.
- CHO, N. Molecular subtypes and imaging phenotypes of breast cancer. **Ultrasonography**, v. 35, n. 4, p. 281, 2016.
- COATES, A. S. et al. Tailoring therapies—improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. **Annals of oncology**, v. 26, n. 8, p. 1533-1546, 2015.
- COCCO, S. et al. Biomarkers in triple-negative breast cancer: State-of-the-art and future perspectives. **International journal of molecular sciences**, v. 21, n. 13, p. 4579, 2020.
- COSTA, F. F. Non-coding RNAs: meet thy masters. **Bioessays**, v. 32, n. 7, p. 599-608, 2010.
- COURNOYER, S. et al. Genotype analysis of tumor-initiating cells expressing CD133 in neuroblastoma. **Genes, Chromosomes and Cancer**, v. 51, n. 8, p. 792-804, 2012.
- D’ALESSANDRO, G. et al. BRCA2 controls DNA: RNA hybrid level at DSBs by mediating RNase H2 recruitment. **Nature communications**, v. 9, n. 1, p. 1-17, 2018.
- DA COSTA VIEIRA, R. A. et al. Breast cancer screening in developing countries. **Clinics**, v. 72, p. 244-253, 2017.
- DA SILVA, J. L. et al. Triple negative breast cancer: A thorough review of biomarkers. **Critical reviews in oncology/hematology**, v. 145, p. 102855, 2020.
- DAGVADORJ, J. et al. Lipopolysaccharide induces alveolar macrophage necrosis via CD14 and the P2X7 receptor leading to interleukin-1 $\alpha$  release. **Immunity**, v. 42, n. 4, p. 640-653, 2015.
- DE ALMEIDA, J. R. C. et al. Marcadores tumorais: revisão de literatura. **Revista Brasileira de Cancerologia**, v. 53, n. 3, p. 305-316, 2007.
- DEBNATH, J. ; MUTHUSWAMY, S. K.; BRUGGE, J. S. Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. **Methods**, v. 30, n. 3, p. 256-268, 2003.
- DENT, R. et al. Triple-negative breast cancer: clinical features and patterns of recurrence. **Clinical cancer research**, v. 13, n. 15, p. 4429-4434, 2007.
- DERRIEN, T. et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. **Genome research**, v. 22, n. 9, p. 1775-1789, 2012
- DEVAUX, Y. et al. Long noncoding RNAs in cardiac development and ageing. **Nat Rev Cardiol**, v. 12, n. 7, p. 415-425, 2015.
- DING, M. et al. Association between transforming growth factor- $\beta$ 1 expression and the clinical features of triple negative breast cancer. **Oncology letters**, v. 11, n. 6, p. 4040-4044, 2016.
- DJEBALI, S. et al. Landscape of transcription in human cells. **Nature**, v. 489, n. 7414, p. 101-108, 2012.

- DONG, Y. et al. Long non-coding RNA HULC interacts with miR-613 to regulate colon cancer growth and metastasis through targeting RTKN. **Biomedicine & Pharmacotherapy**, v. 109, p. 2035-2042, 2019.
- DU, L. et al. Loss of SIRT4 promotes the self-renewal of Breast Cancer Stem Cells. **Theranostics**, v. 10, n. 21, p. 9458, 2020.
- DUBRIDGE, R. B. et al. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Molecular and cellular biology*, v. 7, n. 1, p. 379-387, 1987.
- DÜNDAR, P. E. et al. The knowledge and attitudes of breast self-examination and mammography in a group of women in a rural area in western Turkey. **BMC cancer**, v. 6, n. 1, p. 1-9, 2006.
- EDITION, S. et al. AJCC cancer staging manual. **AJCC cancer staging manual**, 2017.
- ELDEEN, M. A. et al. Molecular Targeting of Breast Cancer Stem Cells: A Promising Strategy For Management And Eradication. **European Journal of Molecular & Clinical Medicine**, v. 8, n. 1, p. 1756-1768, 2021.
- ELIYATKIN, N. et al. Molecular classification of breast carcinoma: from traditional, old-fashioned way to a new age, and a new way. **The journal of breast health**, v. 11, n. 2, p. 59, 2015.
- ELSTON, C. W.; ELLIS, I. O. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. **Histopathology**, v. 19, n. 5, p. 403-410, 1991.
- ELSTON, E. W.; ELLIS, I. O. Method for grading breast cancer. **Journal of clinical pathology**, v. 46, n. 2, p. 189, 1993.
- ENGELMANN, K.; SHEN, H.; FINN, O. J. MCF7 side population cells with characteristics of cancer stem/progenitor cells express the tumor antigen MUC1. **Cancer research**, v. 68, n. 7, p. 2419-2426, 2008.
- ENGQVIST, H. et al. Immunohistochemical validation of COL3A1, GPR158 and PITHD1 as prognostic biomarkers in early-stage ovarian carcinomas. **BMC cancer**, v. 19, n. 1, p. 1-12, 2019.
- EWING, B. et al. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. *Genome research*, v. 8, n. 3, p. 175-185, 1998.
- EWING, B.; GREEN, P. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome research*, v. 8, n. 3, p. 186-194, 1998.
- FATICA, A.; BOZZONI, I. Long non-coding RNAs: new players in cell differentiation and development. **Nature Reviews Genetics**, v. 15, n. 1, p. 7-21, 2014.
- FENG, Y. et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. **Genes & diseases**, v. 5, n. 2, p. 77-106, 2018.
- FILIPPONI, D.; BULAVIN, D. V. Wip1 and ATM in tumor evolution: role for BRCA1. **Oncotarget**, v. 4, n. 12, p. 2170, 2013.
- FOULKES, W. D. et al. Germline BRCA1 mutations and a basal epithelial phenotype in breast cancer. **JNCI: Journal of the National Cancer Institute**, v. 95, n. 19, p. 1482-1485, 2003.
- FOULKES, W. D.; SMITH, I. E.; REIS-FILHO, J. S. Triple-negative breast cancer. **New England journal of medicine**, v. 363, n. 20, p. 1938-1948, 2010.
- FRAGOMENI, S. M.; SCIALLIS, A.; JERUSS, J. S. Molecular subtypes and local-regional control of breast cancer. **Surgical Oncology Clinics**, v. 27, n. 1, p. 95-120, 2018.
- GAO, J. J.; SWAIN, S. M. Luminal a breast cancer and molecular assays: a review. **The oncologist**, v. 23, n. 5, p. 556, 2018.

- GIBB, E. A.; BROWN, C. J.; LAM, W. L. The functional role of long non-coding RNA in human carcinomas. **Molecular cancer**, v. 10, n. 1, p. 1-17, 2011.
- GIBSON, G. R. et al. Metaplastic breast cancer: clinical features and outcomes. **The American surgeon**, v. 71, n. 9, p. 725-730, 2005.
- GINESTIER, C. et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. **Cell stem cell**, v. 1, n. 5, p. 555-567, 2007.
- GINGERAS, T. R. Origin of phenotypes: genes and transcripts. **Genome research**, v. 17, n. 6, p. 682-690, 2007.
- GLUZ, O. et al. Triple-negative breast cancer—current status and future directions. **Annals of Oncology**, v. 20, n. 12, p. 1913-1927, 2009.
- GOLDHIRSCH, A. et al. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. **Annals of oncology**, v. 24, n. 9, p. 2206-2223, 2013.
- GRANGE, C. et al. Sca-1 identifies the tumor-initiating cells in mammary tumors of BALB-neuT transgenic mice. **Neoplasia**, v. 10, n. 12, p. 1433-1443, 2008.
- GRANUCCI, F, and, IVAN. Z. "Role of CD14 in host protection against infections and in metabolism regulation." *Frontiers in cellular and infection microbiology* 3 (2013): 32.
- GREAVES, M.; MALEY, C. C. Clonal evolution in cancer. **Nature**, v. 481, n. 7381, p. 306-313, 2012.
- GREENUP, R. et al. Prevalence of BRCA mutations among women with triple-negative breast cancer (TNBC) in a genetic counseling cohort. **Annals of surgical oncology**, v. 20, n. 10, p. 3254-3258, 2013.
- GUPTA, R. A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. **Nature**, v. 464, n. 7291, p. 1071-1076, 2010.
- HACHIM, I. Y. et al. A dual prognostic role for the TGF $\beta$  receptors in human breast cancer. **Human pathology**, v. 57, p. 140-151, 2016.
- HAFPTY, B. G. et al. Locoregional relapse and distant metastasis in conservatively managed triple negative early-stage breast cancer. **Journal of clinical oncology**, v. 24, n. 36, p. 5652-5657, 2006.
- HAN, C. et al. CCAT1 promotes triple-negative breast cancer progression by suppressing miR-218/ZFX signaling. **Ageing (Albany NY)**, v. 11, n. 14, p. 4858, 2019.
- HAN, M. et al. lncRNA GHET1 knockdown suppresses breast cancer activity in vitro and in vivo. **American journal of translational research**, v. 11, n. 1, p. 31, 2019.
- HANAHAN, D.; WEINBERG, R. A. Hallmarks of cancer: the next generation. **cell**, v. 144, n. 5, p. 646-674, 2011.
- HANSJI, H. et al. Keeping abreast with long non-coding RNAs in mammary gland development and breast cancer. **Frontiers in genetics**, v. 5, p. 379, 2014.
- HASSIOTOU, F.; GEDDES, D. Anatomy of the human mammary gland: Current status of knowledge. **Clinical anatomy**, v. 26, n. 1, p. 29-48, 2013.
- HAUPTMAN, N.; GLAVAČ, D. Long non-coding RNAs in cancer: Diagnostic and prognostic value. **Med Chem (Los Angeles)**, v. 6, p. 422-4, 2016.
- HE, W. et al. Serum soluble CD14 is a potential prognostic indicator of recurrence of human breast invasive ductal carcinoma with Her2-enriched subtype. **PLoS one**, v. 8, n. 9, p. e75366, 2013.
- HE, Y. Y. et al. The human milk oligosaccharide 2'-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation. **Gut**, v. 65, n. 1, p. 33-46, 2016.

- HEILBORN, J. D. et al. Antimicrobial protein hCAP18/LL-37 is highly expressed in breast cancer and is a putative growth factor for epithelial cells. **International journal of cancer**, v. 114, n. 5, p. 713-719, 2005.
- HENRY, N. L.; HAYES, D. F. Cancer biomarkers. **Molecular oncology**, v. 6, n. 2, p. 140-146, 2012.
- HENRY, N. L.; CANNON-ALBRIGHT, L. A. Breast cancer histologic subtypes show excess familial clustering. **Cancer**, v. 125, n. 18, p. 3131-3138, 2019.
- HOLLIDAY, D. L.; SPEIRS, V. Choosing the right cell line for breast cancer research. **Breast cancer research**, v. 13, n. 4, p. 1-7, 2011.
- HOU, P. et al. LincRNA-ROR induces epithelial-to-mesenchymal transition and contributes to breast cancer tumorigenesis and metastasis. **Cell death & disease**, v. 5, n. 6, p. e1287-e1287, 2014.
- HU, W.; ALVAREZ-DOMINGUEZ, J. R.; LODISH, H. F. Regulation of mammalian cell differentiation by long non-coding RNAs. **EMBO reports**, v. 13, n. 11, p. 971-983, 2012.
- HUANG, C.S. et al. The C/EBP $\beta$ -LINC01133 axis promotes cell proliferation in pancreatic ductal adenocarcinoma through upregulation of CCNG1. **Cancer letters**, v. 421, p. 63-72, 2018.
- HUANG, R.Y.J.; GUILFORD, P.; THIERY, J. P. Early events in cell adhesion and polarity during epithelial-mesenchymal transition. **Journal of cell science**, v. 125, n. 19, p. 4417-4422, 2012.
- HUANG, X. et al. LncRNAs in pancreatic cancer. **Oncotarget**, v. 7, n. 35, p. 57379, 2016.
- HWANG-VERSLUES, W. W, et al. Multiple lineages of human breast cancer stem/progenitor cells identified by profiling with stem cell markers. **PLoS One**, v. 4, n. 12, p. e8377, 2009.
- INCA. Instituto Nacional de Câncer José Alencar Gomes da Silva. Estimativa 2014. 2014. Disponível em: <[https://www.inca.gov.br/bvscontrolecancer/publicacoes/Estimativa\\_2014.pdf](https://www.inca.gov.br/bvscontrolecancer/publicacoes/Estimativa_2014.pdf)>. Acesso em: 10 de abril de 2021.
- INCA. Instituto Nacional de Câncer José Alencar Gomes da Silva. **Estimativa 2020: incidência de câncer no Brasil** / Instituto Nacional de Câncer José Alencar Gomes da Silva. – Rio de Janeiro: INCA, p. 120, 2019.
- INCA. Instituto Nacional de Câncer José Alencar Gomes da Silva. Tipos de câncer – câncer de mama. 2021. Disponível em: <<https://www.inca.gov.br/tipos-de-cancer/cancer-de-mama>>. Acesso em: 27 de abril de 2021.
- INUMARU, L. E.; SILVEIRA, É. A.; NAVES, M. M. V. Fatores de risco e de proteção para câncer de mama: uma revisão sistemática. **Cadernos de Saúde Pública**, v. 27, p. 1259-1270, 2011.
- IYER, M. K. et al. The landscape of long noncoding RNAs in the human transcriptome. **Nature genetics**, v. 47, n. 3, p. 199-208, 2015.
- JADALIHA, M. et al. Functional and prognostic significance of long non-coding RNA MALAT1 as a metastasis driver in ER negative lymph node negative breast cancer. **Oncotarget**, v. 7, n. 26, p. 40418, 2016.
- JIN, C. et al. Reciprocal regulation of Hsa-miR-1 and long noncoding RNA MALAT1 promotes triple-negative breast cancer development. **Tumor Biology**, v. 37, n. 6, p. 7383-7394, 2016.
- JONES, C. A. et al. Reduced soluble CD14 levels in amniotic fluid and breast milk are associated with the subsequent development of atopy, eczema, or both. **Journal of allergy and clinical immunology**, v. 109, n. 5, p. 858-866, 2002.
- JONES, H. B. On a new substance occurring in the urine of a patient with mollities ossium. In: **Abstracts of the Papers Communicated to the Royal Society of London**. London: The Royal Society, 1851. p. 673-673.

- KAMANGAR, F.; DORES, G. M.; ANDERSON, W. F. Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world. **Journal of clinical oncology**, v. 24, n. 14, p. 2137-2150, 2006.
- KANCZKOWSKI, W. et al. Abrogation of TLR4 and CD14 expression and signaling in human adrenocortical tumors. **The Journal of Clinical Endocrinology & Metabolism**, v. 95, n. 12, p. E421-E429, 2010.
- KASHIWAGI, S. et al. Significance of E-cadherin expression in triple-negative breast cancer. **British journal of cancer**, v. 103, n. 2, p. 249-255, 2010.
- KIM, C. Y. et al. The LncRNA HOTAIRM1 promotes tamoxifen resistance by mediating HOXA1 expression in ER+ Breast Cancer Cells. **Journal of Cancer**, v. 11, n. 12, p. 3416, 2020.
- KIM, J. et al. Crystal Structure of CD14 and Its Implications for Lipopolysaccharide Signaling\*♦. **Journal of Biological Chemistry**, v. 280, n. 12, p. 11347-11351, 2005.
- KLAASSEN, E. M. M. et al. A systematic review of CD14 and toll-like receptors in relation to asthma in Caucasian children. **Allergy, Asthma & Clinical Immunology**, v. 9, n. 1, p. 1-10, 2013.
- KONG, J. et al. Long non-coding RNA LINC01133 inhibits epithelial–mesenchymal transition and metastasis in colorectal cancer by interacting with SRSF6. **Cancer letters**, v. 380, n. 2, p. 476-484, 2016.
- KONG, J. et al. Long noncoding RNA LINC01133 inhibits oral squamous cell carcinoma metastasis through a feedback regulation loop with GDF15. **Journal of surgical oncology**, v. 118, n. 8, p. 1326-1334, 2018.
- KORDE, L. A. et al. Multidisciplinary meeting on male breast cancer: summary and research recommendations. **Journal of Clinical Oncology**, v. 28, n. 12, p. 2114, 2010.
- LABÉTA, M. O. et al. Innate recognition of bacteria in human milk is mediated by a milk-derived highly expressed pattern recognition receptor, soluble CD14. **Journal of experimental medicine**, v. 191, n. 10, p. 1807-1812, 2000.
- LAKHANI, S. R., et al. "WHO Classification of Tumors of the Breast." (2012).
- LE DU, F. et al. Is the future of personalized therapy in triple-negative breast cancer based on molecular subtype?. **Oncotarget**, v. 6, n. 15, p. 12890, 2015.
- LECCIA, F. et al. ABCG2, a novel antigen to sort luminal progenitors of BRCA1-breast cancer cells. **Molecular cancer**, v. 13, n. 1, p. 1-13, 2014.
- LEE, J. et al. Recombinant soluble CD14 reduces severity of intramammary infection by Escherichia coli. **Infection and immunity**, v. 71, n. 7, p. 4034-4039, 2003.
- LEE, M. et al. Clinical proteomics identifies urinary CD14 as a potential biomarker for diagnosis of stable coronary artery disease. **PLoS one**, v. 10, n. 2, p. e0117169, 2015.
- LEHMANN, B. D.; PIETENPOL, J. A. Identification and use of biomarkers in treatment strategies for triple-negative breast cancer subtypes. **The Journal of pathology**, v. 232, n. 2, p. 142-150, 2014.
- LEHMANN, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. **The Journal of clinical investigation**, v. 121, n. 7, p. 2750-2767, 2011.
- LEHMANN, B. D. et al. Refinement of triple-negative breast cancer molecular subtypes: implications for neoadjuvant chemotherapy selection. **PLoS one**, v. 11, n. 6, p. e0157368, 2016.
- LI, C. et al. lncRNA HOTAIRM1 regulates cell proliferation and the metastasis of thyroid cancer by targeting Wnt10b. **Oncology reports**, v. 45, n. 3, p. 1083-1093, 2021.

- LI, H. et al. Long noncoding RNA linc00617 exhibits oncogenic activity in breast cancer. **Molecular carcinogenesis**, v. 56, n. 1, p. 3-17, 2017.
- LI, J. et al. Up-regulated lncRNA GAS5 promotes chemosensitivity and apoptosis of triple-negative breast cancer cells. **Cell Cycle**, v. 18, n. 16, p. 1965-1975, 2019.
- LI, Q. et al. Over-expressed lncRNA HOTAIRM1 promotes tumor growth and invasion through up-regulating HOXA1 and sequestering G9a/EZH2/Dnmts away from the HOXA1 gene in glioblastoma multiforme. **Journal of Experimental & Clinical Cancer Research**, v. 37, n. 1, p. 1-15, 2018.
- LI, S. et al. Long noncoding RNA GAS5 suppresses triple negative breast cancer progression through inhibition of proliferation and invasion by competitively binding miR-196a-5p. **Biomedicine & Pharmacotherapy**, v. 104, p. 451-457, 2018.
- LI, X. et al. LncRNA HOTAIRM1/HOXA1 axis promotes cell proliferation, migration and invasion in endometrial cancer. **OncoTargets and therapy**, v. 12, p. 10997, 2019.
- LIN, N. U. et al. Sites of distant recurrence and clinical outcomes in patients with metastatic triple-negative breast cancer: high incidence of central nervous system metastases. **Cancer**, v. 113, n. 10, p. 2638-2645, 2008.
- LIU, J. et al. Forkhead box C1 promoter upstream transcript, a novel long non-coding RNA, regulates proliferation and migration in basal-like breast cancer. **Molecular medicine reports**, v. 11, n. 4, p. 3155-3159, 2015.
- LIU, J. et al. E-cadherin expression phenotypes associated with molecular subtypes in invasive non-lobular breast cancer: evidence from a retrospective study and meta-analysis. **World journal of surgical oncology**, v. 15, n. 1, p. 1-13, 2017.
- LIU, Y. et al. Vascular gene expression patterns are conserved in primary and metastatic brain tumors. **Journal of neuro-oncology**, v. 99, n. 1, p. 13-24, 2010.
- LIU, Y. et al. Comprehensive transcriptome analysis identifies novel molecular subtypes and subtype-specific RNAs of triple-negative breast cancer. **Breast Cancer Research**, v. 18, n. 1, p. 1-10, 2016.
- LIU, Y. et al. "Tumor-derived exosomal long noncoding RNA LINC01133, regulated by LIU, Yang et al. Tumor-derived exosomal long noncoding RNA LINC01133, regulated by Periostin, contributes to pancreatic ductal adenocarcinoma epithelial-mesenchymal transition through the Wnt/ $\beta$ -catenin pathway by silencing AXIN2. **Oncogene**, v. 40, n. 17, p. 3164-3179, 2021.
- LLOMBART-CUSSAC, A. et al. HER2-enriched subtype as a predictor of pathological complete response following trastuzumab and lapatinib without chemotherapy in early-stage HER2-positive breast cancer (PAMELA): an open-label, single-group, multicentre, phase 2 trial. **The lancet oncology**, v. 18, n. 4, p. 545-554, 2017.
- LOBBA, A. R. M. et al. Differential expression of CD90 and CD14 stem cell markers in malignant breast cancer cell lines. **Cytometry Part A**, v. 81, n. 12, p. 1084-1091, 2012.
- LOBBA, A. R. M. et al. High CD90 (THY-1) expression positively correlates with cell transformation and worse prognosis in basal-like breast cancer tumors. **PLoS one**, v. 13, n. 6, p. e0199254, 2018.
- LOMBAERTS, M. et al. E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines. **British journal of cancer**, v. 94, n. 5, p. 661-671, 2006.
- LU, P.; WEAVER, V. M.; WERB, Zena. The extracellular matrix: a dynamic niche in cancer progression. **Journal of cell biology**, v. 196, n. 4, p. 395-406, 2012.
- LUKER, K. E. et al. Scavenging of CXCL12 by CXCR7 promotes tumor growth and metastasis of CXCR4-positive breast cancer cells. **Oncogene**, v. 31, n. 45, p. 4750-4758, 2012.

- LV. M. et al. LncRNAs as new biomarkers to differentiate triple negative breast cancer from non-triple negative breast cancer. **Oncotarget**, v. 7, n. 11, p. 13047, 2016.
- MA, L.; BAJIC, V. B.; ZHANG, Z. On the classification of long non-coding RNAs. **RNA biology**, v. 10, n. 6, p. 924-933, 2013.
- MAKKI, J. Diversity of breast carcinoma: histological subtypes and clinical relevance. **Clinical medicine insights: Pathology**, v. 8, p. CPath. S31563, 2015.
- MANI, S. A. et al. The epithelial-mesenchymal transition generates cells with properties of stem cells. **Cell**, v. 133, n. 4, p. 704-715, 2008.
- MASOOD, S. Breast cancer subtypes: morphologic and biologic characterization. **Women's Health**, v. 12, n. 1, p. 103-119, 2016.
- MERCER, T. R.; DINGER, M. E.; MATTICK, J. S. Long non-coding RNAs: insights into functions. **Nature reviews genetics**, v. 10, n. 3, p. 155-159, 2009.
- MITOBE, Y. et al. Prostate cancer-associated lncRNAs. **Cancer letters**, v. 418, p. 159-166, 2018.
- MS. Ministério da Saúde/SVS/DASIS/CGIAE/SIM - Sistema de Informação sobre Mortalidade. **Atlas On-line de Mortalidade**. Instituto Brasileiro de Geografia e Estatística – IBGE (2021). Disponível em <<https://www.inca.gov.br/app/mortalidade>>. Acesso em: 15 de abril de 2021.
- MOHAMED, M. S. et al. Inhibitors of apoptosis: clinical implications in cancer. **Apoptosis**, v. 22, n. 12, p. 1487-1509, 2017.
- MONKS, J. et al. Epithelial cells remove apoptotic epithelial cells during post-lactation involution of the mouse mammary gland. **Biology of reproduction**, v. 78, n. 4, p. 586-594, 2008.
- MORENO, C. S. SOX4: the unappreciated oncogene. In: **Seminars in cancer biology**. Academic Press, p. 57-64, 2020.
- MORRISON, B. J. et al. Breast cancer stem cells: implications for therapy of breast cancer. **Breast Cancer Research**, v. 10, n. 4, p. 1-14, 2008.
- MOURTADA-MAARABOUNI, M. et al. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. **Oncogene**, v. 28, n. 2, p. 195-208, 2009.
- MUKAI, H. Treatment strategy for HER2-positive breast cancer. **International journal of clinical oncology**, v. 15, n. 4, p. 335-340, 2010.
- MURAOKA, R. S. et al. Increased malignancy of Neu-induced mammary tumors overexpressing active transforming growth factor  $\beta$ 1. **Molecular and cellular biology**, v. 23, n. 23, p. 8691-8703, 2003.
- MUTEBI, M. et al. Breast cancer treatment: A phased approach to implementation. **Cancer**, v. 126, p. 2365-2378, 2020.
- NAGATA, T. et al. KLF4 and NANOG are prognostic biomarkers for triple-negative breast cancer. **Breast cancer**, v. 24, n. 2, p. 326-335, 2017.
- NASS, S. J. et al. Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer. **Cancer research**, v. 60, n. 16, p. 4346-4348, 2000.
- NECSULEA, A. et al. The evolution of lncRNA repertoires and expression patterns in tetrapods. **Nature**, v. 505, n. 7485, p. 635-640, 2014.
- NIETO, M. A. The ins and outs of the epithelial to mesenchymal transition in health and disease. **Annual review of cell and developmental biology**, v. 27, p. 347-376, 2011.
- NOVAK, P. et al. Cell-type specific DNA methylation patterns define human breast cellular identity. **PLoS One**, v. 7, n. 12, p. e52299, 2012.

NOVIKOVA, I. V.; HENNELLY, S. P.; SANBONMATSU, K. Y. Tackling structures of long noncoding RNAs. **International journal of molecular sciences**, v. 14, n. 12, p. 23672-23684, 2013.

OBEXER, P.; AUSSERLECHNER, M. J. X-linked inhibitor of apoptosis protein—a critical death resistance regulator and therapeutic target for personalized cancer therapy. **Frontiers in oncology**, v. 4, p. 197, 2014.

OGAWA, Y. et al. Soluble CD14 levels reflect liver inflammation in patients with nonalcoholic steatohepatitis. **PLoS One**, v. 8, n. 6, p. e65211, 2013.

OMRAM, A. R. The epidemiologic transition: a theory of the epidemiology of population change. *Bulletin of the World Health Organization*, v. 79, p. 161-170, 2001.

ONO, M. et al. Exosomes from bone marrow mesenchymal stem cells contain a microRNA that promotes dormancy in metastatic breast cancer cells. **Science signaling**, v. 7, n. 332, p. ra63-ra63, 2014.

PAIK, S. et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. **New England Journal of Medicine**, v. 351, n. 27, p. 2817-2826, 2004.

PAL, B. et al. Construction of developmental lineage relationships in the mouse mammary gland by single-cell RNA profiling. **Nature communications**, v. 8, n. 1, p. 1-14, 2017.

PANDYA, S.; MOORE, R. G. Breast development and anatomy. **Clinical obstetrics and gynecology**, v. 54, n. 1, p. 91-95, 2011.

PANZITT, K. et al. Characterization of HULC, a novel gene with striking up-regulation in hepatocellular carcinoma, as noncoding RNA. **Gastroenterology**, v. 132, n. 1, p. 330-342, 2007.

PATIÑO, R. et al. Circulating monocytes in patients with diabetes mellitus, arterial disease, and increased CD14 expression. **The American journal of cardiology**, v. 85, n. 11, p. 1288-1291, 2000.

PAYER, B.; LEE, J. T. Coupling of X-chromosome reactivation with the pluripotent stem cell state. **RNA biology**, v. 11, n. 7, p. 798-807, 2014.

PECERO, M. L.; SALVADOR-BOFILL, J.; MOLINA-PINELO, S. Long non-coding RNAs as monitoring tools and therapeutic targets in breast cancer. **Cellular Oncology**, v. 42, n. 1, p. 1-12, 2019.

PEPE, M. S. et al. Phases of biomarker development for early detection of cancer. **Journal of the National Cancer Institute**, v. 93, n. 14, p. 1054-1061, 2001.

PEPPERCORN, J.; PEROU, C. M.; CAREY, L. A. Molecular subtypes in breast cancer evaluation and management: divide and conquer. **Cancer investigation**, v. 26, n. 1, p. 1-10, 2008..

PEREIRA, H. et al. Pathological prognostic factors in breast cancer. IV: Should you be a typer or a grader? A comparative study of two histological prognostic features in operable breast carcinoma. **Histopathology**, v. 27, n. 3, p. 219-226, 1995.

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

PHILLIPS, K.; ANDRULIS, I. L.; GOODWIN, P. J. Breast carcinomas arising in carriers of mutations in BRCA1 or BRCA2: are they prognostically different? **Journal of clinical oncology**, v. 17, n. 11, p. 3653-3663, 1999.

PONTING, C. P.; BELGARD, T. Grant. Transcribed dark matter: meaning or myth?. **Human molecular genetics**, v. 19, n. R2, p. R162-R168, 2010.

PONTING, C. P.; OLIVER, P. L.; REIK, Wolf. Evolution and functions of long noncoding RNAs. **Cell**, v. 136, n. 4, p. 629-641, 2009.

- PRADO-VÁZQUEZ, G. et al. A novel approach to triple-negative breast cancer molecular classification reveals a luminal immune-positive subgroup with good prognoses. **Scientific Reports**, v. 9, n. 1, p. 1-12, 2019.
- PRAKASH, R. et al. Homologous recombination and human health: the roles of BRCA1, BRCA2, and associated proteins. **Cold Spring Harbor perspectives in biology**, v. 7, n. 4, p. a016600, 2015.
- PRAT, A. et al. Clinical implications of the intrinsic molecular subtypes of breast cancer. **The Breast**, v. 24, p. S26-S35, 2015.
- PRAT, A.; PEROU, C. M. Deconstructing the molecular portraits of breast cancer. **Molecular oncology**, v. 5, n. 1, p. 5-23, 2011.
- Prat, A. et al. "Molecular features and survival outcomes of the intrinsic subtypes within HER2-positive breast cancer." *Journal of the National Cancer Institute* 106.8 (2014): dju152.
- PRENSNER, J. R.; CHINNAIYAN, A. M. The emergence of lncRNAs in cancer biology. **Cancer discovery**, v. 1, n. 5, p. 391-407, 2011.
- PROVENZANO, E.; ULANER, Gary A.; CHIN, S. Molecular classification of breast cancer. **PET clinics**, v. 13, n. 3, p. 325-338, 2018.
- PROVENZANO, P. P. et al. Collagen reorganization at the tumor-stromal interface facilitates local invasion. **BMC medicine**, v. 4, n. 1, p. 1-15, 2006.
- PROVENZANO, P. P. et al. Collagen density promotes mammary tumor initiation and progression. **BMC medicine**, v. 6, n. 1, p. 1-15, 2008.
- PUCK, T. T.; MARCUS, P. I.; CIECIURA, S. J. Clonal growth of mammalian cells in vitro growth characteristics of colonies from single HeLa cells with and without a feeder layer. **Journal of Experimental Medicine**, v. 103, n. 2, p. 273-284, 1956.
- QUINN, J. E. et al. BRCA1 functions as a differential modulator of chemotherapy-induced apoptosis. **Cancer research**, v. 63, n. 19, p. 6221-6228, 2003.
- RAKHA, E. A. et al. Breast cancer prognostic classification in the molecular era: the role of histological grade. **Breast Cancer Research**, v. 12, n. 4, p. 1-12, 2010.
- REIS, F. J. C. Rastreamento e diagnóstico das neoplasias de ovário: papel dos marcadores tumorais. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 27, p. 222-227, 2005.
- RICCIARDELLI, C. et al. Regulation of stromal versican expression by breast cancer cells and importance to relapse-free survival in patients with node-negative primary breast cancer. **Clinical Cancer Research**, v. 8, n. 4, p. 1054-1060, 2002.
- RICHARDS, E. J. et al. Long non-coding RNAs (lncRNA) regulated by transforming growth factor (TGF)  $\beta$ : lncRNA-hit-mediated TGF $\beta$ -induced epithelial to mesenchymal transition in mammary epithelia. **Journal of Biological Chemistry**, v. 290, n. 11, p. 6857-6867, 2015.
- RIVENBARK, A. G.; O'CONNOR, S. M.; COLEMAN, W. B. Molecular and cellular heterogeneity in breast cancer: challenges for personalized medicine. **The American journal of pathology**, v. 183, n. 4, p. 1113-1124, 2013.
- ROBERTS, M. S. et al. KLF4 defines the efficacy of the epidermal growth factor receptor inhibitor, erlotinib, in triple-negative breast cancer cells by repressing the EGFR gene. **Breast Cancer Research**, v. 22, n. 1, p. 1-14, 2020.
- ROCCA, A. et al. Efficacy of endocrine therapy in relation to progesterone receptor and Ki67 expression in advanced breast cancer. **Breast cancer research and treatment**, v. 152, n. 1, p. 57-65, 2015.
- ROEDIG, H. et al. Danger matrix molecules orchestrate CD14/CD44 signaling in cancer development. In: **Seminars in cancer biology**. Academic Press, p. 31-47, 2020.

- RONSON, G. E. et al. PARP1 and PARP2 stabilize replication forks at base excision repair intermediates through Fbh1-dependent Rad51 regulation. **Nature communications**, v. 9, n. 1, p. 1-12, 2018.
- ROTTENBERG, S. et al. High sensitivity of BRCA1-deficient mammary tumors to the PARP inhibitor AZD2281 alone and in combination with platinum drugs. **Proceedings of the National Academy of Sciences**, v. 105, n. 44, p. 17079-17084, 2008.
- RUSSNES, H. G. et al. Breast cancer molecular stratification: from intrinsic subtypes to integrative clusters. **The American journal of pathology**, v. 187, n. 10, p. 2152-2162, 2017.
- RUTENBERG-SCHOENBERG, M.; SEXTON, A. N.; SIMON, M. D. The properties of long noncoding RNAs that regulate chromatin. **Annual review of genomics and human genetics**, v. 17, p. 69-94, 2016.
- SABATIER, R. et al. Claudin-low breast cancers: clinical, pathological, molecular and prognostic characterization. **Molecular cancer**, v. 13, n. 1, p. 1-14, 2014.
- SAHU, A.; SINGHAL, Udit; CHINNAIYAN, A. M. Long noncoding RNAs in cancer: from function to translation. **Trends in cancer**, v. 1, n. 2, p. 93-109, 2015.
- SAVAGE, K. I.; HARKIN, D. Paul. BRCA 1, a 'complex' protein involved in the maintenance of genomic stability. **The FEBS journal**, v. 282, n. 4, p. 630-646, 2015.
- SCHEEREN, F. A. et al. A cell-intrinsic role for TLR2–MYD88 in intestinal and breast epithelia and oncogenesis. **Nature cell biology**, v. 16, n. 12, p. 1238-1248, 2014.
- SCHEPERS, G. E.; TEASDALE, Rohan D.; KOOPMAN, Peter. Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. **Developmental cell**, v. 3, n. 2, p. 167-170, 2002.
- SEKI, E. et al. TLR4 enhances TGF- $\beta$  signaling and hepatic fibrosis. *Nature medicine*, v. 13, n. 11, p. 1324-1332, 2007.
- SEO, D.; KIM, D.; KIM, W. Long non-coding RNA linc00152 acting as a promising oncogene in cancer progression. **Genomics & informatics**, v. 17, n. 4, 2019.
- SHARMA, P. et al. Germline BRCA mutation evaluation in a prospective triple-negative breast cancer registry: implications for hereditary breast and/or ovarian cancer syndrome testing. **Breast cancer research and treatment**, v. 145, n. 3, p. 707-714, 2014.
- SHEHATA, M. et al. Phenotypic and functional characterization of the luminal cell hierarchy of the mammary gland. **Breast Cancer Research**, v. 14, n. 5, p. 1-19, 2012.
- SHI, F. et al. Long noncoding RNA highly up-regulated in liver cancer predicts unfavorable outcome and regulates metastasis by MMPs in triple-negative breast cancer. **Archives of medical research**, v. 47, n. 6, p. 446-453, 2016.
- SHI, Y.; AI, W.. Function of KLF4 in stem cell biology. **Pluripotent Stem Cells**, 2013.
- SHI, Y. et al. Reduced expression of METTL3 promotes metastasis of triple-negative breast cancer by m6A methylation-mediated COL3A1 up-regulation. **Frontiers in oncology**, v. 10, p. 1126, 2020.
- SHRIKHANDE, S V. et al. Silencing of X-linked inhibitor of apoptosis (XIAP) decreases gemcitabine resistance of pancreatic cancer cells. **Anticancer research**, v. 26, n. 5A, p. 3265-3273, 2006.
- SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2016. **CA: a cancer journal for clinicians**, v. 66, n. 1, p. 7-30, 2016.
- SILVA, M. M.; SILVA, V. H. Envelhecimento: importante fator de risco para o câncer. **Arquivos médicos do ABC**, v. 30, n. 1, 2005.

SKANDALIS, S. S. et al. Versican but not decorin accumulation is related to malignancy in mammographically detected high density and malignant-appearing micro calcifications in non-palpable breast carcinomas. **BMC cancer**, v. 11, n. 1, p. 1-13, 2011.

SNIJDERS, B. E. P. et al. Cytokines and soluble CD14 in breast milk in relation with atopic manifestations in mother and infant (KOALA Study). **Clinical & Experimental Allergy**, v. 36, n. 12, p. 1609-1615, 2006.

SONG, Z. et al. LINC01133 inhibits breast cancer invasion and metastasis by negatively regulating SOX4 expression through EZH2. **Journal of cellular and molecular medicine**, v. 23, n. 11, p. 7554-7565, 2019.

SØRLIE, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. **Proceedings of the National Academy of Sciences**, v. 98, n. 19, p. 10869-10874, 2001.

SORLIE, T. et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 14, p. 8418-8423, 2003.

SOULE, H. D. et al. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. **Cancer research**, v. 50, n. 18, p. 6075-6086, 1990.

SPIZZO, R. et al. Long non-coding RNAs and cancer: a new frontier of translational research? **Oncogene**, v. 31, n. 43, p. 4577-4587, 2012.

STEIN, T. et al. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. **Breast Cancer Research**, v. 6, n. 2, p. 1-17, 2004.

SU, X. et al. Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. **Oncotarget**, v. 5, n. 20, p. 9864, 2014.

SU, Y. et al. Circulating long noncoding RNA as a potential target for prostate cancer. **International Journal of Molecular Sciences**, v. 16, n. 6, p. 13322-13338, 2015.

SUNG, H. et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: a cancer journal for clinicians**, v. 71, n. 3, p. 209-249, 2021.

SZYMICZEK, A.; LONE, A.; AKBARI, M. R. Molecular intrinsic versus clinical subtyping in breast cancer: A comprehensive review. **Clinical Genetics**, v. 99, n. 5, p. 613-637, 2021.

TADDEI, M. L. et al. Anoikis: an emerging hallmark in health and diseases. **The Journal of pathology**, v. 226, n. 2, p. 380-393, 2012.

TAKEDA, K.; AKIRA, S. Toll-like receptors in innate immunity. **International immunology**, v. 17, n. 1, p. 1-14, 2005.

TANAKA, S. et al. Diagnostic and therapeutic avenues for glioblastoma: no longer a dead end?. **Nature reviews Clinical oncology**, v. 10, n. 1, p. 14-26, 2013.

TANAKA, S. et al. Presentation, management, and outcome of newly diagnosed glioblastoma in elderly patients. **Journal of neurosurgery**, v. 118, n. 4, p. 786-798, 2013.

THARMAPALAN, P. et al. Mammary stem cells and progenitors: targeting the roots of breast cancer for prevention. **The EMBO journal**, v. 38, n. 14, p. e100852, 2019.

TIAN, T. et al. Identification of long non-coding RNA signatures in triple-negative breast cancer. **Cancer cell international**, v. 18, n. 1, p. 1-10, 2018.

- TISCORNIA, G.; SINGER, O.; VERMA, I. M. Production and purification of lentiviral vectors. *Nature protocols*, v. 1, n. 1, p. 241-245, 2006.
- TROMBETTA, L. M. et al. Isolation and characterization of novel RECK tumor suppressor gene splice variants. ***Oncotarget***, v. 6, n. 32, p. 33120, 2015.
- TSANG, J.; TSE, G. M. Molecular classification of breast cancer. ***Advances in anatomic pathology***, v. 27, n. 1, p. 27-35, 2020.
- TU, Z. et al. Micro environmental regulation of long noncoding RNA LINC01133 promotes cancer stem cell-like phenotypic traits in triple-negative breast cancers. ***Stem Cells***, v. 37, n. 10, p. 1281-1292, 2019.
- VAN DE VIJVER, M. J. et al. A gene-expression signature as a predictor of survival in breast cancer. ***New England Journal of Medicine***, v. 347, n. 25, p. 1999-2009, 2002.
- VAN GREMBERGEN, O. et al. Portraying breast cancers with long noncoding RNAs. ***Science advances***, v. 2, n. 9, p. e1600220, 2016.
- VANDESOMPELE, J. et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. ***Genome biology***, v. 3, n. 7, p. 1-12, 2002.
- VANGIPURAM, S. D. et al. Wnt pathway activity confers chemoresistance to cancer stem-like cells in a neuroblastoma cell line. ***Tumor biology: the journal of the International Society for Oncodevelopmental Biology and Medicine***, v. 33 p. 2173-83, 2012.
- VAN'T VEER, L. J. et al. Gene expression profiling predicts clinical outcome of breast cancer. ***Nature***, v. 415, n. 6871, p. 530-536, 2002.
- VASSILOPOULOS, A. et al. A critical role of CD29 and CD49f in mediating metastasis for cancer-initiating cells isolated from a Brca1-associated mouse model of breast cancer. ***Oncogene***, v. 33, n. 47, p. 5477-5482, 2014.
- VELAZQUEZ-BERUMEN, A. et al. Defining priority medical devices for cancer management: a WHO initiative. ***The Lancet Oncology***, v. 19, n. 12, p. e709-e719, 2018.
- VERMEULEN, L. et al. Cancer stem cells—old concepts, new insights. ***Cell Death & Differentiation***, v. 15, n. 6, p. 947-958, 2008.
- VIEIRA, A. F.; SCHMITT, F. An update on breast cancer multigene prognostic tests—emergent clinical biomarkers. ***Frontiers in medicine***, v. 5, p. 248, 2018.
- VIKRAM, R.; RAMACHANDRAN, R.; ABDUL, K. S. M. Functional significance of long non-coding RNAs in breast cancer. ***Breast cancer***, v. 21, n. 5, p. 515-521, 2014.
- VISCHIONI, B. et al. Expression and localization of inhibitor of apoptosis proteins in normal human tissues. ***Human pathology***, v. 37, n. 1, p. 78-86, 2006.
- VODUC, K. D. et al. Breast cancer subtypes and the risk of local and regional relapse. ***Journal of clinical oncology***, v. 28, n. 10, p. 1684-1691, 2010.
- VOLOVAT, S. R. et al. MiRNA and LncRNA as potential biomarkers in triple-negative breast cancer: a review. ***Frontiers in Oncology***, v. 10, 2020.
- VUONG, D. et al. Molecular classification of breast cancer. ***Virchows Archiv***, v. 465, n. 1, p. 1-14, 2014.
- WALDO, S. W. et al. Heterogeneity of human macrophages in culture and in atherosclerotic plaques. ***The American journal of pathology***, v. 172, n. 4, p. 1112-1126, 2008.
- WALL, R. et al. Enhanced host immune recognition of mastitis causing *Escherichia coli* in CD-14 transgenic mice. ***Animal Biotechnology***, v. 20, n. 1, p. 1-14, 2009.
- WANG, J. et al. Dysregulation of long non-coding RNA in breast cancer: an overview of mechanism and clinical implication. ***Oncotarget***, v. 8, n. 3, p. 5508, 2017.

- WANG, K. C.; CHANG, H. Y. Molecular mechanisms of long noncoding RNAs. **Molecular cell**, v. 43, n. 6, p. 904-914, 2011.
- WANG, N. et al. Long non-coding RNA HULC promotes the development of breast cancer through regulating LYPD1 expression by sponging miR-6754-5p. **OncoTargets and therapy**, v. 12, p. 10671, 2019.
- WANG, Q. et al. Long noncoding RNAs (lncRNAs) in triple negative breast cancer. **Journal of cellular physiology**, v. 232, n. 12, p. 3226-3233, 2017.
- WANG, Y.; JIANG, T. Understanding high grade glioma: molecular mechanism, therapy and comprehensive management. **Cancer letters**, v. 331, n. 2, p. 139-146, 2013.
- WANG, Y.; LIU, S. LncRNA GHET1 Promotes Hypoxia-Induced Glycolysis, Proliferation, and Invasion in Triple-Negative Breast Cancer Through the Hippo/YAP Signaling Pathway. **Frontiers in Cell and Developmental Biology**, v. 9, p. 566, 2021.
- WEE, P.; WANG, Z. Epidermal growth factor receptor cell proliferation signaling pathways. **Cancers**, v. 9, n. 5, p. 52, 2017.
- WEIGELT, B.; REIS-FILHO, J. S. Histological and molecular types of breast cancer: is there a unifying taxonomy? **Nature reviews Clinical oncology**, v. 6, n. 12, p. 718-730, 2009.
- WEIGELT, B.; PETERSE, J. L.; VAN'T VEER, L. J. Breast cancer metastasis: markers and models. **Nature reviews cancer**, v. 5, n. 8, p. 591-602, 2005.
- WEINSTEIN, I. B.; JOE, A. K. Mechanisms of disease: oncogene addiction—a rationale for molecular targeting in cancer therapy. **Nature clinical practice Oncology**, v. 3, n. 8, p. 448-457, 2006.
- WEISS, A. et al. Validation study of the American Joint Committee on Cancer eighth edition prognostic stage compared with the anatomic stage in breast cancer. **JAMA oncology**, v. 4, n. 2, p. 203-209, 2018.
- WENG, Y. C. et al. Long non-coding RNA LINC01133 silencing exerts anti-oncogenic effect in pancreatic cancer through the methylation of DKK1 promoter and the activation of Wnt signaling pathway. **Cancer biology & therapy**, v. 20, n. 3, p. 368-380, 2019.
- WILUSZ, J. E.; SUNWOO, H.; SPECTOR, D. L. Long noncoding RNAs: functional surprises from the RNA world. **Genes & development**, v. 23, n. 13, p. 1494-1504, 2009.
- World Health Organization (WHO). Breast Cancer. (2021). Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/breast-cancer>>. Acesso em: 17 de abril de 2021.
- World Health Organization (WHO). Global Health Estimates 2020: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2019. (2020). Disponível em: <[https://who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-leading-causes-of-death](https://who.int/data/gho/data/themes/mortality-and-global-health-estimates/global-health-estimates/global-leading-causes-of-death)>. Acesso em: 10 de abril de 2021.
- WRIGHT, M. H. et al. Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics. **Breast Cancer Research**, v. 10, n. 1, p. 1-16, 2008.
- WRIGHT, S. D. et al. CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. **Science**, v. 249, n. 4975, p. 1431-1433, 1990.
- WU, C. C. et al. Candidate serological biomarkers for cancer identified from the secretomes of 23 cancer cell lines and the human protein atlas. **Molecular & cellular proteomics**, v. 9, n. 6, p. 1100-1117, 2010.
- WU, W. et al. Prognostic significance of CXCL12, CXCR4, and CXCR7 in patients with breast cancer. **International journal of clinical and experimental pathology**, v. 8, n. 10, p. 13217, 2015.
- WU, Z. H. et al. Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. **Oncology reports**, v. 32, n. 1, p. 395-402, 2014.

- WU, Z. et al. CD14: Biology and role in the pathogenesis of disease. **Cytokine & growth factor reviews**, v. 48, p. 24-31, 2019.
- XIAN, H. P. et al. Circulating long non-coding RNAs HULC and ZNF1-AS1 are potential biomarkers in patients with gastric cancer. **Oncology letters**, v. 16, n. 4, p. 4689-4698, 2018.
- XIN, Q. et al. CXCR7/CXCL12 axis is involved in lymph node and liver metastasis of gastric carcinoma. **World journal of gastroenterology**, v. 23, n. 17, p. 3053, 2017.
- XIONG, G. et al. Prolyl-4-hydroxylase  $\alpha$  subunit 2 promotes breast cancer progression and metastasis by regulating collagen deposition. **BMC cancer**, v. 14, n. 1, p. 1-12, 2014.
- YADAV, B. S.; CHANANA, P.; JHAMB, S. Biomarkers in triple negative breast cancer: A review. **World journal of clinical oncology**, v. 6, n. 6, p. 252, 2015.
- YANG, F. et al. An androgen receptor negatively induced long non-coding RNA ARNILA binding to miR-204 promotes the invasion and metastasis of triple-negative breast cancer. **Cell Death & Differentiation**, v. 25, n. 12, p. 2209-2220, 2018.
- YANG, L. et al. LGR5 Promotes Breast Cancer Progression and Maintains Stem-Like Cells Through Activation of Wnt/ $\beta$ -Catenin Signaling. **Stem cells**, v. 33, n. 10, p. 2913-2924, 2015.
- YARMISHYN, A. A.; KUROCHKIN, I. V. Long noncoding RNAs: a potential novel class of cancer biomarkers. **Frontiers in genetics**, v. 6, p. 145, 2015.
- YEH, Y.C.; LIN, H.H.; TANG, M.J. A tale of two collagen receptors, integrin  $\beta$ 1 and discoidin domain receptor 1, in epithelial cell differentiation. **American Journal of Physiology-Cell Physiology**, v. 303, n. 12, p. C1207-C1217, 2012.
- YIN, L. et al. Triple-negative breast cancer molecular subtyping and treatment progress. **Breast Cancer Research**, v. 22, n. 1, p. 1-13, 2020.
- YOTOVA, I. et al. LINC01133 Inhibits Invasion and Promotes Proliferation in an Endometriosis Epithelial Cell Line. **International Journal of Molecular Sciences**, v. 22, n. 16, p. 8385, 2021.
- YOUNESS, R. A. et al. The long noncoding RNA SONE represses triple-negative breast cancer aggressiveness through inducing the expression of miR-34a, miR-15a, miR-16, and let-7a. **Journal of cellular physiology**, v. 234, n. 11, p. 20286-20297, 2019.
- YU, F. et al. Kruppel-like factor 4 (KLF4) is required for maintenance of breast cancer stem cells and for cell migration and invasion. **Oncogene**, v. 30, n. 18, p. 2161-2172, 2011.
- YU, K. D. et al. Identification of prognosis-relevant subgroups in patients with chemoresistant triple-negative breast cancer. **Clinical cancer research**, v. 19, n. 10, p. 2723-2733, 2013.
- ZENG, H. F.; QIU, H. Y.; FENG, F. B. Long noncoding RNA LINC01133 functions as an miR-422a sponge to aggravate the tumorigenesis of human osteosarcoma. **Oncology research**, v. 26, n. 3, p. 335, 2018.
- ZHANG, J. H.; LI, A. Y.; WEI, N. Downregulation of long non-coding RNA LINC01133 is predictive of poor prognosis in colorectal cancer patients. **Eur Rev Med Pharmacol Sci**, v. 21, n. 9, p. 2103-7, 2017.
- ZHANG, J. et al. SOX4 induces epithelial-mesenchymal transition and contributes to breast cancer progression. **Cancer research**, v. 72, n. 17, p. 4597-4608, 2012.
- ZHANG, J. et al. SOX4 promotes the growth and metastasis of breast cancer. **Cancer cell international**, v. 20, n. 1, p. 1-11, 2020.
- ZHANG, J.; ZHU, N.; CHEN, X. A novel long noncoding RNA LINC01133 is upregulated in lung squamous cell cancer and predicts survival. **Tumor Biology**, v. 36, n. 10, p. 7465-7471, 2015.
- ZHANG, K. et al. The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis. **Nature cell biology**, v. 15, n. 6, p. 677-687, 2013.

- ZHANG, L. et al. Androgen receptor, EGFR, and BRCA1 as biomarkers in triple-negative breast cancer: a meta-analysis. **BioMed research international**, v. 2015, 2015.
- ZHANG, M. et al. Role of transforming growth factor- $\beta$ 1 in triple negative breast cancer patients. **International Journal of Surgery**, v. 45, p. 72-76, 2017.
- ZHANG, W.; GUAN, X.; TANG, J. The long non-coding RNA landscape in triple-negative breast cancer. **Cell Proliferation**, v. 54, n. 2, p. e12966, 2021.
- ZHANG, X. et al. lncRNA AFAP1-AS1 promotes triple negative breast cancer cell proliferation and invasion via targeting miR-145 to regulate MTH1 expression. **Scientific reports**, v. 10, n. 1, p. 1-11, 2020.
- ZHANG, X. et al. FOXCUT promotes the proliferation and invasion by activating FOXC1/PI3K/AKT pathway in colorectal cancer. **Cancer Management and Research**, v. 12, p. 6269, 2020.
- ZHANG, X. F. et al. Overexpression of long non-coding RNA CCAT1 is a novel biomarker of poor prognosis in patients with breast cancer. **International journal of clinical and experimental pathology**, v. 8, n. 8, p. 9440, 2015.
- ZHAO, S. F. et al. Biglycan as a potential diagnostic and prognostic biomarker in multiple human cancers. **Oncology letters**, v. 19, n. 3, p. 1673-1682, 2020.
- ZHENG, J.; WATSON, A. D.; KERR, D. E. Genome-wide expression analysis of lipopolysaccharide-induced mastitis in a mouse model. **Infection and Immunity**, v. 74, n. 3, p. 1907-1915, 2006.
- ZHENG, J. Is SATB1 a master regulator in breast cancer growth and metastasis? **Women's Health**, v. 4, n. 4, p. 329-332, 2008.
- ZHOU, M. et al. Circulating levels of the innate and humoral immune regulators CD14 and CD23 are associated with adult glioma. **Cancer research**, v. 70, n. 19, p. 7534-7542, 2010.
- ZHU, A. et al. SPARC overexpression in primary tumors correlates with disease recurrence and overall survival in patients with triple negative breast cancer. **Oncotarget**, v. 7, n. 47, p. 76628, 2016.
- ZIEGLER-HEITBROCK, H. W. L.; ULEVITCH, R. J. CD14: cell surface receptor and differentiation marker. **Immunology today**, v. 14, n. 3, p. 121-125, 1993.

## ANEXO – Súmula Curricular

### Dados Pessoais

**Nome:** Henrique César de Jesus Ferreira

**Data de nascimento:** 16/08/1987

**Naturalidade:** Natal, Rio Grande do Norte

### Formação Acadêmica:

- 2015 - 2021** Doutorado em Ciência Biológicas (Bioquímica)  
Universidade de São Paulo, São Paulo, Brasil  
Orientadora: Proa. Dra. Mari Cleide Sogayar  
Bolsista do CNPq
- 2011 - 2014** Mestrado em Bioquímica e Biologia Molecular  
Universidade Federal do Rio Grande do Norte, Natal, Brasil  
Orientadora: Profa. Dra. Silvia Regina Batistuzzo de Medeiros  
Bolsista do CNPq
- 2010 - 2011** Graduação em Ciências Biológicas (Bacharelado)  
Universidade Federal do Rio Grande do Norte, Natal, Brasil  
Orientadora: Profa. Dra. Luciana Duarte Martins da Matta
- 2006 - 2010** Graduação em Ciências Biológicas (Licenciatura)  
Universidade Federal do Rio Grande do Norte, Natal, Brasil
- 2005 - 2008** Técnico em Controle Ambiental / Meio Ambiente  
Instituto F. de Educação, Ciência e Tecnologia do RN, Natal, Brasil

### Artigos publicados em periódicos:

Velloso, F. J., Bianco, A. F., Farias, J. O., Torres, N. E., Ferruzo, P. Y., Anschau, V., Jesus-Ferreira, H.C., .. & Correa, R. G. The crossroads of breast cancer progression: insights into the modulation of major signaling pathways. *OncoTargets and therapy*, v. 10, p. 5491, 2017.

Deocesano-Pereira, C., Machado, R. A. C., Jesus-Ferreira, H. C., Marchini, T., Pereira, T. F., Carreira, A. C. O., & Sogayar, M. C. Functional impact of the long non-coding RNA MEG3 deletion by CRISPR/Cas9 in the human triple negative metastatic Hs578T cancer cell line. *Oncology letters*, v. 18, n. 6, p. 5941-5951, 2019.

**Experiência Profissional:**

- 2019 - 2021** Vendedor Técnico na área de Cromatografia  
Easycrom Comércio e Manutenção de Equipamentos, São Paulo, Brasil
- 2015 - 2015** Professor Substituto de Bioquímica  
Universidade Federal do Rio Grande do Norte, Natal, Brasil  
Disciplinas Ministradas: Bioquímica, Bioquímica para Ecólogos
- 2014 - 2015** Professor Substituto de Anatomia Humana  
Universidade Federal do Rio Grande do Norte, Natal, Brasil  
Disciplinas Ministradas: Anatomia Humana, Anatomia XIII, Anatomia para o Curso de Farmácia, Anatomia para o Curso de Enfermagem, Anatomia para o Ensino de Ciência e Biologia

**Apresentação de Trabalhos em Eventos:**

Congresso da Sociedade Brasileira de Biologia Celular, 2018, São Paulo. Identification and functional characterization of lncRNA LINC01133 by CRISPR/Cas9 editing system in human breast cancer cell lines, 2018.

XX Encontro de Genética do Nordeste, 2014, Campina Grande - PB. Metagenoma de comunidades microbianas expostas à radiação natural e a metais, 2014.

65ª Reunião Anual da SBPC, 2013, Recife-PE. Caracterização de quitinases com importância biotecnológica extraída da lagosta *Panulirus argus*, 2013.

Congresso Brasileiro de Microbiologia, 2013, Natal-RN. Análise metagenômica inicial de comunidades microbianas expostas à radiação natural, 2013.

4th International Symposium in Biochemistry of Macromolecules and Biotechnology, 2012, Recife-PE. Identification and partial isolation of a chitinase extracted from lobsters of the specie *Panulirus argus*, 2012.

4th International Symposium in Biochemistry of Macromolecules and Biotechnology, 2012, Recife-PE. Partial characterization of a chitinase isoform obtained from the lobster *Panulirus argus*, 2012.

4th International Symposium in Biochemistry of Macromolecules and Biotechnology, 2012, Recife-PE. Isolation, microbicidal and bioinsecticide activite of chitinase from shrimp *Litopenaeus schmitti*, 2012.

XXII Congresso de Iniciação Científica da UFRN, 2011, Natal-RN. Extração, caracterização cinética e produção de oligossacarídeos, 2011.

XXI Congresso de Iniciação Científica, 2010, Natal-RN. Identificação e caracterização de glicosidases e sulfatases em músculo do camarão marinho *Litopenaeus schmitti*, 2010.

**Monitoria de disciplinas:**

- 2019** QBQ5717 – Biologia Molecular da Transformação Maligna. Pós-graduação. Departamento de Bioquímica. Instituto de Química (IQ) – USP.
- 2018** QBQ1252 – Estrutura de Biomoléculas e Metabolismo. Graduação. Departamento de Bioquímica. Instituto de Química (IQ) – USP.
- 2017** QBQ5717 – Biologia Molecular da Transformação Maligna. Pós-graduação. Departamento de Bioquímica. Instituto de Química (IQ) – USP.
- 2016** QBQ0215 – Bioquímica Metabólica. Graduação. Departamento de Bioquímica. Instituto de Química (IQ) – USP.
- 2011** Monitor da disciplina de Embriologia Humana para os cursos de Biologia, Farmácia, Fonoaudiologia, Odontologia, Medicina, Fisioterapia, Biomedicina e Nutrição. Departamento de Anatomia. Centro de Biociências – UFRN.

**Orientação e Supervisão:****Supervisor:**

1º Estágio de Inverno – Temas Avançados em Bioquímica e Biologia Molecular – Sistema de Edição Gênica CRISPR/Cas9. Departamento de Bioquímica. Instituto de Química (IQ) – USP.

**Iniciação Científica:**

Isabele Yumi Araújo Osawa – Nocauteamento do RNA Longo Não Codificador FOXCUT na linhagem humana de câncer de mama triplo negativa Hs578T através do Sistema CRISPR/Cas9 (Graduação em Química) – Junho de 2018 à Dezembro de 2018. Universidade de São Paulo.

**Outros Cursos:**

- 2017** Citometria de Fluxo e Formação de Operadores do Equipamento FACS Aria II. NUCEL, Universidade de São Paulo, USP, Brasil.
- 2016** Curso de Treinamento do Citômetro de Fluxo Accuri. NUCEL, Universidade de São Paulo, USP, Brasil.

**Prêmios e Títulos:**

- 2008** Mérito Estudantil, IFRN – Técnico em Controle Ambiental, Rio Grande do Norte.
- 2008** Bolsa de Mobilidade Estudantil (Universidade de Caxias do Sul), Universidade Federal do Rio Grande do Norte - Banco Santander.