



UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

MELISSA QUINTERO ESCOBAR

**“Transcriptômica e metabolômica como ferramentas para o
entendimento do processo de adaptação metabólica tumoral dos
tumores de mama triplo-negativo”**

Campinas

2016

MELISSA QUINTERO ESCOBAR

**“Transcriptômica e metabolômica como ferramentas para o
entendimento do processo de adaptação metabólica tumoral dos
tumores de mama triplo-negativo”**

Tese apresentada ao Instituto de
Biologia da UNICAMP para a
obtenção do título de Doutora em
ciências na área de Fármacos,
Medicamentos e Insumos para
Saúde.

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELA ALUNA MELISSA
QUINTERO ESCOBAR E ORIENTADA PELA DRA.
SANDRA MARTHA GOMES DIAS

Orientadora: Dra. Sandra Martha Gomes Dias

Co-orientadora: Dra. Ana Carolina de Mattos Zeri

Campinas

2016

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2012/09452-9

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Gustavo Lebre de Marco – CRB 8/7977

Quintero Escobar, Melissa, 1986-
Q457t Transcriptômica e metabolômica como ferramentas para o entendimento do processo de adaptação metabólica tumoral dos tumores de mama triplo negativo / Melissa Quintero Escobar. – Campinas, SP : [s.n.], 2016.

Orientador: Sandra Martha Gomes Dias.
Coorientador: Ana Carolina de Mattos Zeri.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Biologia.

1. Mamas - Câncer. 2. Câncer. 3. Expressão gênica. 4. RMN. 5. RNA-seq. 6. Metabolômica. I. Dias, Sandra Martha Gomes. II. Zeri, Ana Carolina de Mattos. III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título

Informações para Biblioteca Digital

Título em outro idioma: Transcriptomics and metabolomics of triple negative breast tumor reveal their metabolic adaptation

Palavras-chave em inglês:

Breast cancer

Cancer

Gene expression

NMR RNA-

seq

Metabolomics

Área de concentração: Fármacos, Medicamentos e Insumos para Saúde

Titulação: Doutora em ciências

Banca examinadora:

Sandra Marta Gomes Dias [Orientadora]

Juliana Helena Smetana

Katlin Bauer Massirer

Mariana Maschietto

Daniela Trivella

Data de defesa: 22-11-2016

Programa de Pós-Graduação: Biociências e tecnologia de produtos bioativos

Campinas, 22 de Novembro de 2016

COMISSÃO EXAMINADORA

Dra. Sandra Marta Gomes Dias (Orientadora)

Dra. Juliana Helena Costa Smetana

Dra. Katlin Bauer Massirer

Dra. Mariana Maschietto

Dra. Daniela Barreto Barbosa Trivella

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

AGRADECIMENTOS

Aos distantes, porém constantes em minha saudade: minha família, meus pais, irmãos, cunhadas e meus queridos sobrinhos, Alejandro, Manuel e Manuela, quem foram meu motor.

A minha orientadora, Dra. Sandra, pela oportunidade de trabalhar em seu laboratório, pelas discussões sinceras, pelos conselhos, e por tudo o que eu pude aprender enquanto estive no laboratório.

A minha coorientadora, Dra. Ana, pela oportunidade de trabalhar na ressonância, pelos conselhos, apoio nas ideias e pela motivação cientista.

À Douglas e a Larissa pela paciência e disposição para me ajudar com os mais diversos problemas. Sem eles não teria sido possível a finalização deste projeto. Muito obrigada.

Aos meus amigos do grupo da ressonância, Silvana, Mauricio, Rafael, Marcos, Pati, Malu. Quem sempre estiveram dispostos a me ajudar, pelas discussões serias e, principalmente por todo o apoio.

Aos colegas do grupo da Bioinformática, em especial ao Marcelo Falsarella, por todo o suporte e ideias novas.

Aos meus colegas e amigos do LNBio Ju, Ângela, Igor, Carol, Maruska e Kali. Obrigada pelos momentos divertidos e por todo o apoio.

Aos companheiros do grupo de metabolismo tumoral pela colaboração.

Aos amigos que já não estão mais no grupo, mas que aprendi deles também, Alexandre, Emerson, Jean e o Patrick.

Aos meus amigos, Andrea, Camilo, Camila, Carol, Cheo, Cinti, Isa, Javi, Jonathan, Miguel, Mónica, Silvinha, Tere, e aos que faltam, agradeço pelos momentos de descontração e pelo apoio nos momentos difíceis. Foram a minha família por vários anos, aprendi as melhores coisas da minha vida com vocês.

Aos meus amigos, da Colômbia. Mauricio e Wilo que de uma ou outra forma sempre estiveram presentes com palavras de conforto.

À Rafael, quem chegou no momento complicado e soube sobrelevar esses momentos de estresse dando conforto e apoio sempre.

À FAPESP pelo importante apoio financeiro com os projetos (FAPESP, 2011/02912-1 e 2012/09452-9), sem esse apoio não houvesse sido possível a realização deste trabalho.

Ao LNBio e a UNICAMP por permitir me formar como pesquisador em duas excelentes instituições. Ao CNPEM por permitir usar todas as *facilities* disponíveis, o RMN, o laboratório de bioensaios, o laboratório de vectores virais e o laboratório de bioinformática.

Ao Brasil por abrir as portas aos estrangeiros que querem sair na frente e não contam com o apoio nos seus países de origem.

Muito obrigada!

RESUMO

Grandes avanços foram feitos ao longo dos anos em nossa compreensão sobre o câncer, mas muitos detalhes relativos aos mecanismos responsáveis pelo desenvolvimento tumoral ainda não são completamente entendidos. Recentemente, a capacidade de reprogramar o metabolismo energético foi reconhecida como um dos *hallmarks* da transformação tumoral. Seu estudo e compreensão tem aberto novas e promissoras portas para a busca de alvos terapêuticos. Cerca de 10-20% dos casos de câncer de mama diagnosticados correspondem ao subtipo triplo-negativo (TN), o qual não expressa os receptores dos hormônios estrogênio e progesterona e não apresenta amplificação de HER2. Os tumores TN apresentam um perfil molecular único, comportamento agressivo e pior prognóstico em relação a outros subtipos de câncer de mama. Além, os tumores TN não respondem aos tratamentos convencionais desenvolvidos para a doença, o que ressalta a necessidade de se procurar novas estratégias terapêuticas. Tumores TN tem captação aumentada de 2-deoxi-2- (18F) fluoro-D-glicose (FDG) e dependência de glutamina, o que se correlaciona com seu alto índice proliferativo e evidencia a intensificação da glicólise aeróbica ou efeito Warburg. Neste trabalho, objetivamos analisar a expressão gênica diferencial e de metabólitos de linhagens celulares e tecidos tumorais de mama TN e não-TN de maneira a se caracterizar a assinatura metabólica do subtipo TN. Por fim, objetivamos identificar potenciais novo alvos terapêuticos para o tratamento deste subtipo tumoral. Como resultados, verificamos que tumores TN apresentam características que potencialmente lhes garantem uma alta eficiência no uso de nutrientes, com alta captação e processamento de glicose, alto processamento de glutamina, re-aproveitamento de lactato e utilização de acetaldeídos como fonte energética e biossintética. No último caso, lhe grantindo, uma potencial via de quimioresistência. Há ainda a indicação de ativação de vias de síntese/modificação de lipídeos que podem ter uma função sinalizadora nas células. Por fim, foi gerada uma lista de 73 genes com expressão aumentada em tumores de mama TN (em relação a não-TN). Desta lista, *GBP1*, uma *guanylate-binding protein*, mostrou diminuir seletivamente a proliferação celular de linhagens TN quando silenciado. Sua função para o estabelecimento/progressão/sobrevivência do subtipo TN precisa ser avaliada de maneira a confirmá-la como alvo para o tratamento desta doença.

ABSTRACT

Great advances have been made over the years in our understanding of cancer, but many details of the mechanisms involved for tumor development are not yet fully understood. Recently, the ability to reprogram the energy metabolism was recognized as one of the hallmarks of tumoral transformation. His study and understanding has opened promising new doors in the search for therapeutic targets. About 10-20% of cases of breast cancer diagnosed corresponds to triple negative (TN) subtype, which does not express the hormones receptors of estrogen and progesterone and does not have HER2 amplification. TN tumors have a unique molecular profile, an aggressive behavior and a poor prognosis compared to others breast cancer subtypes. In addition, TN tumors do not respond to conventional treatments developed for the disease, which highlights the need to look for new therapeutic strategies. TN tumors increase the uptake of 2-deoxy-2-(18F) fluoro-D-glucose (FDG) and has a glutamine dependence, which correlates with their high proliferative index and demonstrates the enhancement of aerobic glycolysis or Warburg effect. In this study, we aimed to analyze differential gene expression and metabolomics from cell lines and tumor breast tissues TN and not TN (N-TN) subtypes in order to characterize the metabolic signature. Finally, we aimed to identify potential therapeutic targets for treatment of this tumor subtype (TN). As a result, we found that TN tumors potentially have characteristics that guarantee them a high efficiency in the use of nutrients, with high uptake and processing of glucose, high glutamine processing, re-use lactate and use of acetaldehyde as energy and biosynthetic source. This last resource could be an important way of chemoresistance. There is also activation of synthesis and modification pathways of lipids that can have important signaling function in cells. Finally, we generated a list of 73 genes with increased expression in TN breast tumors (relative to non-TN) and from this list, *GBP1* a guanylate-binding protein, shown to reduce selectively the cell proliferation in TN cell lines when muted. Its function for the establishment / progression / survival TN subtype must be evaluated to confirm it or not as a potential target to treat this disease.

ÍNDICE DE FIGURAS

Figura 1. Hallmarks em câncer.....	17
Figura 2. Diferenças chaves entre o metabolismo de células normais quiescentes e células proliferativas normais ou tumorais.	23
Figura 3. Diferenças entre o metabolismo de células quiescentes diferenciadas e de células proliferativas no tocante ao destino do piruvato	24
Figura 4. Metabolismo central do carbono.	25
Figura 5. Anaplerose e cataplerose do ciclo do TCA.....	27
Figura 6. Ilustração de diferentes níveis de informação que podem ser medidos em uma célula.	30
Figura 7. Experimento típico de RNA-Seq.....	37
Figura 8. Integridade do RNA total extraído	41
Figura 9. Correlação dos dados de expressão gênica dos genes escolhidos para validação. Comparou-se os valores de $1/\Delta CT$ (obtidos pelo qPCR) com o de $\log_2 RSEM+1$ da amostra sequenciada <i>in-house</i>	42
Figura 10. Avaliação da qualidade dos <i>reads</i> da linhagem SKBR3 antes e após o <i>trimming</i>	44
Figura 11. A. Função da expressão dos genes <i>ESR1</i> , <i>PGR</i> e <i>ERBB2</i> B. <i>Status</i> da expressão dos marcadores por IHC C. <i>Status dos pacientes após reclassificação</i> . D. Porcentagem de concordância com a classificação imunohistoquímica (IHC) disponível. E. <i>Boxplots</i> da expressão gênica para cada um dos marcadores em tecidos normais, TN e N-TN.	45
Figura 12. Análise de expressão diferencial de tecidos tumorais de mama TN e não-TN. ..	46
Figura 13. Valores RSEM para os genes <i>ESR1</i> , <i>PGR</i> e <i>ERBB2</i>	47
Figura 14. Análise de expressão gênica diferencial entre linhagens celulares TN e não-TN	48
Figura 15. Interseção de genes DE entre linhagens celulares <i>versus</i> tecidos tumorais.....	49
Figura 16. Genes diferencialmente expressos em tecidos tumorais TN em relação aos não-TN.....	56
Figura 17. Genes diferencialmente expressos em linhagens tumorais TN em relação aos não-TN.....	59
Figura 18. Vetores lentivirais usados para silenciamento gênico.....	66
Figura 19. Genes <i>GBP1</i> , <i>ARL4C</i> e <i>B3GNT5</i>	68
Figura 20. qPCR dos genes selecionados em linhagens TN e não-TN.	68
Figura 21. qPCR para verificar a atenuação da expressão dos genes <i>ARL4C</i> , <i>B3GNT5</i> e <i>GBP1</i>	69
Figura 22. Porcentagem de proliferação (ou morte) de linhagens não-TN e TN após <i>knock down</i> dos alvos <i>ARL4C</i> , <i>B3GNT5</i> e <i>GBP1</i>	70
Figura 23. Porcentagem de proliferação (ou morte) de linhagens não-TN e TN após <i>knock down</i> do alvo <i>GBP1</i>	70
Figura 24. Expressão gênica de <i>GBP1</i> conforme analisado pela base de dados Oncomine	72
Figura 25. Expressão gênica de <i>GBP1</i> em tecidos de mama sadios e tecidos de câncer de mama separados entre não-TN e TN.....	72
Figura 26. Cavidades na superfície da estrutura de <i>GBP1</i> disponíveis para a interação com pequenas moléculas.	73
Figura 27. Distribuição <i>Ligand-based druggability</i> do alvo <i>GBP1</i>	74

ÍNDICE DE TABELAS

Tabela 1 Subtipos intrínsecos de câncer de mama	21
Tabela 2. Linhagens celulares utilizadas nos experimentos e seu <i>status</i> quanto à presença dos receptores de estrógeno (ER), progesterona (PR) e Her2.....	33
Tabela 3. Lista de primers utilizados para as confirmações de expressão gênica diferencial.	40
Tabela 4. <i>Reads</i> alinhados com o programa STAR (%).	43
Tabela 5. Porcentagem de sequências mantidas depois do corte com o programa <i>Skewer</i>	43
Tabela 6. <i>Pathways maps</i> mais afetados tanto em tecido quanto em linhagens TN.	50
Tabela 7. <i>Metabolic networks</i> mais afetadas tanto em tecido quanto em linhagens TN.....	51
Tabela 8 <i>Metabolic networks</i> mais afetadas nas linhagens tumorais TN <i>versus</i> não-TN (dados de expressão gênica e metabolômica).....	52
Tabela 9. Lista de <i>primers</i> utilizados para verificação da expressão gênica dos alvos em linhagens tumorais TN e não-TN.	63
Tabela 10. Sequências dos oligos para silenciamento para cada um dos alvos selecionados.....	65

Lista de abreviaturas e siglas

AKT- Proteína quinase B.

ATP - Adenosina Trifosfato.

BL- Basal - Subtipo de câncer triplo negativo.

BPTES - Bis-2- (5-fenil acetamido-1,3,4-tiadiazol-2-il) acetato de sulfureto.

CCDS- A sequência de codificação de consenso.

cDNA- DNA complementar. **DEPC**-Dietilpirocarbonato.

DHAP- Fosfato de diidroxiacetona. **DHAP**- Fosfato de diidroxiacetona.

EDTA - Ácido etilenodiamino tetra-acético.

EGFR- receptor do fator de crescimento epidérmico.

EMBL- *European Molecular Biology Laboratory*.

ER- Receptor de estrogênio.

ErBb- Família de receptores.

ERBB2- gene que codifica Her2.

FDG-PET – Tomografia por emissão de pósitrons associado à captação de 2-deoxi-2-(18F) fluoro-D-glicose.

FH- Fumarato hidratase.

GAP-Gliceraldeído3-fosfato

GDH- Glutamato desidrogenase.

GLS- Glutamina sintetase.

GLUT- Transportadores de glucose.

H-DBAS- *Human transcriptome DataBase for Alternative splicing*.

HER2 or Neu, ErbB-2, CD420 ou p185 - receptor do fator de crescimento epidérmico humano.

HIF 1-Fator de transcrição induzido por hipóxia tipo 1.

HK- Hexoquinase.

HMDB- Base de dados do metaboloma humano.

HTSF- *High Throughput Sequencing Facility* da Universidade de Carolina do Norte em Chapel Hill.

IDH1- Isocitrato desidrogenase 1.

IM- Imunomodulador -Subtipo de câncer triplo negativo.

KEGG- *Kyoto Encyclopedia of Genes and Genomes*.

KG-Alfa-cetoglutarato.

LACTAD- Laboratório Central de Tecnologias de Alto Desempenho

LAR- Receptor de andrógeno luminal - subtipo de câncer triplo negativo.

LDH-A-Lactato desidrogenase A.

LKB1- Supressor de tumor no fígado quinase B1.

M- Mesenquimal - Subtipo de câncer triplo negativo.

m/v- Massa/volume.

MAPK-Proteína-quinases ativadas por mitógenos.

MCT- Transportador monocarboxilato

mL- Mililitros.

mM – Milimolar.

MOPS- 3-N-morfolino ácido propanosulfônico.

mRNA-RNA mensageiro.

MS- Espectrometria de massas.

MSL- Subclasse do subtipo de câncer triplo negativo mesenquimal.

NADPH - Dinucleotídeo de Nicotinamida e Adenina fosfato reduzido.

NCBI- *National Center for Biotechnology information*,

N-TN- Câncer de mama não triplo negativo.

OAA- Oxaloacetato.

ONCOMINE- *Cancer Microarray Database and Integrated Data-mining Platform*.

p53- Gene supressor tumoral.

PARP-Poly (ADP-ribose) polimerases.

pb- Pares de bases.

PCR- Reação em cadeia da polimerase.

PDH-Piruvato desidrogenase.

PDK1-Piruvato quinase desidrogenase 1.

PEP- Fosfoenolpiruvato.

PHRED-*Phred quality score*.

PI3K-Fosfoinositida 3-quinase.

PKC- Proteína quinase C.

PKM2- Piruvato quinase M2.

ppm- Partes por milhão.

PPEE: Probabilidade posterior expressão genética.

PR- Receptor de progesterona.

PRA, PRB- Receptores de progesterona A e B.

RMN- Ressonância magnética nuclear.

RNA-ácido ribonucleico.

RNA-seq- Sequenciamento do RNA.

ROS- Espécies reativas de oxigênio.

RPKM- *Reads* alinhados no gene por kilobase por milhão de *reads* sequenciados na biblioteca.

RSEM-*RNA-Seq by Expectation Maximization*.

rRNA- RNA do ribossomo.

SDH-Succinato desidrogenase.

STAT-Transdutor de sinal e ativador de transcrição.

TCA - Ciclo do Ácido Tricarboxílico.

TIGAR- Regulador de apoptose.

TN- Câncer de mama triplo negativo.

tRNA- RNA de transferência.

UCSC- *Genome browser of university of California - Santa Cruz.*

UNICAMP- Universidade Estadual de Campinas.

VHL-Supressor de tumor Von Hippel-Lindau.

μL – Microlitros.

μM – Micromolar.

Sumário

1	INTRODUÇÃO	17
1.1	OS MARCOS DA TRANSFORMAÇÃO TUMORAL	17
1.2	CÂNCER DE MAMA	18
1.3	CÂNCER DE MAMA TRIPLO NEGATIVO	19
1.3.1	Subclassificação Molecular de TN Baseado no Perfil de Expressão Gênica	20
1.4	METABOLISMO TUMORAL	22
1.4.1	Efeito Warburg	22
1.4.2	Ciclo do TCA	24
1.4.3	Alterações metabólicas dos tumores TN	27
1.5	BIOLOGIA DE SISTEMAS	28
2	OBJETIVOS	31
2.1	OBJETIVO GERAL	31
2.2	OBJETIVOS ESPECÍFICOS	31
3	CAPÍTULO I: diferenças metabólicas entre tumores TN e não-TN	32
3.1	METODOLOGIA	32
3.1.1	Cultura de células	32
3.1.2	Extração de metabólitos intracelulares	33
3.1.3	Aquisição de espectros de RMN e quantificação de metabólitos	34
3.1.4	Extração de RNA	35
3.1.5	<i>Next generation sequencing</i> com Hiseq2000 (Illumina)	35
3.1.6	Obtenção dos dados de RNA-Seq de bancos públicos	37
3.1.7	Processamento dos dados de RNA-Seq	38
3.1.8	Classificação dos tecidos tumorais de mama entre TN e não-TN	38
3.1.9	Determinação da expressão gênica diferencial	39
3.1.10	Validação dos dados de RNA-Seq por qPCR	39
3.2	RESULTADOS E DISCUSSÃO	41
3.2.1	Extração de RNA e sequenciamento	41
3.2.2	Validação dos dados de RNA-Seq <i>in house</i> por qPCR	41
3.2.3	Análises da qualidade dos dados do sequenciamento	43
3.2.4	Definição da lista de genes diferencialmente expressos entre tecidos e linhagens TN versus não-TN	44

3.2.5	Análises de vias alteradas pelo Metacore.....	48
3.2.6	Análises de vias metabólicas curadas manualmente.....	53
3.3	CONCLUSÕES	60
4	CAPITULO II: Identificação e validação de potenciais novos alvos para o tratamento de tumores de mama triplo negativo	62
4.1	METODOLOGIA	62
4.1.1	Definição de genes <i>up-regulated</i> em linhagens e tecidos tumorais de mama TN em relação a não-TN	62
4.1.2	Confirmação da expressão diferencial dos genes em dados de microarranjo de tecidos tumorais de mama	62
4.1.3	Validação da expressão de genes alvos por qPCR em linhagens de tumor de mama.....	63
4.1.4	<i>Knock down</i> de genes alvos em linhagens celulares.....	63
4.1.5	Avaliação da ação silenciamento gênico sobre proliferação celular por microscopia de fluorescência.....	66
4.2	RESULTADOS E DISCUSSÃO	67
4.2.1	Confirmação da expressão dos genes alvos selecionados por qPCR ...	67
4.2.2	Análise da influência do <i>knock down</i> dos genes alvos sobre a proliferação celular	69
4.3	CONCLUSÕES	76
5	REFERENCIAS	77
6	ANEXO I	85

1. INTRODUÇÃO

1.1 OS MARCOS DA TRANSFORMAÇÃO TUMORAL

Grandes avanços foram feitos ao longo dos anos em nossa compreensão sobre diferentes doenças humanas, incluindo o câncer. No entanto, muitos detalhes relativos aos mecanismos responsáveis pelo desenvolvimento tumoral ainda não são completamente entendidos dado à complexidade das interações genéticas e ambientais que estão envolvidas no processo. Sabe-se hoje que a aquisição de várias características são necessárias para que uma célula possa se tornar cancerosa. Entre estas particularidades, conhecidas como os *hallmarks* da transformação tumoral, se destacam a auto-suficiência em sinais de crescimento, evasão de apoptose, insensibilidade aos sinais inibidores de proliferação, potencial replicativo ilimitado, angiogênese sustentada, invasão de tecidos e metástase¹. Recentemente, a capacidade de reprogramar o metabolismo energético foi reconhecida como um dos *hallmarks* da transformação tumoral² (Figura 1). Sendo um campo relativamente novo de pesquisa em biologia tumoral, o processo de reprogramação metabólica abre novas e promissoras portas para a busca de alvos terapêuticos para o tratamento do câncer. Além, conhecê-lo em detalhes é de iminente importância para a compreensão do desenvolvimento e progressão da doença.



Figura 1. Hallmarks em câncer. Adaptado de Ji Luo et al 2009.³

1.2 CÂNCER DE MAMA

O câncer de mama representa 25% de todos os cânceres em mulheres, sendo, segundo a Organização Mundial da Saúde (Agência Internacional de Investigação do Câncer), o mais frequente entre as mulheres no mundo. No Brasil, segundo o Instituto Nacional do Câncer (INCA), excluindo os tumores de pele não melanoma, o câncer de mama também é o mais incidente em mulheres de todas as regiões, exceto na região Norte, onde o câncer do colo do útero ocupa a primeira posição. A estimativa de incidência em 2016 será de 57.960 casos novos, que representam uma taxa de incidência de 56,2 casos por 100.000 mulheres. A taxa de mortalidade por câncer de mama ajustada pela população mundial apresenta uma curva ascendente e representa a primeira causa de morte por câncer na população feminina brasileira, com 12,66 óbitos/100.000 mulheres em 2013 (INCA, atlas de mortalidade). Embora as taxas de mortalidade continuem sendo muito menores do que as taxas de incidência, o câncer de mama ainda é a causa mais frequente de morte relacionada a câncer em mulheres.

Alguns fatores de risco para o desenvolvimento do câncer de mama são bem conhecidos, tais como o envelhecimento, fatores relacionados à vida reprodutiva da mulher, histórico familiar de câncer de mama, consumo de álcool, excesso de peso, sedentarismo e alta densidade do tecido mamário⁴. Além destes, fatores hormonais e reprodutivos como a menarca precoce e a menopausa tardia⁵, a nuliparidade⁶ e terapias de reposição hormonal⁷ também são considerados como fatores de risco para esta doença.

Os hormônios sexuais, como os esteroides (estrogênio e progesterona), têm um papel importante no crescimento e desenvolvimento da glândula mamária. Tem sido demonstrada uma correlação clara entre patogênese do câncer de mama e a exposição cumulativa a estrogênios. Aproximadamente 60% dos pacientes com câncer de mama estão na pré-menopausa e 75% daqueles que estão na pós-menopausa têm a doença dependente de estrógenos⁸.

Em seu papel duplo de ligante e substrato, os estrógenos podem, simultaneamente, estimular a proliferação celular e a expressão gênica assim como aumentar o número de erros na replicação do DNA através de seus produtos de oxidação^{8,9}. A progesterona, através dos seus receptores (PRA e PRB), tem um duplo

papel no tecido mamário podendo tanto estimular a proliferação celular como ter efeitos diferenciadores¹⁰.

Um terceiro receptor hormonal, também importante para o desenvolvimento da doença, é o receptor de membrana HER2 (*human epidermal growth factor receptor 2*), também conhecido como *Neu*, *ErbB-2*, *CD420* ou *p185*. Her2 é membro da família do receptor do fator de crescimento epidérmico (*EGFR/ErbB*), e, em humanos, é codificado pelo gene *ERBB2*. A amplificação gênica ou superexpressão deste gene tem sido relacionada com a patogênese e progressão de certos tipos de câncer de mama agressivos e, portanto, HER2 é um biomarcador importante e alvo da terapia para a doença¹¹. Ao contrário dos outros membros da família, HER2 é considerado como um receptor órfão dado que não tem nenhum ligante conhecido. Os outros três receptores ErbB conhecidos tem ligantes descritos e podem formar homo ou heterodímeros após a ligação do ligante¹².

HER2 pode se heterodimerizar com qualquer um dos outros três receptores, sendo ele o parceiro de dimerização preferencial dos receptores ErbB. A dimerização resulta na autofosforilação de resíduos de tirosina no interior do domínio citoplásmico da proteína, o que inicia uma variedade de vias de sinalização (MAPK, PI3K/Akt, fosfolipase C, PKC e *STAT*). A sinalização através da família dos receptores ErbB promove a proliferação de células e opõe-se à apoptose, e, por conseguinte, deve ser altamente regulado para impedir que o crescimento celular descontrolado ocorra¹³.

1.3 CÂNCER DE MAMA TRIPLO NEGATIVO

Cerca de 10-20% dos casos de câncer de mama diagnosticados correspondem a um subtipo que não expressa os receptores dos hormônios estrogênio e progesterona e não possui amplificação de HER2. Dado a estas características, este subtipo tumoral é conhecido como câncer de mama triplo-negativo ou TN^{13,15}, o qual apresenta um perfil molecular único, comportamento agressivo e pior prognóstico^{8,17}. Ao mesmo tempo, os tumores TN são normalmente maiores que os não-TN⁸, apresentam padrões distintos de metástase¹³, associação com mutações em *BRCA1*¹⁴, expressão aberrante de moléculas de reparo de DNA⁶ e desenvolvimento na idade jovem, principalmente em mulheres.¹⁵⁻¹⁷

Os tumores TN não respondem aos tratamentos convencionais como tamoxifeno ou inibidores da aromatase, nem a terapias que visam os receptores HER2, tais como o emprego do anticorpo Herceptin (trastuzumab)¹¹. Além disso, possuem caráter heterogêneo e não apresentam alvos moleculares bem definidos¹⁸. Algumas vias alteradas relacionadas à doença estão sendo estudadas e usadas na busca de novas abordagens terapêuticas, como inibidores de *EGFR*¹⁸ e de *PARP*¹⁹, porém, ainda é necessário um conhecimento mais aprofundado de suas bases moleculares para o desenvolvimento de tratamentos eficientes²⁰. Tem-se verificado que o grau de captação de 2-deoxi-2- (18F) fluoro-D-glicose (FDG), associado à técnica empregada para a visualização de tumores sólidos em pacientes por tomografia por emissão de pósitrons, se correlaciona com o alto índice proliferativo de tumores TN²¹. Esta técnica evidencia uma característica destes tumores ligada a um reconhecido marco da adaptação metabólica sofrido pelas células tumorais: o aumento da glicólise aeróbica ou efeito Warburg²¹.

1.3.1 Subclassificação Molecular de TN Baseado no Perfil de Expressão Gênica

Sorlie e colaboradores¹⁵ dividiram o câncer de mama em cinco subtipos moleculares intrínsecos (luminal A, luminal B, HER-2 enriquecido, de tipo normal e basal), ou seja, não modulado por tratamento, com base em análise de agrupamento hierárquico de aproximadamente 500 genes. O estudo utilizou dados de microarranjo de mRNA de 65 tumores de mama obtidos de 42 indivíduos diferentes. Mais tarde, reportou-se também o subtipo *claudin-low*²² (Tabela 1).

Tabela 1 Subtipos intrínsecos de câncer de mama²³

Subtipos	Características
Luminal A	Alta expressão dos genes de ER e ER associados, com bom prognóstico clínico.
Luminal B	Baixa expressão dos genes de ER e ER associados. Esta expressão está associada a uma maior taxa de proliferação de células tumorais e um pior prognóstico clínico em comparação com o subtipo luminal A.
HER2 enriquecido	Alta expressão dos genes de HER2 e <i>GRB7</i> . Dado a existência de terapia focada com anticorpos monoclonais direcionados a HER2, hoje em dia apresenta bom prognóstico.
Basal	Positivo para a expressão de citoqueratina basal, mas negativo para a expressão de genes relacionados a luminal e HER2 com nível normal; associado a uma elevada taxa de proliferação de células tumorais e a um mal prognóstico clínico.
Normal	Expressão semelhante comparada à mama normal, suspeita de contaminação de células normais.
<i>Claudin-low</i>	Sem expressão das proteínas claudinas que estão implicadas na adesão célula-célula; alta expressão de marcadores relacionados a fenótipo mesenquimal e de células-tronco; sem expressão de ER e expressão normal de <i>HER2</i> , mas com baixa expressão basal de citoqueratina.

Cerca de 70% dos TN pertencem ao subtipo de tipo basal, e 76% dos cânceres basais foram considerados triplos negativos. Muitos estudos usaram os dois indistintamente, no entanto, e é importante notar que, embora exista uma sobreposição significativa, o subtipo de tipo basal não abrange todo o TN²³.

Os tumores TN apresentam vários subtipos que são classificados por seus padrões imuno-histológicos²⁴, e pela expressão diferencial de genes²⁰. A utilização desta última estratégia permitiu a identificação de seis subtipos: os basais (1 e 2), o imunomodulador (IM), o mesenquimal (M) e mesenquimal *stem-like* (MSL) e o receptor de androgénio luminal (LAR). A identificação destes diversos subtipos em tumores TN evidencia o alto grau de heterogeneidade da doença²⁰ e ressalta a necessidade de se procurar novas estratégias terapêuticas⁸.

1.4 METABOLISMO TUMORAL

1.4.1 Efeito Warburg

Em mamíferos, a proliferação celular é necessária para os processos de embriogênese e crescimento, sendo essencial para a correta função de vários tecidos adultos. Entretanto, a proliferação celular também está na gênese do processo tumoral, onde, diferente das células saudáveis, há perda de seu controle fino.

A pesquisa sobre a proliferação celular tem sido focada no entendimento dos mecanismos que regulam o estado proliferativo, o que tem levado à identificação das vias de tradução de sinais estimuladas por fatores de crescimento e das redes transcricionais que permitem o início e manutenção do ciclo celular. Entretanto, a iminência da proliferação é um desafio para o metabolismo celular uma vez que a produção de duas células filhas requer a duplicação da biomassa celular (proteínas, lipídeos e ácidos nucleicos). Células em estado proliferativo, em geral, absorvem nutrientes em excesso às suas necessidades bioenergéticas e desviam metabólitos para vias que suportam a plataforma de síntese de macromoléculas (Figura 2). Desta maneira, sinais que estimulam a proliferação celular precisam também participar na reorganização da atividade metabólica que permitirá o início da proliferação de células até então quiescentes.

Nas últimas décadas, o estudo de diversos tipos de células proliferativas, incluindo as tumorais, tem permitido cada vez mais entender o metabolismo no estado proliferativo. Sabe-se hoje que este metabolismo difere do metabolismo das células quiescentes pela alta taxa de glicólise, com concomitante produção de lactato, e a biossíntese de lipídeos e outras macromoléculas²⁵.

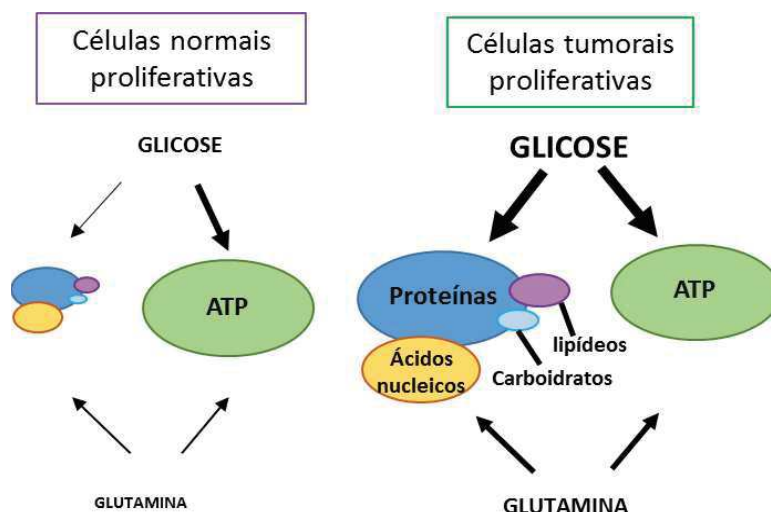


Figura 2. Diferenças chaves entre o metabolismo de células normais quiescentes (esquerda) e células proliferativas normais ou tumorais (direita). Células quiescentes captam e oxidam nutrientes visando a produção maximizada de ATP. Por outro lado, em células proliferativas, há aumento da captação de nutrientes (glicose e glutamina em sua grande maioria), os quais são metabolizados, tanto para a produção de ATP, quanto para abastecer vias biossintéticas que culminam com a síntese de proteínas, ácidos nucleicos e lipídeos. Adaptado de Matthew Vander Heiden Webinar (2012)²⁶.

Na década de 1920, Otto Warburg observou que células tumorais de rápida proliferação consumiam glicose a taxas surpreendentemente rápidas em comparação a células normais e secretavam muito do carbono derivado da glicose como lactato (tanto na presença quanto na ausência de oxigênio), ao contrário de optar por sua completa oxidação²⁷. A este fenômeno se deu o nome de efeito Warburg, ou glicólise aeróbica, o qual foi reconhecido posteriormente como um novo *hallmark* do câncer² (Figura 3). O processo de glicólise é muito mais rápido e confere às células capacidade de produzir energia e sobreviver na ausência de oxigênio, um fator crucial para o crescimento do tumor²⁸. Sabe-se que células tumorais expressam quantidades aumentadas de transportadores de glicose e enzimas da via glicolítica, favorecendo a captação de glicose e a velocidade de catálise^{29,30}. Desta forma, células transformadas podem superar a produção de energia de uma célula normal³¹ (Figura 3).

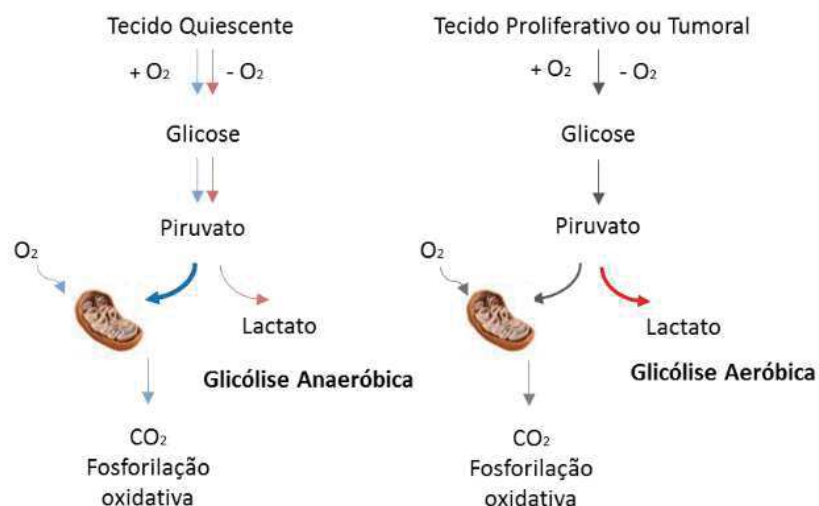


Figura 3. Diferenças entre o metabolismo de células quiescentes diferenciadas (esquerda) e de células proliferativas (direita) no tocante ao destino do piruvato. Adaptado de Vander Heiden et al. (2009) ³².

As células em estado de proliferação, e especialmente as células tumorais, dependem fortemente da disponibilidade de precursores metabólicos para a síntese dos blocos constituintes celulares (lipídeos, proteínas e ácidos nucleicos). Neste sentido, é importante ressaltar que a intensificação da via glicolítica é importante por fornecer diretamente intermediários para a síntese de lipídeos e do açúcar ribose para os nucleotídeos (Figura 4). Além disso, o aumento da captação de glicose e da secreção de lactato propicia vantagem competitiva às células neoplásicas em relação às não transformadas vizinhas³¹, como por exemplo, a acidificação do meio extracelular causado pelo lactato. Para células normais, o baixo pH provoca danos que levam a morte celular. Neste mesmo ambiente, as células neoplásicas podem sobreviver e proliferar, devido a modificações inerentes às mesmas, como aumento no número de transportadores de prótons. Outra vantagem da acidose extracelular é que ela promove angiogênese (formação de novos vasos sanguíneos), processo importante para a nutrição do tumor^{33,34}.

1.4.2 Ciclo do TCA

Um efeito direto do Efeito Warburg é a diminuição da entrada de piruvato glicolítico no ciclo do ácido tricarboxílico (*Tricarboxylic Acid Cycle*, TCA). Entretanto, para sintetizar lipídeos, proteínas e ácidos nucleicos, as células proliferativas também

usam precursores do TCA (Figura 4). Desta maneira, o papel chave do ciclo TCA nestas células seria o de agir como uma fonte de metabólitos para a biossíntese³⁵. O resultado é o contínuo efluxo de intermediários metabólicos (cataplerose), o que representa uma diferença importante em relação ao metabolismo não proliferativo de tecidos oxidativos, como o cardíaco, por exemplo, onde a visão tradicional do ciclo TCA é o de servir para extração máxima de ATP a partir de substratos oxidáveis.

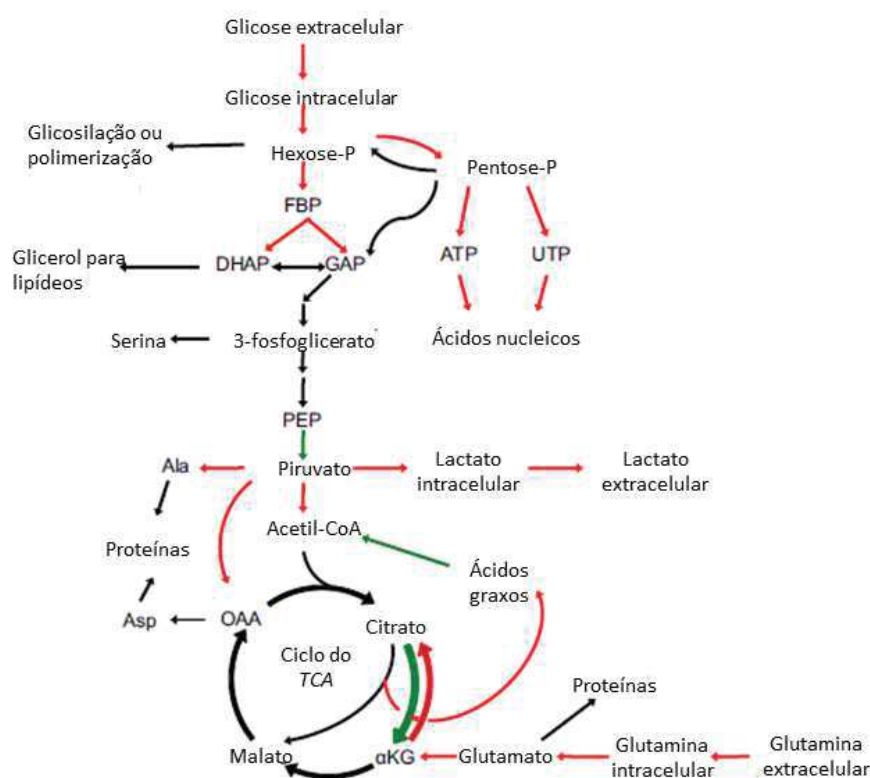


Figura 4. Metabolismo central do carbono. Reações metabólicas que tendem a ser mais rápidas em tumores estão identificadas por setas em vermelho, e reações mais lentas são denotadas por setas em verde. DHAP: fosfato de diidroxiacetona; GAP: gliceraldeído3-fosfato; KG: alfa-cetoglutarato; OAA: oxaloacetato; PEP: fosfoenolpiruvato. Ciclo do TCA: ciclo

do ácido tricarboxílico. Adaptado de Coller, H. (2014)³⁶.

A glicose metabolizada que entra no ciclo TCA dentro da mitocôndria na forma de piruvato é convertida a acetil-CoA, e então em citrato. O citrato é transferido para fora da mitocôndria e, no citosol, é convertido à oxaloacetato (OAA) e o precursor lipogênico, acetil-CoA (Figura 5). A exportação de citrato para a síntese lipídica tem impacto sobre a função geral do ciclo resultando no que é chamado de ciclo

“truncado”, devido à diminuição da quantidade relativa da fração de citrato mitocondrial que é oxidada. Outros intermediários do ciclo do TCA também são usados para a biossíntese de diferentes macromoléculas. OAA e α -cetoglutarato suplementam o *pool* de aminoácidos não essenciais que serão usados para a síntese de proteínas e nucleotídeos, contribuindo para o fenômeno de cataplerose.

De maneira a sustentar o funcionamento do ciclo TCA em face de cataplerose, as células precisam lançar mão de um influxo de metabólitos intermediários, fenômeno conhecido como anaplerose (Figura 5). Uma fonte importante de anaplerose é o metabolismo de aminoácidos, particularmente o da glutamina, o aminoácido mais abundante em mamíferos. As células podem oxidar a glutamina parcialmente de uma maneira análoga à oxidação parcial da glicose através da glicólise aeróbica. Esta via, conhecida como glutaminólise, usa diversos passos do ciclo TCA, e torna a glutamina uma fonte também energética para as células proliferativas^{37,38}. Além disto, há a formação de glutamato e aspartato, os quais, junto com a glutamina por si, são precursores de ácido nucleico e serina. Outros importantes destinos da glutamina são a síntese de lipídeos, glutatona e a via de síntese de hexosaminas³⁹. Tomados em conjunto, estas observações sugerem que o metabolismo da glutamina permite às células proliferativas manter um fluxo anaplerótico compatível com o uso dos intermediários metabólicos do ciclo TCA como precursores biossintéticos. De fato, sabe-se que hoje que a glutamina (assim como a enzima glutaminase, primeira na via glutaminolítica) são importantes para a proliferação de diversos tipos tumorais^{40,41,42,43}, assim como também tem sido mostrada ser importante para o processo de metástase⁴⁴.

Como não são todos os tipos de tumores que apresentam metabolismo glutaminolítico, especula-se que a glutaminólise não seja uma simples consequência da transformação celular, mas seja resultado da ação direta de oncogenes específicos²⁹. A família de genes *MYC* (*c-myc*, *L-myc*, *s-myc* e *N-myc*), os quais normalmente apresentam maior número de cópias em processos tumorais (70% dos tumores humanos)⁴⁵, são fatores de transcrição que regulam crescimento e entrada no ciclo celular. Em células normais, a estimulação dada por mitógenos é necessária para um aumento da expressão de *c-myc* na fase G1, facilitando a entrada das células na fase S através do estímulo da expressão de ciclinas e CDK4⁴⁵. C-myc controla a expressão de enzimas glicolíticas, tal como a *LDH-A*, além de estar também envolvido

no metabolismo de nucleotídeos, necessário para a duplicação do genoma²⁵. A superexpressão de Myc na linhagem celular de glioma SF188 levou a um processo de reprogramação genética que culminou com a expressão de genes relacionados com o catabolismo de glutamina e tornou as células dependentes deste nutriente²⁹. Outro trabalho revelou que a presença de c-myc promove o aumento dos níveis proteicos da glutaminase, através de um mecanismo que envolve a supressão de miRNAs, os quais são reguladores negativos da expressão desta enzima⁴⁶.

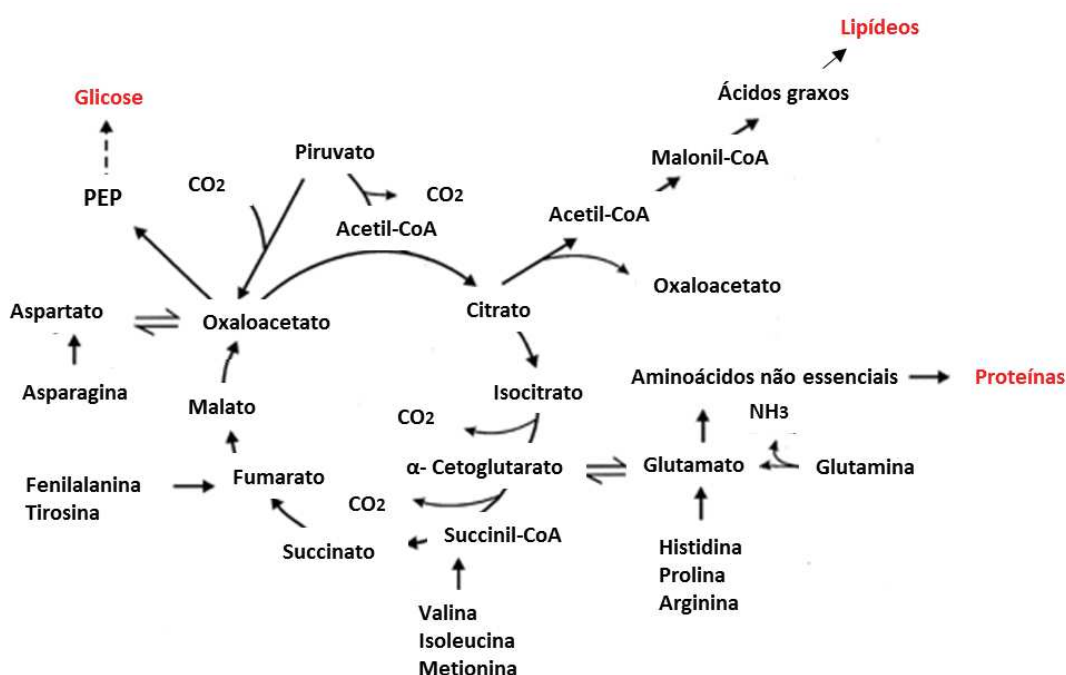


Figura 5. Anaplerose e cataplerose do ciclo do TCA. Devido, entre outros, a menor entrada de piruvato glicolítico, as células tumorais sofrem o que é chamado truncamento do TCA. Intermediários do TCA são, então, repostos a partir de fontes anapleróticas tais como aminoácidos (especialmente a glutamina) e a geração de oxaloacetato diretamente a partir de piruvato (via piruvato carboxilase). As reações catapleróticas são importantes tanto para a gluconeogênese quanto para a lipogênese. Adaptado de Owen et al. 2002⁴⁷.

1.4.3 Alterações metabólicas dos tumores TN

Em um estudo realizado por McClelland e colaboradores (2012)⁴⁸, foram identificados novos alvos terapêuticos para tumores de mama triplo-negativo a partir do uso de RNAi. Um desses alvos é a enzima *LDHB* (Lactado Desidrogenase isoforma B), encontrada mais expressa no tumor em relação a células normais. Esta enzima é a última da via glicolítica e está envolvida na conversão reversível de piruvato a

lactato. O bloqueio de *LDHB* reduziu a proliferação em câncer triplo-negativo *in vitro* e estabilizou o tumor *in vivo*. A partir dessas descobertas, possibilidades terapêuticas estão surgindo, como o uso de galoflavina, um inibidor de *LDH* testado em linhagens de câncer de mama triplo-negativo e que apresentou um grande potencial na diminuição da agressividade deste tumor⁴⁹.

Em outro estudo, Timmerman e colaboradores (2013)⁵⁰ traçaram o perfil de dependência de glicose e glutamina em diferentes linhagens de triplo-negativo, encontrando uma alta dependência de glutamina em um subconjunto destas células, formado principalmente por células do tipo basal. Os autores demonstraram também que o antiporter xCT, responsável pela entrada de cistina na célula e expresso em um terço dos tumores triplo-negativos, é um importante alvo terapêutico para a doença. Além disso, os autores propuseram o sulfasalazine (SASP), um anti-inflamatório inibidor de xCT, como um possível fármaco para a terapia contra o câncer de mama triplo-negativo.

A análise de um painel de linhagens de câncer de mama mostrou que a expressão da isoforma C da glutaminase (GAC), e sua atividade glutaminolítica, encontra-se elevada na maioria dos TN em relação a células que expressam algum dos receptores. O tratamento de linhagens triplo-negativo de câncer de mama com CB-839 *in vitro* e *in vivo* reduziu o consumo de glutamina e a produção de glutamato, diminuiu o consumo de oxigênio e de diversos intermediários do ciclo do TCA⁴¹, indicando potencial terapêutico para compostos que inibem esta enzima. De fato, GAC tem sido proposto como um biomarcador potencial para a sensibilidade do inibidor CB-839, o qual está atualmente em estudos de Fase Clínica I para diversos tumores sólidos (entre eles o TN) e hematológicos^{41,50}.

Dado aos exemplos já bem-sucedidos denotados, avançar no entendimento do processo de adaptação metabólica deste subtipo tumoral e explorar os achados para a proposição de novos alvos metabólicos pode indicar um caminho promissor no tratamento desta doença.

1.5 BIOLOGIA DE SISTEMAS

A biologia de sistemas tem como objetivo dar uma visão biológica mais integrativa entre as múltiplas entidades moleculares que governam o funcionamento

de um organismo e todas suas conexões ao longo de seus diferentes níveis de organização⁵¹. Um dos grandes desafios deste tipo de estudo é inferir o máximo de informação biológica, relevante e concreta, a partir dos dados coletados. Para tanto, se utilizam as plataformas “ômicas”: Genômica, Transcriptômica, Proteômica, Metabolômica e Fluxômica⁵² (Figura 6).

Até pouco tempo, era comum uma abordagem reducionista para o entendimento do câncer, com pesquisas focadas em mutações específicas e/ou vias de sinalização única. Entretanto, a aplicação de abordagens ômicas para o estudo do câncer é uma promessa para o melhoramento do diagnóstico, prognóstico, e para a elucidação de novos alvos terapêuticos através da construção de uma taxonomia molecular da doença. Neste sentido, por exemplo, com a obtenção do perfil de expressão gênica é possível especular que proteínas representam potenciais alvos terapêuticos. No entanto, para a determinação mais efetiva de alvos, é necessária a complementação da técnica com estudos fenotípicos. Assim, os estudos de metabolômica podem complementar a abordagem de transcriptômica, dado que é a última instância (mais do que o perfil de transcritos ou protéico) na determinação das rotas metabólicas acionadas ou suprimidas em uma célula.

A transcriptômica permite quantificar o nível de expressão dos genes utilizando técnicas que podem analisar milhares de moléculas de RNA mensageiro (mRNA) ao mesmo tempo⁵³. Atualmente, as pesquisas em transcriptômica buscam a investigação de genes que se expressem diferencialmente (maior ou menor expressão), por exemplo, em células normais e tumorais, gerando um perfil gênico que possa apontar para o caminho que a célula adota para sua proliferação e encontrar, desta maneira, informação que leve a novos alvos para intervenção terapêutica.

Uma metodologia que tem sido cada vez mais aplicada para a caracterização e quantificação do nível de expressão gênica de células é o RNA-Seq (*RNA Sequencing*). A técnica, em vantagem a outras tais como o microarranjo, permite quantificação precisa e absoluta nos níveis de mRNA, a detecção de eventuais mutações e de variantes de transcritos antes desconhecidos. Nesta técnica, em linhas gerais, uma população de RNA (total ou fracionado) é convertida em bibliotecas de fragmentos de cDNA com adaptadores aderidos nas extremidades. Cada molécula é sequenciada em modo *High-Throughput* para obter sequências curtas a partir dos dois extremos, chamadas *reads*. Seguido do sequenciamento, os *reads* são alinhados com

um genoma de referência ou transcritos de referência para produzir um mapa de transcrição em grande escala⁵⁴.

A metabolômica faz referência ao estudo do perfil metabólico de uma célula (ou tipo celular), ou seja, o conjunto completo de metabólitos ou intermediários de baixo peso molecular, em especial daqueles envolvidos nas vias metabólicas dos lipídios, açúcares, aminoácidos e nucleotídeos. Estas vias, por sua vez, são altamente dependentes de variações fisiológicas, do desenvolvimento ou de estados patológicos, e podem ajudar na descrição do fenótipo celular⁵⁵. A metabolômica seria, portanto, a “leitura funcional do estado fisiológico” de uma célula ou de um organismo⁵⁶.

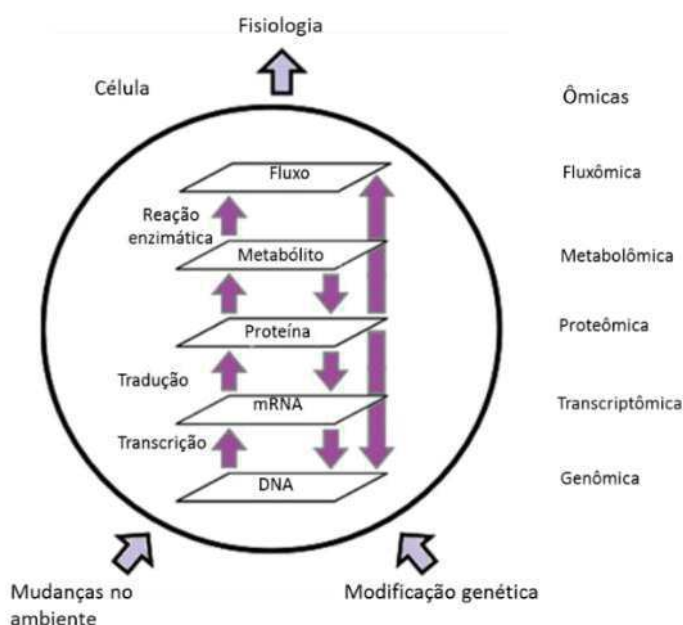


Figura 6. Ilustração de diferentes níveis de informação que podem ser medidos em uma célula. Adaptado de Matsuoka & Shimizu (2012)⁵⁷.

Na era pós-genômica, ferramentas de perfil metabólico celular foram desenvolvidas para proporcionar uma visão global do desenvolvimento e progressão tumoral. Entre elas se encontram técnicas de separação (cromatografia), técnicas de detecção do tipo espectrometria de massas (MS) e a espectroscopia por ressonância magnética nuclear (RMN), além das técnicas mistas. Cada técnica combina estratégias para identificação e quantificação dos metabólitos celulares usando tecnologias analíticas sofisticadas com a aplicação de métodos estatísticos e de análise multivariada para a extração da informação e para a interpretação dos dados⁵⁸. Além disso, cada uma apresenta suas próprias particularidades, de modo

que suas vantagens e limitações devem ser levadas em consideração em cada estudo metabolômico^{54,57}.

2 OBJETIVOS

2.1 OBJETIVO GERAL

O objetivo deste projeto foi analisar a expressão gênica diferencial (mediante obtenção de dados de RNA-Seq realizados *in house* e disponíveis em bancos públicos) de linhagens celulares e tecidos tumorais de mama triplo negativo e não-triplo negativo de maneira a se compreender as alterações na expressão de genes de vias metabólicas, e, assim, caracterizar a assinatura gênica metabólica do subtipo TN. Além, planejamos também obter dados de metabolômica de linhagens celulares TN e não-TN por ressonância magnética nuclear, de maneira a complementar as informações de expressão gênica, e entender as estratégias empregadas pelo subtipo TN para adaptar seu metabolismo. Por fim, objetivamos identificar potenciais novos alvos terapêuticos para o tratamento deste subtipo tumoral.

2.2 OBJETIVOS ESPECÍFICOS

- Definir a expressão gênica diferencial entre linhagens TN e não-TN e entre tecidos tumorais TN e não-TN. Para tanto, vamos obter dados de RNAseq de linhagens celulares disponíveis em bancos públicos e sequenciadas *in-house* e de tecidos tumorais obtidos de bancos públicos.
- Validar dados de RNA-seq sequenciados *in-house* por qPCR.
- Definir as particularidades metabólicas de tumores TN em relação a não-TN através da construção de mapas de integração bioquímica dos dados de RNAseq, e, quando presentes, de metabólitos obtidos de linhagens tumorais.
- Definir genes potencialmente importantes para a sobrevivência de tumores TN através da análise de expressão diferencial TN *versus* não-TN. Avaliação dos mesmos por *knock down* e estudos de proliferação celular e linhagens.

A apresentação dos resultados e discussão seguirá em 2 capítulos para facilitar o entendimento dos achados.

3 CAPITULO I: DIFERENÇAS METABÓLICAS ENTRE TUMORES TN E NÃO-TN

3.1 METODOLOGIA

3.1.1 Cultura de células

Seis linhagens de células tumorais de mama provenientes do repositório ATCC, 4 TN (MDA-MB231, MDA-MB436, MDA-MB468, BT-549) e 2 não-TN (SKBR3, MCF7), foram sequenciadas pela técnica de RNA-Seq (Tabela 1). Para as análises de metabolômica, além das já citadas, as linhagens MDA-MB157, HCC1806, Hs578T, ZR-75-1, também provenientes da ATCC, foram utilizadas (Tabela 2). Com exceção da MDA-MB436, MDA-MB468 e MCF7, todas as linhagens foram cultivadas em meio RPMI com 10% de soro fetal bovino (SFB) e mantidas em estufa a 37°C e 5% de CO₂. A linhagem MDA-MB436 foi mantida no meio Lebowitz L-15 (cultilab) suplementado com 10% SFB, 10 µg/ml insulina e 16 µg/ml glutatona a 37°C em uma atmosfera úmida, sem controle de CO₂. A linhagem MDA-MB468 foi mantida no meio Lebowitz L-15 (cultilab) suplementado com 10% SFB a 37°C em uma atmosfera úmida, sem controle de CO₂. Por fim, MCF7 foi mantida no Meio MEM (Cultilab) com 4,5 g glicose, suplementado com 10% soro fetal bovino (SFB), 10 µg/ml insulina a 37°C em atmosfera úmida, 5% de CO₂.

Para os estudos de metabolômica, as células foram semeadas em placas de 100 mm, e cultivadas com troca de meio a cada dois dias, até atingirem confluência de 80%. Após, 2X10⁶ células foram plaqueadas em placas de 100 mm, incubadas na estufa a 37°C por 24 horas e coletadas para as análises. Coletaram-se amostras de três passagens de células.

Tabela 2. Linhagens celulares utilizadas nos experimentos e seu *status* quanto à presença dos receptores de estrógeno (ER), progesterona (PR) e Her2. O subtipo de câncer para cada uma das linhagens foi baseada na classificação segundo Lehmann et al 2011²⁰ para as TN e de Neve et al 2009⁶³ para as N-TN. ME: metastática; NR: Não reportado; M: Mesenquimal; MSL: *Mesenquimal stem-like*; BL: *Basal-like*; L: Luminal. Há também informação se a linhagem foi usada para RNA-seq e/ou RMN.

Linhagem	Histologia	Origem das células	ER	PR	HER2	Subtipo ^{20, 63}	perda alélica de <i>BRCA1</i> ^{6 4}	RNA-Seq	RMN
MDA-MB157	Carcinoma medular	NR	ER-	PR-	Her2-	MSL	sim		X
MDA-MB231	Adenocarcinoma	ME	ER-	PR-	Her2-	MSL	sim	X	X
MDA-MB436	Adenocarcinoma	ME	ER-	PR-	Her2-	MSL	sim	X	X
MDA-MB468	Carcinoma	ME	ER-	PR-	Her2-	BL	sim	X	X
HCC1806	Carcinoma escamoso	NR	ER-	PR-	Her2-	BL	NR		X
BT-549	Carcinoma ductal	Primário	ER-	PR-	Her2-	M	sim	X	X
Hs578t	Carcinossarcoma	Primário	ER-	PR-	Her2-	M	sim		X
SKBR3	Adenocarcinoma	ME	ER-	PR-	Amplificado	L	sim	X	X
ZR-75-1	Carcinoma ductal	NR	ER+	PR-	Não amplificado	L	não		X
MCF7	Carcinoma ductal	ME	ER+	PR+	Não amplificado	L	sim	X	X

3.1.2 Extração de metabólitos intracelulares

As células foram lavadas três vezes com PBS 1X para remover os componentes do meio de cultura, seguido de adição de 1 mL de acetonitrila (CH₃CN) para realização do *quenching* celular (parada dos processos metabólicos). Em seguida, as células foram colocadas em tubo falcon de 15 ml adicionando posteriormente 1 ml de clorofórmio grau HPLC (Fisher Scientific). Com este procedimento, obteve-se uma camada superior, que corresponde à fração de metabólitos polares, uma camada inferior com a fração apolar e uma camada intermediária contendo proteínas. Coletou-se as fases superior e intermediária as

quais foram, em seguida, congeladas em nitrogênio líquido, liofilizadas e armazenadas no -80°C. A fase superior foi analisada por RMN, enquanto que a intermediária foi quantificada pela reação de Bradford para determinação de massa proteica empregando-se o kit *Bradford Bio-Rad Assays*.

3.1.3 Aquisição de espectros de RMN e quantificação de metabólitos

Os espectros de RMN de prótons, ou ^1H , são rotineiramente empregados para a definição de perfis metabólicos de diferentes amostras através da identificação e quantificação de diversas moléculas em misturas complexas. As identidades e concentrações de cada molécula podem ser obtidas por meio da comparação com sinais de RMN de moléculas individuais, previamente medidos em condições fisiológicas e constantes, e disponíveis em bibliotecas de metabólitos tais como a HMDB (*Human Metabolome Database*)⁶⁵, ou por comparação com dados disponíveis na literatura científica⁵⁸.

Para a aquisição de espectros ^1H , foi adicionado aos extratos celulares liofilizados 28 mM de NaH_2PO_4 e 72 mM de Na_2HPO_4 diluídos em água deuterada em pH final de 7.4. Como referência interna foi adicionado 0,5 mM de TSP (ácido 3-(trimetil-silil) -2,2',3,3'-tetradeuteropropiônico ou TMSP-d4). A espectroscopia foi realizada a 25°C com espectrômetro de marca *Agilent* operando a 600 MHz com a sonda criogênica de ressonância tripla de 5mm. Espectros unidimensionais foram adquiridos utilizando um volume de 600 μl , uma sequência de pulso de pre-saturação com os seguintes parâmetros: tempo de aquisição de 4s, *delay time* de 1,5 s e 256 números de transientes.

A identificação e quantificação dos metabólitos presentes a partir dos sinais espectrais de ^1H foram realizadas com o auxílio do software *Chenomx NMR suite* (v 7.5, Edmonton, Canadá). Esta abordagem, denominada *targeted profiling*, ou determinação do perfil de moléculas na amostra, baseia-se na comparação dos sinais obtidos com uma biblioteca de espectros de metabólitos obtidos individualmente (o que permite a identificação) bem como na utilização de uma referência interna de concentração (no caso, o TSP), para a quantificação metabólica. Os dados obtidos (em mols) foram normalizados pela massa proteica quantificada pelo método de Bradford.

3.1.4 Extração de RNA

As extrações do RNA total foram realizadas com a utilização do *kit* RNeasy (Qiagen), seguindo as instruções fornecidas pelo fabricante. Verificou-se a concentração e a qualidade do RNA obtido por espectrofotometria a 260 nm e em gel de agarose 1% (m/v) preparado em solução 20 mM MOPS (3-N-morfolino ácido propanosulfônico), 5 mM acetato de sódio, 1 mM EDTA pH8,0, 0,74% de formaldeído em água tratada com DEPC (Dietilpirocarbonato) em um volume final de 500ml. O gel foi equilibrado nesta solução durante 30 minutos antes de ser utilizado, e a amostra (RNA em 15 µl de solução saturada de bromofenol, 4 mM EDTA pH 8,0, 2.66 % formaldeído, 20% glicerol, 4X solução corrida, 9,25% de formamida v/v, água tratada com DEPC) foi incubada a 65°C por 5 minutos e imediatamente aplicada. A eletroforese foi realizada por 1 hora a 60 V (3,75 V/cm).

3.1.5 Next generation sequencing com Hiseq2000 (Illumina)

Dez microgramas de RNA total das amostras MDA-MB231 e SKBR3 foram enviados pelo correio (armazenadas em gelo seco) para o laboratório *High Throughput Sequencing Facility* (HTSF) da Universidade de Carolina do Norte em Chapel Hill sob a direção do Prof. Dr. Piotr Mieczkowski, enquanto que dois microgramas de RNA total das amostras MDA-MB436, MDA-MB468, BT549, MCF7 foram enviados (armazenadas em gelo seco) para o Laboratório Central de Tecnologias de Alto Desempenho (LACTAD) da Universidade Estadual de Campinas (UNICAMP). As amostras foram sequenciadas utilizando-se a técnica *Paired End* x 100 pares de bases no Hiseq2000 (Illumina/Solexa), segundo princípios de RNA-Seq (Figura 7). Somente 1 amostra de cada linhagem foi sequenciada.

Para o sequenciamento realizado neste projeto foi separado o mRNA do RNA total por meio de *kits* de purificação específico: *Dynabeads mRNA purification kit* (Life Technologies) e *TrueSeq RNA sample preparation kit v2* (Illumina) usados pelo laboratório HTSF e o LACTAD, respectivamente. Após isolados, os mRNAs foram fragmentados na presença de cátions divalentes e temperaturas elevadas e, em seguida, empregados para a síntese de cDNA com *primers* randômicos empregando-se o *kit Superscript II Reverse Transcriptase* (Life Technologies).

A segunda fita foi sintetizada utilizando-se a DNA polimerase I e Rnase H. T4 DNA polimerase e *Klenow fragment* e *AMPure XP beads-Beckman coulter genomics*

foram usadas para se obter extremidades de cDNA abruptas segundo protocolo do HTSF e do LACTAD, respectivamente. Ao final, uma adenina foi adicionada às extremidades 3' das moléculas de cDNA usando a *Klenow* exo-enzima no HTSF ou a *tailing mix* no LACTAD, seguido da adição de adaptadores contendo uma timina na extremidade 5'. Adaptadores que não se ligaram foram removidos por eletroforese e fragmentos entre 200-300 pares de bases foram purificados, correspondendo aos cDNAs ligados aos adaptadores. Finalmente, as bibliotecas foram enriquecidas por 15 ciclos de PCR usando *primers* complementares às sequências dos adaptadores. A qualidade e quantificação das bibliotecas foram feitas empregando-se Bionalyzer Chip DNA 1000 series II (Agilent). Durante o sequenciamento, foram geradas pontes de amplificação, onde as moléculas desnaturadas de fita dupla são aderidas na superfície sólida de uma placa e amplificadas *in situ*, gerando agrupamentos de moléculas de DNA idênticas (*reads*). O sequenciamento da Illumina utiliza quatro nucleotídeos análogos fluorescentes, os quais funcionam como terminadores do sequenciamento. Quando o nucleotídeo correto é incorporado à cadeia crescente, a sua identidade é revelada pela cor de sua fluorescência. O grupo 3-OH do nucleotídeo é então bloqueado para evitar mais incorporação. Após a etapa de imagem, a placa é lavada e o bloqueio invertido, permitindo assim a adição do nucleotídeo seguinte.

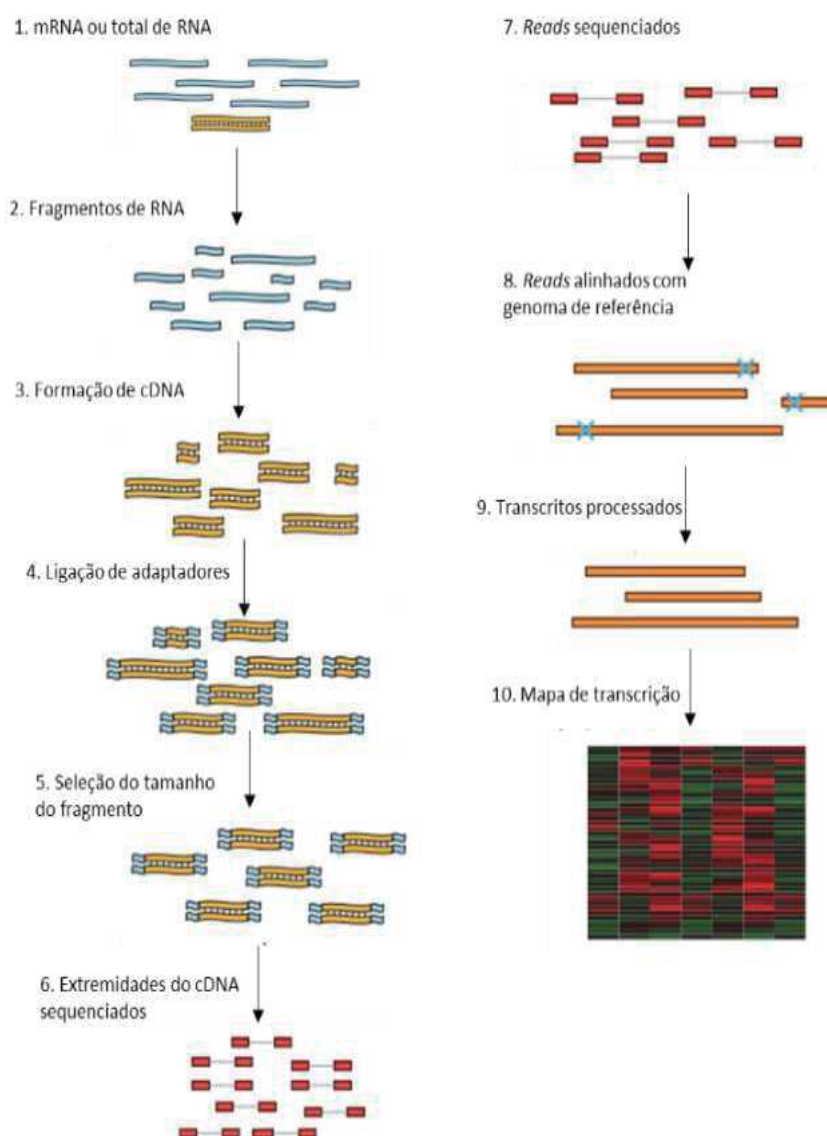


Figura 7. Experimento típico de RNA-Seq. RNAs longos (1) são primeiro fragmentados (2) e, em seguida, convertidos em uma biblioteca de fragmentos de cDNA (3). Adaptadores de sequenciamento (azul) são subsequentemente adicionados a cada fragmento de cDNA (4). O tamanho dos fragmentos a serem sequenciados são selecionados (5). Uma pequena sequência é obtida a partir de cada cDNA usando tecnologia high-throughput de sequenciamento (6,7). As sequências resultantes (reads) são alinhadas com o genoma ou transcriptoma de referência (8), depois o processamento dos transcritos é realizado (9) gerando como resultado o perfil de expressão para cada gene (10). Adaptado de Martin & Wang 2011⁶⁶.

3.1.6 Obtenção dos dados de RNA-Seq de bancos públicos

Dados de RNA-Seq de 1100 amostras de pacientes e dados clínicos respectivos foram obtidos do projeto *The Cancer Genome Atlas* (TCGA), através do seu portal (<https://tcga-data.nci.nih.gov/tcga/>), com os valores de RSEM (RNA-Seq by Expectation Maximization) não normalizados⁶⁷. Dados publicados de RNA-Seq de amostras de diferentes linhagens celulares foram obtidas a partir do portal *Gene Expression Omnibus* (GEO), nos acessos GSE58135 (*Dataset* Varley et al)⁶⁸ e GSE48213 (*Dataset* Daemen et al)⁶⁹, utilizando os servidores DDBJ *sequence read*

archive (DRA)⁷⁰ ou mediante os arquivos de leitura das sequências do NCBI (SRA)⁷¹ para a execução do *download*.

3.1.7 Processamento dos dados de RNA-Seq

Na metodologia experimental empregada, dezenas de milhões de pequenos fragmentos mRNA foram sequenciados com alta precisão, sendo gerados *reads* do tamanho de 100 pb. Os dados destes sequenciamentos foram analisados pelo Dr. Marcelo Falsarella Carazzolle e o estudante de doutorado do nosso grupo Douglas Adamoski. A análise completa dos dados do RNA-Seq foi dividida em etapas, sendo cada uma realizada por um conjunto de programas e metodologias específicas. Estas etapas foram aplicadas tanto para os dados coletados *in-house* quanto para os obtidos de bancos públicos. O *pipeline* foi iniciado pela avaliação da qualidade dos *reads* e das sequências usando o software *FastQC*⁷², seguido do corte das sequências de baixa qualidade utilizando o *Skewer*⁷³ empregando-se como parâmetros tamanho mínimo final de cada *read* de 30pb, qualidade média do *read* de PHRED 30 e *trimming* para obter último nucleotídeo da sequência com qualidade PHRED 20. O alinhamento contra o genoma humano de referência (hg19) foi realizado pelo software *STAR*⁷⁴ e o arquivo SAM gerado foi corrigido com o auxílio de *samtools*. As abundâncias de transcritos foram estimadas pelo método do RSEM (RNA-Seq *by Expectation Maximization*)⁶⁷.

3.1.8 Classificação dos tecidos tumorais de mama entre TN e não-TN

Definimos o perfil de expressão dos marcadores ER, PR e Her2 nos dados de RNA-Seq de tecidos para podermos realizar a classificação dos tumores entre TN e não-TN. Para tanto, primeiramente, os valores de expressão foram normalizados usando a abordagem do quartil superior⁷⁵ e transformados na escala log2. Os dados de expressão dos marcadores foram avaliados com o pacote R *mclust*⁷⁶ para a definição de modelos bimodais de expressão e classificação em positivo ou negativo de acordo com a estratégia previamente utilizada^{20,77}. A escolha entre variância fixa ou distinta entre as duas modais foi baseada no modelo que mais retornasse acerto com os dados clínicos disponíveis. Tal abordagem foi necessária uma vez que os dados de imunohistoquímica (IHC) não estavam disponíveis para todas as amostras.

3.1.9 Determinação da expressão gênica diferencial

A análise de expressão gênica diferencial das amostras TN e não-TN foi realizada com os pacotes do R EBSeq⁷⁸ and DESeq2⁷⁹, empregando-se, em ambos, a normalização pelo método do quartil superior⁷⁵. Após a normalização, os genes foram filtrados de maneira a se manter aqueles com RSEM maior que 10 no terceiro quartil (ou seja, pelo menos 25% dos pacientes/linhagens devem expressar o gene com valor de RSEM maior que 10). Para o EBSeq, foram realizadas gerações o suficiente para ao menos 3 indicarem convergência dos parâmetros numéricos. A lista de genes diferencialmente expressos (DE) gerada pelo EBSeq foi limitada de maneira a conter os seguintes parâmetros: genes DE cuja significância estatística (PPEE) fosse menor ou igual que 0.05 e com *fold change* $\geq +2$ ou ≤ -2 (ou $\log_2 \text{Fold Change} \geq +1$ ou ≤ -1). Os *heatmaps* foram construídos com o pacote do R pheatmap⁸⁰, os diagramas de Venn com o pacote do R VennDiagram⁸¹ e os gráficos de PCA com o pacote do R ggbiplot⁸². Vias canônicas e funções moleculares e biológicas alteradas foram definidas usando o *software* METACORE (<http://www.metacore.com>).

3.1.10 Validação dos dados de RNA-Seq por qPCR

O qPCR foi utilizado para a validação da expressão de alguns genes selecionados (com foco em genes metabólicos de interesse do estudo) em função de terem expressão aumentada ou diminuída nas linhagens TN versus não-TN. Esta validação foi importante uma vez que as amostras sequenciadas não o foram em replicata.

Todos os *primers* foram desenhados com o algoritmo da IDT (www.idtdna.com) de modo a se detectar todas as isoformas dos genes, e de ter pelo menos um dos *primers* em junção exon-exon. Os *primers* foram confirmados quanto à especificidade utilizando-se o software Primer Blast (<http://www.ncbi.nlm.nih.gov/tools/primer-blast/>). A Tabela 2 apresenta a lista de *primers* utilizados. O RNA total extraído pelo kit Rneasy (Qiagen) foi copiado em cDNA utilizando-se *primers* hexaméricos randômicos e o kit Superscript III *Reverse Transcriptase* (Life Technologies). Um *mix* contendo SYBR Green 2X, 3 μM de *mix* de *primers forward* e *reverse*, cDNA em concentração aparente de 1,875 ng/ μL e água ultrapura q.s.p 20 μL , foi acondicionado em placa de 96 poços (MicroAmp® 96-Well Plates – Invitrogen) e analisado no termociclador 7500 *Real-Time PCR System* (Applied Biosystems). Os seguintes ciclos foram empregados:

95°C por 10 minutos, seguido de 40 ciclos a 95°C por 15 segundos e extensão a 60°C por 60 segundos. O ensaio foi realizado em triplicata técnica. Para a comparação com os dados de RNA-Seq, para cada linhagem, graficamos os valores de $\log_2 \text{RSEM}+1$ para RNA-Seq versus $1/\Delta\text{CT}$ para qPCR. Utilizou-se a função de correlação de spearman para avaliar a equivalência entre os dados de qPCR e RNA-Seq.

Estabeleceu-se primeiramente a eficiência dos pares de *primers* de alguns dos genes, selecionados ao acaso. Para isto, realizou-se uma diluição seriada dos cDNA e a reação de qPCR. Os valores de CT obtidos foram graficados contra o Log da concentração parente de cDNA e fez-se uma regressão linear dos dados. Para cálculo de eficiência utilizou-se $\{[10^{(-1/\text{inclinação da reta})} - 1] \cdot 100\}$. Para uma diluição seriada 1:1, quanto mais próximo o valor de 2 maior a eficiência da reação (100%). Uma curva de *melting* foi aplicada aos amplicons para confirmar-se a obtenção de somente um pico de *melting*. Empregou-se o *housekeeping* rRNA 18S.

Tabela 3. Lista de primers utilizados para as confirmações de expressão gênica diferencial.

GENE		PRIMERS	GENE		PRIMERS
rRNA 18S	F R	ATTCCGATAACGAACGAGAC TCACAGACCTGTTATTGCTC	MDH1	F R	CAGCCTTAGATAAATACGCCAAG GATCCAAACGAGTCAAGCAAC
GAC	F R	GATCAAAGGCATTCTTTGG TACTACAGTTGTAGAGATGTCC	GLUL	F R	CCCCTTTTCGGTGACAGAAG GGCTTAGGATCAAAGGTTGC
KGA	F R	TGGTGATCAAAGGGTAAAGTC TGCTGTTCTAGAATCATAGTCC	PFKFB2	F R	CAAGCCCTGACTATCCTGAAAG GCCCACGTTTATCACCTTGA
ENO3	F R	AATGTGGGTGATGAAGGTGG TGCGATAGAACTCAGATGCTG	ENO2	F R	CATCAAGGACAAATACGGCAAG CGATCTTTCCGTGTAGCCAG
FBP1	F R	TGAGAAGGATGCTCTGCAAC ACCAAATGAACTCCCCGATG	ACACB	F R	ACAGTTACCGAATTACCAT CATCTTCATCACCTCCAT
PCK2	F R	GAGCTGTGGATGAGAGGTTTC CACATAGGCTGAGTCAGTGAG	ACADM	F R	CACCGAACAGCAGAAAGAA TCCAAGACCTCCACAGTT
PGM2	F R	TGCTGATAGACTTGCTGTGG TGTACGTGTCTTTGAGAGCAC	ACADS	F R	CGATTGTGCTGTGAACTA CTGCCTCCTTGATGAAAG
ALDH2	F R	GACCGTGGTTACTTCATCCAG CTCCCAACAACCTCCTCTATG	ACADSB	F R	GTGATGGCAAATGTAGACCCTA TGGCTTCTGGAACCTTGAC
ALDH9A1	F R	CTACCCCTTTTCAGATTGCCTC AGCCTCACTGTAGATTTTCAGC	ACADVL	F R	CACGGACTTCCAGATAGA CTACTCCAGGTTCTTCAT
ACSS1	F R	CTCTTCATGCTGTACACCTCAG ACAGCCAAAGATGTCACCTG	IDH3G	F R	AAGAGTATCGCCAATAAG AGTGTGCATATTCTCATT
IDH2	F R	AAAACATCCCACGCCTAGTC GACCATTTTGAAAGTGCCGG	ACOT7	F R	TACACCTCCAAGCACTCT CGGGAATACACAACAGGA
IDH3B	F R	GGGTGTGATTGAGTGTGTAAG ACCCATCCCCAAGTTTCATG	GALM	F R	GACCAACCATTCTTACTTCAAC GGCAACTTCTCCTGTAGG
ACSS3	F R	GGCAAACAGAGACTGGAT AGGTGGCAATGGTAACTT	GFPT1	F R	CGAACGAGACGAGAAATCC TCCATATCTTGTGCTTGTGAA
ADCY1	F R	CAATGTGAGCATCCTGTT GTTCTCCGTGGCTAATTC	GNPDA2	F R	GGAGGAATTGGTCCAGATG GATCATTACTTCTCTAGCATCCAT
ADCY7	F R	GAATGTCAGCATCCTCTA CGATACACAGTAGTAGCA	GPI	F R	GCTGCCACATAAGGTCTT TGTTGATGTCCCAGATGATG
AK2	F R	CACAGGAAGGCTGATTCA TGGTTGAGTGTGGTAGG	GUK1	F R	TTCAGCGTGTCCCATACC CCTCCCTGGTTACAAAGTAGT

AMDHD2	F R	TGGTTTGAGTGTGGTAGG AGGAACAACCTTGTGATAA	ME3	F R	TGACAGCAGAGCAAATTG GTAGTAGGAAGCCAGGTT
DCTPP1	F R	AGAACTCTTTCAGTGGAA ACTAAGCTCCTCTTGAAG	NME1	F R	CAAGTTGGCAGGAACATTATAC GCTCGTGTAAATCTACCAGTT
DUT	F R	GCCTATGATTACACAATACCACCTA TCTATGACACCAGCTCCTACA	POLA1	F R	ACAGAGGATGGAGAACAAGAA CGAATGTCATACTGAAGAATAAGGT
ENPP1	F R	AAGTCTTAACCACCTTCTAA TGCTTAATAATCTTCTCTTCTG	POLR1B	F R	TGAGGAGGCAGAGGAAATGG TTCATATTGACAGCGGAATGTTCTT
ENTPD2	F R	CTCAACCTGACCAATCCA CCAGCCGTACTTGATGAA	TKT	F R	CATTATTATGAAGGTGGCATTGGT CAGGTGGGTGACAGTGAT
ENTPD3	F R	CAGATCCACAAGCAAGAG GATTCCAGAGCCTTTTCAC	RRM2	F R	CAAGAAGAAGGCAGACTG AGTGTAACCCTCATCTCT
FHIT	F R	CAGTCTGTAAAGGTCCGTAGTG GGTCTAAGCAGGCAGGTATTC	SLC2A12	F R	TGACTGTAAGTATCTTATTGG ATTCTTCTTGGTGATGACTC
POLR3K	F R	ACCCGCAAGGTAACAAAT GAAGTAAGCACGAGGATGT	TK1	F R	GGCAGATCCAGGTGATTC GCGAGTGTCTTTGGCATA
PPT1	F R	CGCAACCACAGCATCTTCT AAATCCAAACCACTCCGAATCTAC	RBKS	F R	ATAACTCCAGCAACTTCTT AATCTCAGCCTCACTTTC

3.2 RESULTADOS E DISCUSSÃO

3.2.1 Extração de RNA e sequenciamento

Verificou-se a concentração e qualidade dos RNAs obtidos por espectrofotometria a 260 nm/280 nm e eletroforese em gel de agarose 1% corrido em condições desnaturantes, respectivamente. A integridade do RNA total extraído foi julgada baseado na qualidade das bandas correspondentes às subunidades de rRNA 28S e 18S (Figura 8).

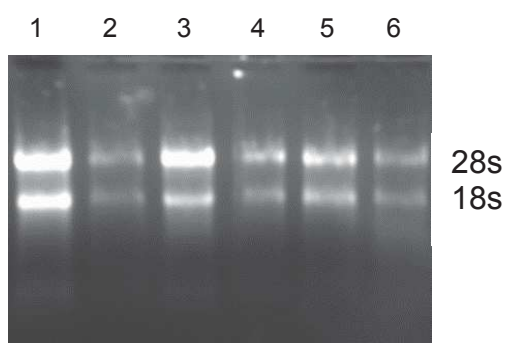


Figura 8. Integridade do RNA total extraído das linhagens indicadas através de eletroforese em gel de agarose 1% seguido de coloração com brometo de etídeo. Banda maior e menor detectadas correspondem às subunidades ribossomais 28S e 18S, respectivamente. 1, MDA-MB436; 2, MDA-MB468; 3, BT-549; 4, MCF7; 5, MDA-MB231; 6, SKBR3.

3.2.2 Validação dos dados de RNA-Seq *in house* por qPCR

O mesmo lote de extração de RNA utilizado para as análises de RNA-Seq foi empregado nos experimentos de qPCR com o objetivo de confirmar os dados obtidos.

Este passo foi importante dado que o sequenciamento das linhagens não foi realizado em replicata.

Estabeleceu-se primeiramente a eficiência dos pares de *primers* de alguns dos genes, selecionados ao acaso, tendo sido obtido valores maiores que 90% (tabela 5 do anexo.). Uma curva de *melting* foi aplicada aos amplicons para confirmar-se a obtenção de somente um pico de *melting* (dados não mostrados).

Os valores de RSEM obtidos para uma dada linhagem e um dado gene (Tabela 3) foram comparados, na forma $\log_2(RSEM+1)$, com os valores de $1/\Delta Ct$ para mesma linhagem e gene obtidos por qPCR. Com esta comparação, obteve-se coeficientes de correlação de Spearman entre 0.39 a 0.66 (Figura 9).

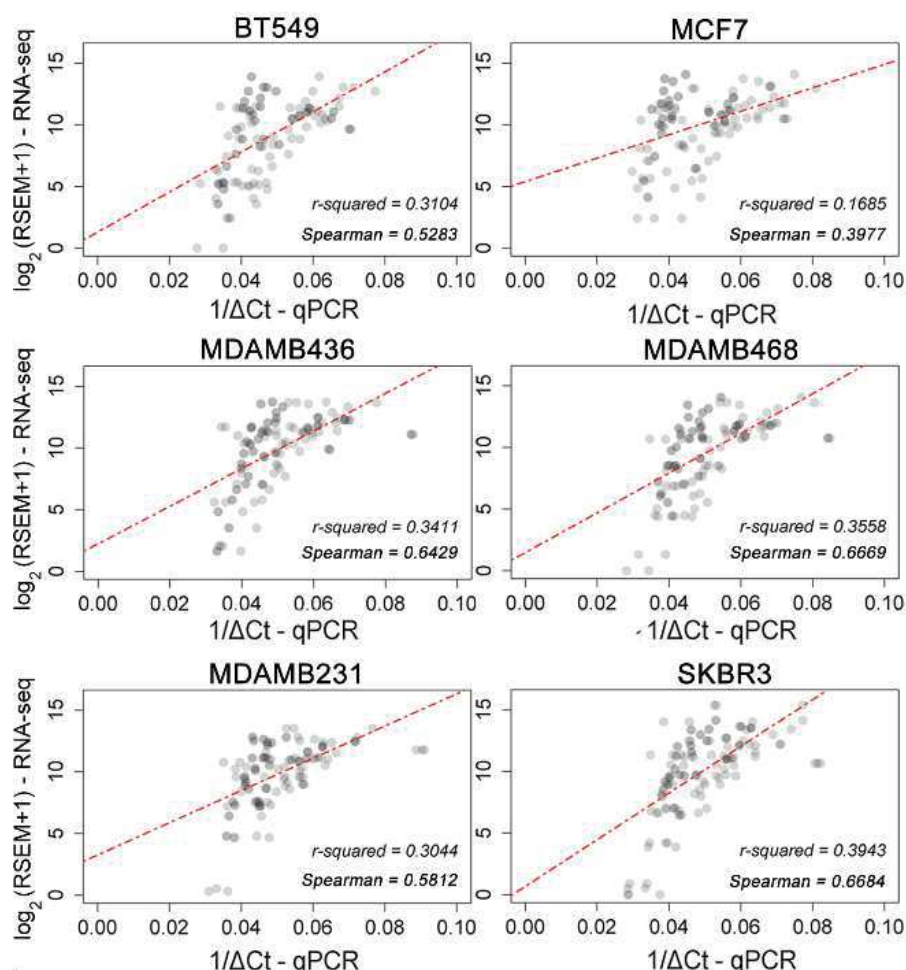


Figura 9. Correlação dos dados de expressão gênica dos genes escolhidos para validação. Comparou-se os valores de $1/\Delta Ct$ (obtidos pelo qPCR) com o de $\log_2 RSEM+1$ da amostra sequenciada *in-house*.

3.2.3 Análises da qualidade dos dados do sequenciamento

Com o emprego do programa STAR verificamos que mais de 90% dos *reads* obtidos foram alinhados (Tabela 4). Primeiramente, empregando-se o programa *skewer*, foi realizado o *trimming* das sequências com base na qualidade da última base sequenciada, progressivamente removendo bases até a primeira com qualidade superior a PHRED⁸³ 20, mantendo-se o restante da sequência até um limite de, no mínimo, 30 bases. Após, isto, avaliou-se a qualidade média das sequências de maneira a manter-se somente aquelas com qualidade média PHRED acima de 30. Para ilustrar o ganho de qualidade de sequência obtido com o procedimento antes e após o *trimming* realizado com o *Skewer*, o programa FastQC, mediante o índice de qualidade PHRED, foi utilizado (Figura 10, exemplo para SKBR3; demais linhagens Figura 1 do Anexo I). Verifica-se que, após o *trimming*, obteve-se uma utilização de mais de 90% dos *reads* (Tabela 5).

Tabela 4. *Reads* alinhados com o programa STAR (%).

Linhagens	STAR
MDA-MB-231	95,97%
SKBR3	95,80%
MCF7	92,21%
BT549	94,97%
MDA-MB-468	94,72%
MDA-MB-436	94,75%

Tabela 5. Porcentagem de sequências mantidas depois do corte utilizando-se o programa *Skewer*.

Trimming com Skewer				
Origem dos dados	Linhagem	Número de reads		Mantidas
		Input	Output	
Laboratório HTSF, Universidade de Carolina do Norte, Chapel Hill, EUA.	MDA_MB231	135033857	130043459	96,30%
	SKBR3	74159211	71605957	96,56%
Lactad, Unicamp, SP, Brasil	MDA_MB468	38240082	36918188	96,54%
	MDA_MB436	35782056	34642593	96,82%
	BT549	26908190	26066982	96,87%
	MCF7	29220227	28357542	97,05%

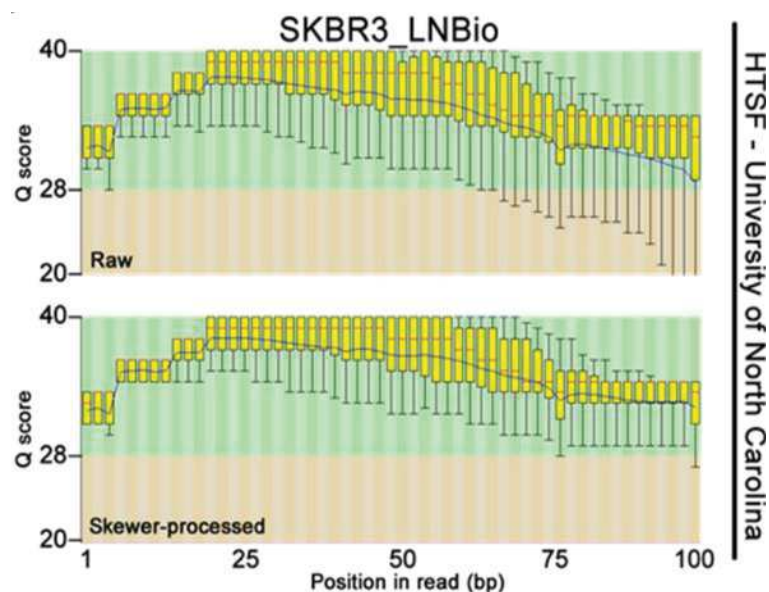


Figura 10. Avaliação da qualidade dos *reads* da linhagem SKBR3 antes (RAW) e após o *trimming* (*skewer-processed*). *Box-plot* dos valores de PHRED (Q score) para cada posição de base para todas as sequências lidas tanto antes (acima) quanto depois (abaixo) do *trimming*. Sem o *trimming* (RAW) evidencia-se uma redução da qualidade no final das posições dos *reads* por pb, típica do processo de sequenciamento. Dados com *trimming* apresentam melhora significativa nos valores dos scores.

3.2.4 Definição da lista de genes diferencialmente expressos entre tecidos e linhagens TN versus não-TN

Para obter o perfil de expressão gênica de tecidos TN *versus* não-TN, classificamos 1100 casos do banco TCGA de acordo aos níveis de expressão dos receptores de estrogênio (ER, gene *ESR1*), progesterona (PR, gene *PGR*) e do receptor 2 do fator de crescimento epidermal (HER2, gene *ERBB2*), utilizando os níveis de expressão normalizados de cada gene (Figura 11A). Quatrocentos e setenta e oito casos não apresentavam a análise imunohistoquímica (IHC) (Figura 11B). Assim, utilizamos a nossa metodologia de classificação por expressão gênica para poder inclui-los nas análises (Figura 11C). Esta metodologia foi comparada com a classificação imunohistoquímica disponível, tendo sido verificado 95.3% de acerto com o fenótipo TN (Figura 11D). Apresentamos na Figura 11E *boxplots* representativos da expressão gênica de cada um dos marcadores em tecidos normais TNBC e não TNBC.

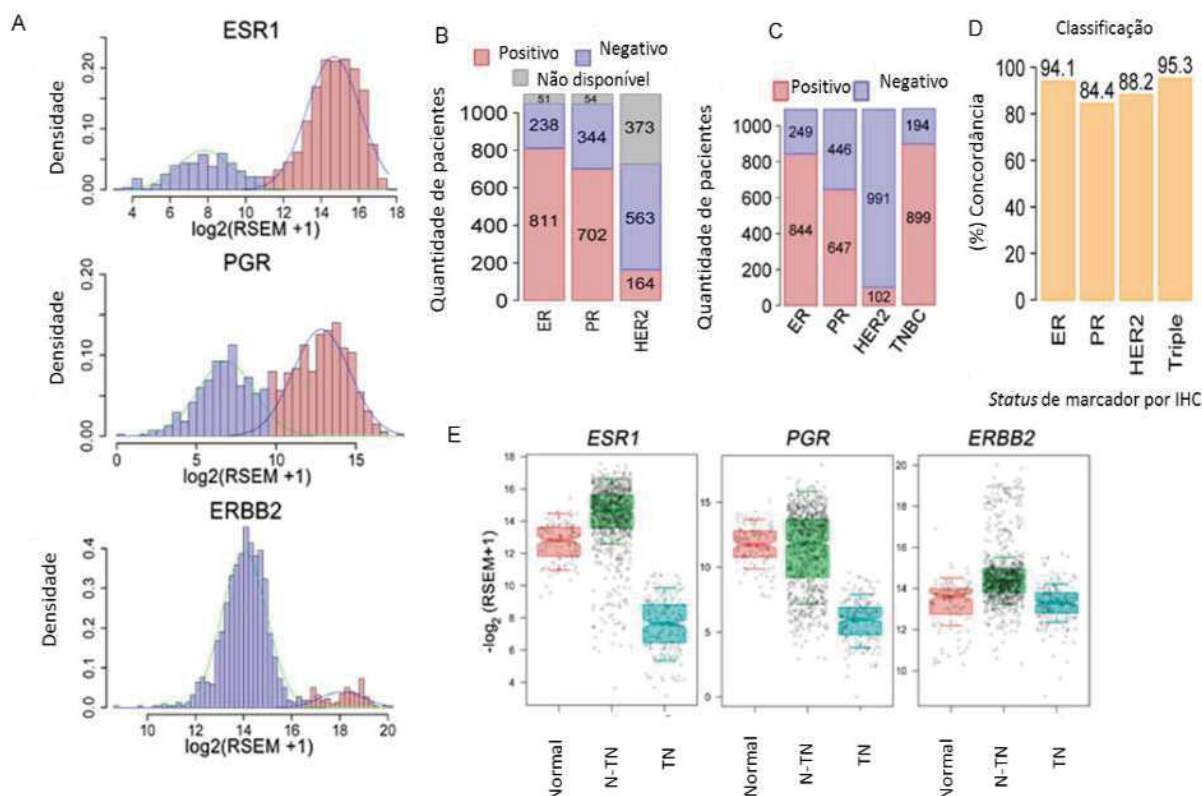


Figura 11. A. Dados processados a partir de tumores de mama obtidos a partir do TCGA foram usados para definir amostras em função da expressão dos genes *ESR1*, *PGR* e *ERBB2* após ajuste bimodal sobre os graficos de densidade de expressão dos mesmos. B. *Status* da expressão dos marcadores por IHC; amostras em cinza foram incluídas em nossas análises após a classificação por expressão gênica C. *Status dos pacientes após reclassificação*. D. Porcentagem de concordância com a classificação imunohistoquímica (IHC) disponível. E. *Boxplots* da expressão gênica para cada um dos marcadores em tecidos normais, TN e N-TN. $\log_2 \text{RSEM} + 1$ do *ESR*, *PGR* e *ERBB2*. Círculos em cinza denota cada amostra. Entalhes denotam o intervalo de confiança de 95% da mediana.

Após as classificações, o perfil de expressão gênica dos tumores TN foram comparados com o dos não-TN para definir os genes que estão DE usando as rotinas empregadas pelo DESeq e pelo EBSeq. Ambos métodos indicaram uma grande quantidade de genes alterados sendo que muitos se sobrepuseram entre as duas análises (Figura 12A). Desta maneira, nós prosseguimos com as análises do EBSeq. Após à aplicação dos parametros (*Fold Change* $\geq +2$ e ≤ -2 , e PPEE < 0.05) para selecionar os genes diferencialmente expressos, obteve-se 1404 genes *down regulated* e 1555 *up-regulated* (Figura 12B).

Para confirmar se a lista de genes diferencialmente expressos era capaz de distinguir entre os grupos de pacientes TN e não-TN realizou-se uma análise de componentes principais (PCA), que mostrou uma separação clara entre os dois grupos (Figura 12C). Igualmente, usando outro tipo de agrupamento hierárquico, verificamos que os genes diferencialmente expressos também foram capazes de separar os casos TN dos não-TN (Figura 12D).

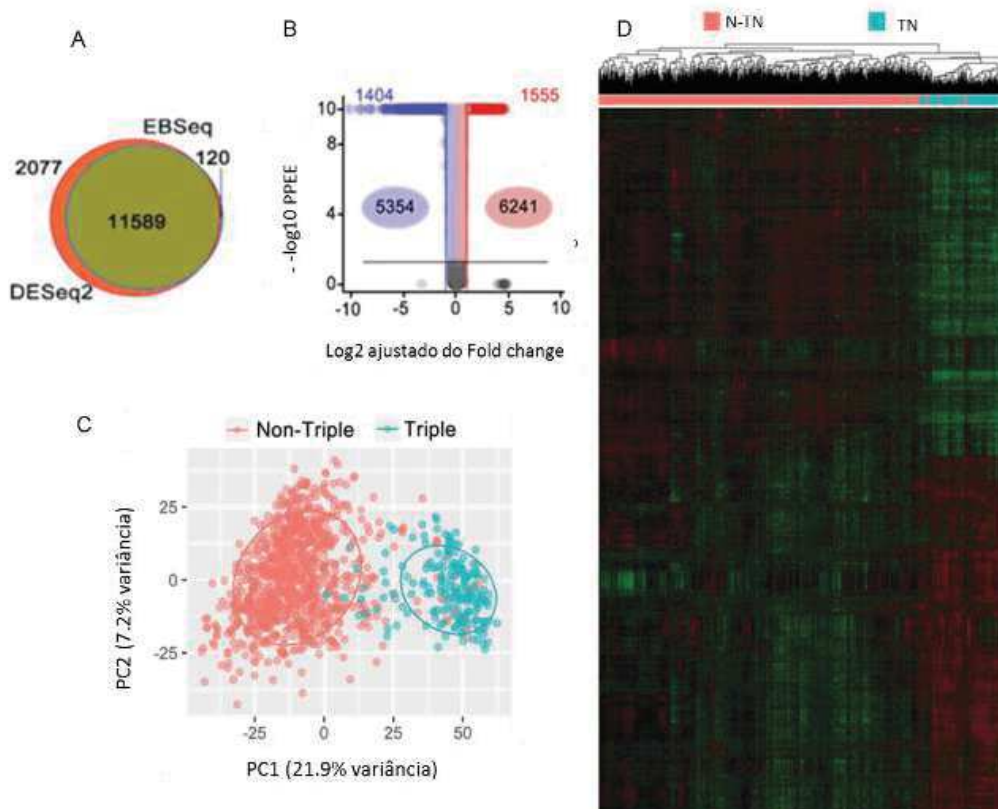


Figura 12. Análise de expressão diferencial de tecidos tumorais de mama TN e não-TN. (A) Comparação de genes diferencialmente expressos segunda análise do EBSeq e DESeq2 indicando uma concordância entre os programas na definição de genes alterados. (B) *Volcano-plot* dos dados expressão gênica versus PPEE: genes em azul genes reprimidos e em vermelho genes induzidos. Abaixo da linha horizontal preta, em cinza genes sem significancia estatística $PPEE > 0.05$; genes não diferenciais ($fold\ change \geq -2$ e $\leq +2$ e $PPEE \leq 0.05$) estão como pontos azuis claros (5354 genes) ou vermelhos (6241 genes); genes com expressão diferencial ($fold\ change \geq +2$ ou ≤ -2 e $PPEE \leq 0.05$) estão representados em pontos azul escuro quando reprimidos (1404 genes) ou vermelho escuro quando super-expressos (1555 genes). (C) Análise de componentes principais usando os 2959 genes diferencialmente expressos mostrando uma clara divisão entre pacientes TN e não-TN. (D) Agrupamento hierárquico e *heatmap* de genes induzidos (vermelhos) ou reprimidos (verdes) também evidenciando a divisão entre TN e não-TN. As correlações foram obtidas através do coeficiente de Pearson; O agrupamento foi feito pelo método completo, e os valores de ambos eixos foram escalonados por linha \log_2 (RSEM + 1).

As linhagens celulares cujos RNA-seq foram feitos *in-house* ou foram obtidos de bancos de dados foram classificadas de acordo com os dados da literatura e confirmados pelo perfil de expressão dos mesmos. Conforme esperado, as mesmas apresentaram perfil de expressão dos genes *ESR1*, *PGR* e *ERBB2* condizente com sua classificação (Figura 13). Após aplicação de um filtro para considerar somente os genes DE com *Fold Change* $\geq +2$ e ≤ -2 , e com *PPEE* ≤ 0.05 , obteve-se 1699 genes *down regulated* e 2309 *up-regulated* (Figura 14A).

A lista de genes DE entre as linhagens TN e não-TN foram utilizados para um agrupamento no *cluster* hierárquico (Figura. 14B), e, conforme esperado, foi possível confirmar a identidade das linhagens celulares sequenciadas *in-house*, dado que as mesmas se agruparam próximo às mesmas linhagens cujos dados de RNA-seq estavam disponíveis publicamente. Desta maneira, nos estudos descritos a frente de vias biológicas e metabólicas alteradas utilizando as informações de linhagens celulares, as obtidas de bancos públicos foram utilizadas como replicatas das sequenciadas por nós.

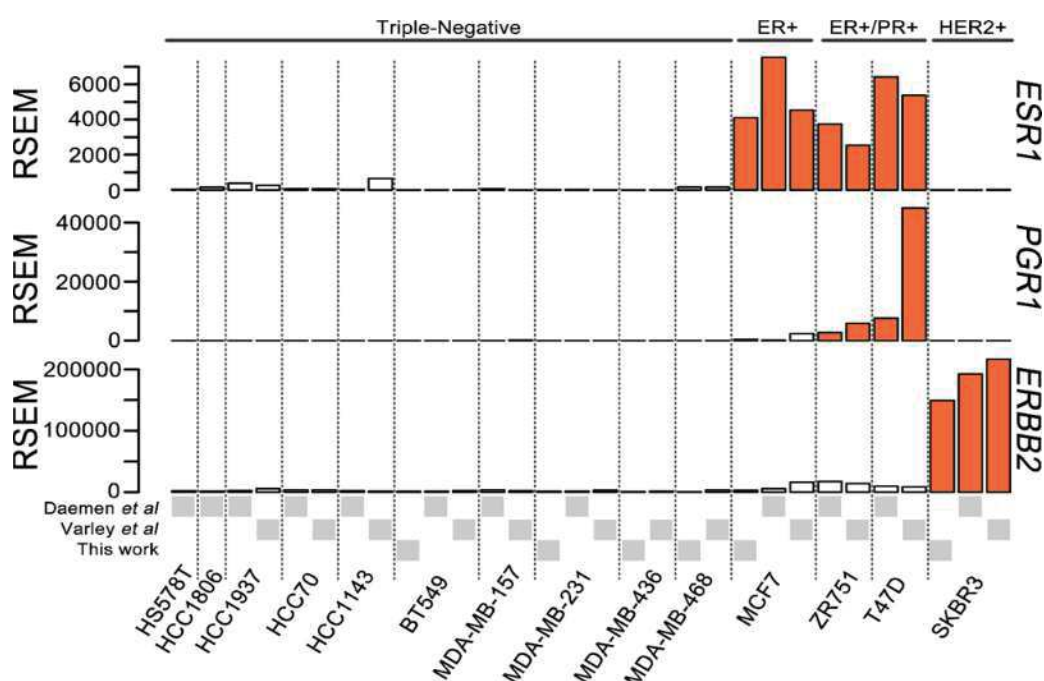


Figura 13. Valores RSEM para os genes *ESR1*, *PGR* e *ERBB2* dos dados obtidos de nosso sequenciamento e/ou de banco de dados (Daemen et al e Varley et al). O quadrado cinza indica a fonte dos dados. Linhagens não-TN (MCF7; ZR751; T47D; SKBR3) e linhagens TN (HSS578T; HCC1806; HCC1937; HCC70; HCC1143; BT549; MDA-MB157; MDA-MB-231; MDA-MB-436; MDA-MB-468).

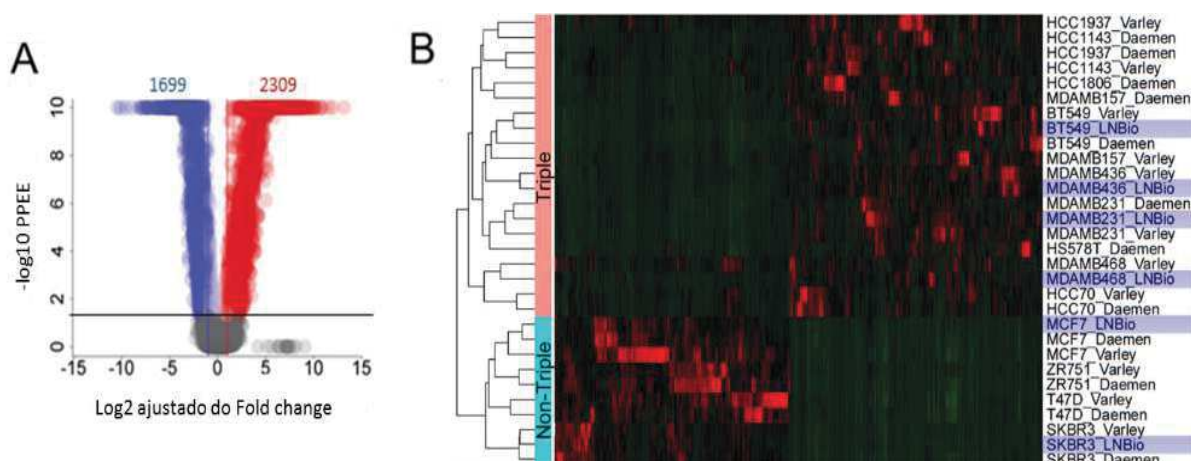


Figura 14. Análise de expressão gênica diferencial entre linhagens celulares TN e não-TN obtidas *in house* (LNBio) e do portal GEO (Varley e Daemen)^{68,69}. (A) *Volcano-plot* dos dados expressão gênica versus PPEE: genes em azul genes reprimidos e em vermelho genes induzidos. Abaixo da linha horizontal preta, em cinza genes sem significancia estatística $PPEE > 0.05$; genes diferenciais e acima do corte estatístico estão como pontos azuis escuros (1699 genes) quando reprimidos ou vermelhos (2309 genes) quando iduzidos. (B) *Cluster hierárquico* denotando genes induzidos ($fold\ change \geq +2$) em vermelho e reprimidos ($fold\ change \leq -2$) em verde com $PPEE \leq 0.05$. As correlações foram obtidas através do coeficiente de Pearson; O agrupamento foi feito pelo método completo, e os valores de ambos eixos foram escalados por linha \log_2 (RSEM + 1). É possível verificar que o perfil de expressão das nossas linhagens celulares (caixas azuis, LNBio) permitem as mesmas serem agrupadas com as linhagens correspondentes sequenciadas em outros laboratórios.

3.2.5 Análises de vias alteradas pelo Metacore

O estudo de linhagens celulares tumorais estabelecidas, apesar de carregar inerentes e conhecidos artefatos que não os torna uma amostra fidedigna do que acontece nos tecidos *in vivo*, ainda é uma ferramenta muito usada para o entendimento da biologia do câncer e para a validação de alvos tumorais em amostras humanas. Assim, é importante conhecer e comparar processos tumorais tanto em tumores extraídos de pacientes como em linhagens representativas.

Desta maneira, os dados de expressão gênica diferencial entre tecidos tumorais TN e não-TN, assim como entre linhagens TN e não-TN cultivadas e sequenciadas *in-house* ou obtidas de bancos públicos foram avaliados no Metacore para entendermos as alterações de processos biológicos e moleculares que ocorrem no subtipo TN. Mais, neste segundo grupo, os resultados de metabolômica foram incorporados à análise, também no Metacore, para complementar os achados.

3.2.5.1 Análises de processos biológicos e vias metabólicas

A lista de genes diferencialmente expressos obtidos entre tecidos e linhagens TN *versus* não-TN foram cruzadas e os genes em comum estabelecidos (Figura 15).

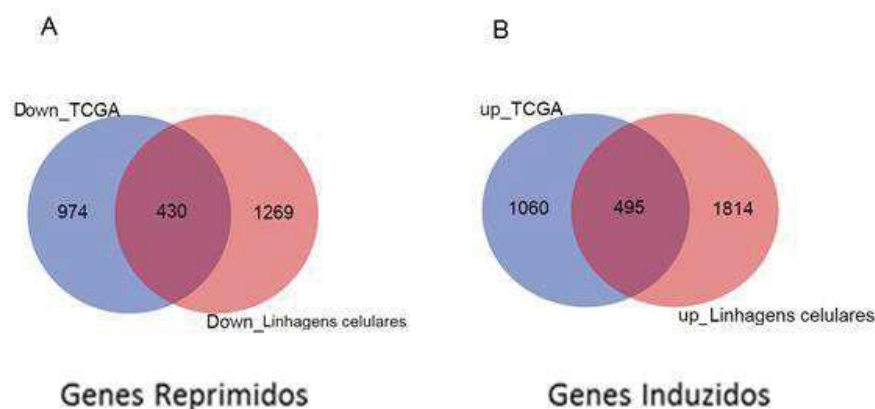


Figura 15. Interseção de genes DE entre linhagens celulares *versus* tecidos tumorais (TCGA) após a filtragem do $IFC \leq -2$ e $\geq +2$ e corte de $PPEE \leq 0.05$. A. Genes reprimidos. B. Genes induzidos.

Ao realizar as análises de enriquecimento pelo *software* Metacore da lista de genes DE em tecidos tumorais e linhagens celulares, verificamos que alguns processos são enriquecidos de maneira em comum aos dois tipos de amostras, dado um corte de p-valor de 0.05 (Tabela 6). Três processos (Adesão celular e remodelamento ECM; Remodelamento dos filamentos do citoesqueleto; Adesão celular _matriz celular glicoconjugados) estão potencialmente ligados a processos de reestruturação de citoesqueleto, adesão celular, remodelamento de matriz e potencialmente migração/invasão, indicando que estes processos podem estar mais alterados em tumores TN.

Há ainda enriquecimento na via de coregulação da transcrição mediada por coactivadores YAP/TAZ, os quais, quando não fosforilados, interagem com fatores de transcrição da família TEF (*Transcription enhancer factor*), SMAD, RUNX2, KLF5, p73, entre outros, regulando processos celulares tais como proliferação, sobrevivência, diferenciação, renovação de *stem-cells* e embriogênese⁸⁴. Genes como Oct-3/4, SOX9 estão com expressão aumentada indicando que, em TN, características *stem-like* estão ativadas. De fato, sabe-se que tumores TN são majoritariamente basais e

claudin-low e apresentam características mesenquimais, com alta incidência de câncer *stem-cells*⁸⁵.

Por fim, vale a pena notar que tumores TN (tanto tecido quanto linhagens) apresentaram genes diferencialmente expressos dentro da via de “Resposta imune_ IL - 4 induzidos, reguladores do crescimento, sobrevivência, metabolismo e diferenciação celular”. Em células não hematopoiéticas, IL-4 sinaliza através do receptor IL-4R tipo II e regula várias funções celulares, coordenando a resposta imune com a atividade metabólica de vários tecidos. Dentro desta via verifica-se o aumento na expressão de várias ciclinas (A, E e D), CDK6, CDC25A, potencialmente indicando ativação de ciclo celular, o que é coerente com a alta taxa de proliferação destes tumores. Ainda, dentro desta via, há aumento na expressão de PPAR γ e α , PGAR (*angiopoietina like 4*) e ACADL (*long-chain specific acyl-CoA desidrogenase*), indicando aumento de beta oxidação.

Tabela 6. *Pathways maps* mais afetados tanto em tecido quanto em linhagens TN versus não-TN.

<i>Pathways maps</i>	Linhagens celulares	#genes/#genes alterados	Tecidos (TCGA)	#genes/#genes alterados
	p-valor		p-valor	
Adesão celular e remodelamento ECM	3,233E-21	52/38	2,425E-07	52/20
Co-regulação da transcrição mediada por coativadores YAP/TAZ	7,019E-11	56/28	4,415E-06	56/19
Remodelamento dos filamentos do citoesqueleto	4,031E-10	36/21	2,386E-06	36/15
Ativação independente dos receptores de ESR1 e ESR2	1,025E-05	63/18	8,422E-04	63/21
Adesão celular _matriz celular glicoconjugados	2,772E-04	45/18	2,787E-05	45/11
Resposta imune_ IL - 4 induzidos, reguladores do crescimento, sobrevivência, metabolismo e diferenciação celular	1,384E-03	38/12	7,891E-06	38/14

Uma outra análise foi feita focada em *metabolic networks* (Tabela 7). Da lista de vias enriquecidas tanto em tecidos quanto em linhagens TN em relação a não-TN, destaca-se a alta incidência de processos relacionados a sinalização/metabolismo de

lipídeos como a *via da fosfatidilcolina*, um dos principais componentes das membranas, fonte de segundos mensageiros e determinante da progressão do ciclo celular⁸⁶. A *via da fosfatidilinositol -4,5 – difosfato*, um dos componentes principais da via de sinalização da PI3K, também está enriquecida. A *via de metabolismo da N-acil-esfingosina fosfato*, um dos reguladores importantes do metabolismo de esfingolipídios, é outra via também enriquecida⁸⁷. Nesta via aparecem diversos genes de colágeno, integrinas e proteases de remodelamento de matriz com expressão aumentada, sugerindo aumento de migração celular. De fato, como já comentado, TN são tumores altamente agressivos e invasivos⁸⁸.

Tabela 7. *Metabolic networks* mais afetadas tanto em tecido quanto em linhagens nos tumores TN

	Linhagens celulares	#genes/#genes alterados	Tecidos (TCGA)	#genes/#genes alterados
	p-valor		p-valor	
Via fosfatidilcolina	8,902E-09	130/37	2,052E-10	130/35
Via fosfatidilinositol -4,5 - difosfato	3,189E-08	92/29	1,475E-03	92/17
Via de fosfato de N - acil- esfingosina	4,280E-07	108/30	6,874E-06	108/24
Via L - fosfatidilserina	2,239E-04	81/20	2,836E-05	81/19
Via fosfatidiletanolamina	2,790E-03	91/19	4,804E-05	91/20
Via -1 – docosaheptaenoyl -glicerol-3-fosfocolina	5,159E-05	128/29	2,089E-03	128/21
Via 1,2- dioleoil - sn- glicerol-3-fosfato	8,922E-05	94/23	1,884E-03	94/17
Via fosfatidilinositol- 3,4,5 - trifosfato	3,245E-03	66/15	2,875E-04	66/15
Via 2 - oleoil glicerol-3 - fosfato	5,692E-04	45/13	3,508E-03	45/10

3.2.5.2 Expressão gênica diferencial associado a dados de metabolômica de linhagens celulares indicam vias alteradas nos tumores TN

Dados de metabolômica obtidos no laboratório por espectros de ¹H medidos por RMN foram incorporados aos dados de expressão gênica (*in-house* e de bancos públicos para compor replicatas) das linhagens TN MDA-MB157, MDA-MB231, MDA-MB436, MDA-MB468, BT549, HCC1806, HS578T e as não-TN SKBR3, MCF7, ZR751. Diferentes processos metabólicos como o metabolismo de aminoácidos, o metabolismo de lipídeos, entre outros, estão alterados nas células tumorais TN em relação as não-TN (Tabela 8). A concentração medida de cada um dos metabólitos

intracelulares encontra-se na Tabela 1 do Anexo I. Os esquemas de alguns dos *metabolic networks* alterados são mostrados nas Figuras 2-6 do Anexo 1.

Tabela 8 *Metabolic networks* mais afetadas nas linhagens tumorais TN *versus* não-TN. Vias em negrito são mostradas nas Figuras 2-6 do Anexo I.

<i>Metabolic networks</i>	Min (p - Valor)
Metabolismo de aminoácidos Ala,Ser,Cis,Met,His,Pro,Gli,Glc,Gln	3,124E-11
Metabolismo de Lipídeos Glicosfingolipidos	3,124E-11
Metabolismos de aminoácidos Arginina	5,566E-10
Metabolismo de aminoácidos Asparagina, ácido aspártico	7,096E-10
Via da fosfatidilcolina	1,225E-09
Via da n-acil esfingosina fosfato	3,702E-09
Metabolismo de aminoácidos Asparagina, ácido aspártico, Arginina e transportadores	1,086E-08
Glutaminólise e transporte do glutamato.	1,626E-08
Via da L-ornitina e transportador	2,230E-07
Metabolismo de aminoácidos Alanina, Glicina, Cisteína e transportadores	2,902E-06
Via e transporte do L-glutamato	3,375E-06
Via e transporte da L-serina	5,604E-06
Metabolismo da esfingomielina	7,973E-06
Via da (L)-valina	1,450E-05
Via do (L)-fenilalanina	5,797E-05
Glutaminólise.	1,183E-04
Via da Alanina	1,269E-04
Via 2-oleoil-glicerol-3-fosfato	1,867E-04
Metabolismo da beta alanina	4,177E-04
Via da L-citrullina	4,444E-04
Via da tirosina	4,444E-04
Metabolismo de aminoácidos fenilalanina, metionina, triptofano	2,106E-03
Via 1-linoleoil-glicerol_3-fosfato	2,915E-03
Via 1-acil-glicerol_3-fosfatoenalonamina	3,047E-03
Via do ácido d-glucurônico	4,413E-03
Via do 1-docosahexaenoil-glicerol_3-fosfocolina	5,016E-03
Metabolismo de carboidratos; metabolismo de propionato	8,304E-03
Via 1-icosatrienoil-sn-glicero-3-fosfocolina	1,119E-02
Via GalNAc β 1-3Gal	1,155E-02
Vias -treonina e transportes	1,374E-02
Vias 6'-sialil-lactose e transportadores	1,728E-02
Metabolismo de carboidratos; TCA e transporte de ácido tricarbóxico	1,894E-02
Via (L) prolina e transportadores	2,276E-02
Metabolismo lipídico; Triacilglicerol	2,304E-02
Via da leucina	2,618E-02
Biossíntese de CoA	3,445E-02
Via 1-hexadecanoil-glicerol_3-fosfato	3,475E-02
Via 1-oleoil-glicerol-3-fosfato	3,736E-02

3.2.6 Análises de vias metabólicas curadas manualmente

Por fim, realizamos uma análise de vias metabólicas alteradas em TN em relação a não-TN utilizando mapas construídos por nós e focados em vias selecionadas. Oito vias metabólicas (glicolítica, gliconeogênica, glutaminolítica, ciclo do TCA, metabolismo de lipídeos, de síntese de ribose 5-fosfato, de hexosamina e degradação do etanol), com um total de 200 genes, foram avaliadas (Tabela 2 do Anexo 1). Dos genes avaliados, 17 deles (*HK3*, *PFKP*, *ENO1*, *LDHB*, *ALDH1A3*, *ALDH3A1*, *GLS1*, *ACSL6*, *ELOVL3*, *ELOVL4*, *SLC2A1*, *SLC2A6*, *SLC2A12*, *SLC2A14*, *SLC16A1*, *SLC16A8*, *SLC16A10*) estavam induzidos e 15 genes (*FBP1*, *PCK1*, *GLUL*, *GLS2*, *ACACB*, *ACADSB*, *ACOT4*, *ACOX2*, *ELOVL2*, *ELOVL5*, *SLC1A1*, *SLC1A2*, *SLC1A4*, *SLC2A10*, *SLC5A4*) estavam com expressão diminuída em tecidos tumorais TN em relação a não-TN (Figura 16). Por outro lado, nas linhagens tumorais, 22 genes (*HKDC1*, *PGM1*, *ALDH1A1*, *ALDH1A3*, *ALDH1B1*, *ALDH3A1*, *ACO1*, *ME3*, *GLS1*, *PRPS2*, *UAP1*, *GFPT2*, *ACOT9*, *ACSL4*, *CPT1C*, *SLC1A3*, *SLC1A6*, *SLC2A3*, *SLC2A9*, *SLC5A1*, *SLC16A1*, *SLC16A7*) estavam induzidos e 25 genes (*ALDOC*, *LDHD*, *FBP1*, *ACSS1*, *ACSS3*, *ALDH2*, *ALDH3B2*, *IDH2*, *MDH1B*, *GLS2*, *G6PD*, *ACACB*, *ACOT4*, *ACOX2*, *CPT1A*, *ELOVL3*, *FASN*, *PPT1*, *SLC1A2*, *SLC2A4*, *SLC2A8*, *SLC2A10*, *SLC2A11*, *SLC2A14*, *SLC16A8*) estavam com expressão diminuída (Figura 17).

3.2.6.1 Tecidos tumorais

Os tecidos tumorais TN apresentam um aumento da expressão de vários transportadores de açúcar, principalmente glicose (família GLUT ou *SLC2A1*, *SLC2A6*, *SLC2A12* e *SLC2A14*) indicando uma alta demanda deste nutriente para este subtipo tumoral (Figura 16). De acordo com o aumento na expressão de transportadores de glicose, há um aumento da expressão de enzimas da via glicolítica hexoquinase 3 (*HK3*), fosfofrutoquinase P (*PFKP*), enolase I (*ENO1*) e lactato desidrogenase B (*LDHB*) nos tecidos tumorais TN.

Diferente das outras isoenzimas de hexokinase, a isoenzima 3 tem maior afinidade aparente pelo substrato glicose, menor afinidade pelo inibidor (e regulador) ATP, assim como pelo produto, glicose 6 fosfato (também inibidor da reação)^{89,90}, indicando que esta isoenzima é a mais ativa das variantes. Sabe-se que a super-expressão de *HK3* reduz a morte celular induzida por estresse oxidativo, aumenta os

níveis de ATP, reduz a produção espécies reativas de oxigênio (ROS) e preserva o potencial de membrana mitocondrial⁹¹. HK3 está induzida em amostras de leucemia linfoblástica juvenil⁹² mas não há relatos sobre seu *status* de expressão em tumores TN.

O aumento na expressão da fosfofrutokinase isoforma de plaquetas (PFKP, *Phosphofructokinase isoforme platelet*) tem sido reportada como uma característica de tecidos malignos⁹³, sendo, em tumores de mama, regulado pelo fator de transcrição KLF4⁹⁴. Em acordo com o verificado por nós, recentemente foi mostrado que este gene tem expressão aumentada em linhagens TN em relação a linhagens não-TN⁹⁵.

A enzima ENO1, apesar de estar envolvida na via da glicólise, tem sido reportada estar presente na superfície de células tumorais, sendo importante para a migração e a invasão tumoral⁹⁶. O aumento da expressão de ENO1 em TN não foi ainda descrito, de maneira que explorar a importância da mesma para este subtipo tumoral tem o potencial de relevar um novo alvo para o tratamento destes tumores, com potencial impacto sobre o processo de metástase.

Um achado interessante foi o aumento na expressão dos transportadores de lactato MCT1(SLC16A1) e MCT3 (SLC16A8) e do gene LDHB. LDHB é mais eficiente na conversão de lactato a piruvato, enquanto que LDHC e LDHA favorecem o sentido convencional (piruvato a lactato). Enquanto estes dois últimos favorecem a reciclagem de NAD e manutenção do fluxo glicolítico, especula-se que células com expressão aumentada de LDHB apresentem uma vantagem na utilização de lactato como fonte energética e biossintética. Já foi reportado que, no microambiente tumoral, células em hipóxia e em aerobismo comunicam nutriente entre si, de maneira que células hipóxicas produzem e secretam lactato que acaba por ser consumido pelas células com suprimento de oxigênio⁹⁷. De fato, já se sabe que tumores TN tem a expressão de LDHB elevada, e que seu *knock down* diminui a proliferação celular *in vitro*, assim como o crescimento de tumor *in vivo* ⁴⁸.

Observamos uma curiosa diminuição de GLS2 (e aumento de GLS1) nas células TN. De fato, aumento na expressão e dependência de GLS1 já foi reportado em tumores TN⁴¹. Outro gene diminuído é o da glutamina sintetase (*Glutamate ligase*, GLUL). De acordo, sabe-se que tumores TN são altamente dependentes de glutamina exógena^{41,50}. O porquê a diminuição de GLS2 em TN é desconhecido por nós, mas dados recentes ainda não publicados do nosso grupo mostram que a expressão de

GLS2 pode ser induzida pela diminuição na atividade de GLS1, o que leva ao resgate na proliferação de células TN (dados não publicados).

Além do metabolismo da glicose e glutamina, o aumento da biossíntese de macromoléculas, particularmente lipídeos, foi reconhecido como um componente da reprogramação metabólica em células tumorais⁹⁸. Os lipídeos sintetizados, além de serem fontes de energia e ativadores de sinalização celular, contribuem na geração de estruturas que são importantes para a formação de membranas biológicas. Um dos genes com expressão aumentada em TN foi o ACSL6 que codifica para a proteína de membrana acil-CoA sintetase de cadeia longa. Estas proteínas são necessárias para a degradação de ácidos graxos, remodelação de fosfolipídios e produção de ésteres de acil-CoA de cadeia longa que regulam muitos processos fisiológicos⁹⁹. Não foi ainda reportada nenhuma relação entre tumores TN e a expressão deste gene.

Outras proteínas com expressão aumentada nos tumores TN foram as proteínas de alongação dos ácidos graxos ELOVL3 e ELOVL4. O papel destas enzimas na tumorigênese/manutenção de TN ainda não foi estabelecido, e o estudo da sua importância para este subtipo tumoral pode indicar novos alvos. Por fim, é válido comentar sobre o aumento de enzimas da via de oxidação (degradação) de aldeídos, a ALDH1A3 e a ALDH3A1, nos tumores TN. A ALDH1A3, em específico, produz ácido retinóico (AR), uma molécula de sinalização ativadora de um fator de transcrição. Foi recentemente mostrado que a sinalização por AR induzida pela expressão de ALDH1A3 tem um papel ativo na tumorigênese de câncer de mama, com indicação de prognóstico ruim e associação com o subtipo TN^{100,101}. Já o ALDH3A1 está envolvido na oxidação de vários aldeídos aos seus correspondentes ácidos, sendo importante para a detoxificação de acetaldeídos derivados de álcoois e no metabolismo de corticosteroides, neurotransmissores e peroxidação de lipídeos. O acetaldeído pode ser metabolizado a um subproduto menos ativo, o acetato, que é então decomposto em água e dióxido de carbono para facilitar a eliminação¹⁰², se tornando, desta maneira, fonte energética para as células. A super-expressão de ALDH3A1 em TN está relacionada a inativação de ciclofosfamida e quimioresistência¹⁰³, o que torna esta proteína um potencial interessante alvo para tumores TN (Figura 16).

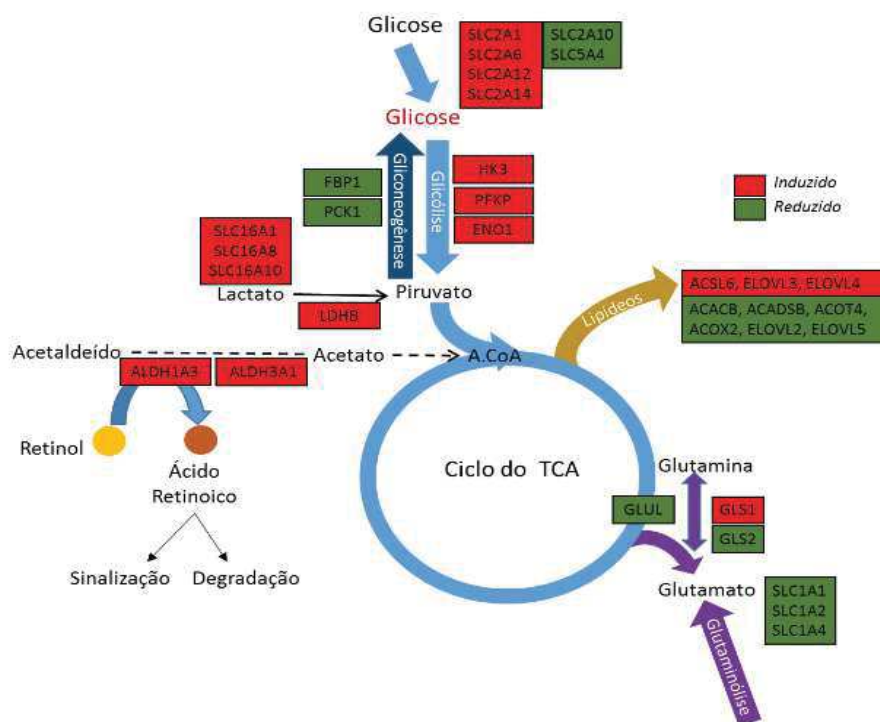


Figura 16. Genes diferencialmente expressos em tecidos tumorais TN em relação aos não-TN (dados foram obtidos do TCGA). A variação da expressão relativa de cada gene é representada pela cor vermelha quando aumentada ($\log_2 \text{fold change} \geq +1$) e verde quando diminuída ($\log_2 \text{fold change} \leq -1$). SLC2A: transportadores de glicose; SLC16A: transportadores de lactato; SLC1A: transportadores de glutamato. HK3: Hexoquinase 3; PFKP: Fosfofrutoquinase; ENO1: Enolase 1; PCK1: piruvato carboxilase 1; FBP1: frutose bifosfatase 1; LDHB: Lactato desidrogenase B. ALDH: Aldeído desidrogenase. GLUL: glutamina sintetase; GLS1: Glutaminase1; GLS2: Glutaminase 2; OGDHL: oxoglutarato desidrogenase isoforma L; ACSL6: acil-CoA sintetase de cadeia longa; ELOVL: proteínas de alongação dos ácidos graxos; ACOT: acil-CoA tioesterases; ACACB: acetil CoA carboxilase beta; ACADSB: acil-CoA desidrogenase.

3.2.6.2 Linhagens celulares

Da mesma maneira que os tecidos tumorais TN, as linhagens TN apresentaram um aumento na expressão de 3 transportadores de açúcares (SLC2A3, SLC2A9, SLC5A1), entre eles, glicose, frutose e galactose, tendo sido a variante do transportador diferente, entretanto. Em acordo com o aumento na expressão de transportadores de açúcares, evidenciou-se o aumento da expressão da *Hexokinase Domain Containing 1* (HKDC1) a quinta hexoquinase descoberta em mamíferos⁸⁹. Esta hexoquinase foi só recentemente descoberta e é sugerida desempenhar um papel chave no metabolismo de glicose em mulheres grávidas. Além disto, tem sido descrita como super-expressa em diversos tipos tumorais e foi reportada

recentemente como um possível alvo terapêutico para câncer de pulmão¹⁰⁴. Sua expressão diferencial em tumores TN ainda não havia sido descrito.

Verificamos também aumento na expressão da fosfoglucomutase 1 (*PGM1*), uma fosfotransferase que catalisa a conversão reversível do produto de degradação de glicogénio, glicose-1-fosfato, em glicose-6-fosfato, substrato que é usado para a glicólise e a via das pentoses fosfato (*pentose phosphate pathway*, PPP). No braço não-oxidativo desta via também verificamos aumento da expressão da enzima fosforibosil pirofosfato sintetase (*PRPS*) que catalisa a síntese de fosforibosil pirofosfato (*PRPP*) e é essencial para a síntese de nucleotídeos. Estas alterações podem indicar que, em cultura de células, possa haver um mecanismo de metabolização de açúcares extra glicólise que seja necessário para o abastecimento do braço não-oxidativo da PPP com o intuito de aumentar a síntese de nucleotídeos. Linhagens TN, assim como os tecidos, também apresentaram aumento na expressão dos transportadores de lactato MCT1(*SLC16A1*) e MCT2 (*SLC16A7*) indicando que metabolização de lactato também possa ser importante para as linhagens em cultura.

Igual ao verificado em tecidos, nas linhagens TN há também um potencial aumento da biossíntese de lipídeos também se encontraram induzidos. Neste sentido, há aumento na expressão da aconitase 1 (*ACO1*) (que converte reversivelmente) isocitrato em citrato, um precursor lipogênico. Também está aumentada a expressão da acil-CoA sintetase de cadeia longa *ACSL4*, envolvida na degradação de ácidos graxos, remodelação de fosfolipídios e produção de ésteres de acil-CoA de cadeia longa que regulam muitos processos fisiológicos⁹⁹. Ao contrário do *ACSL6* (com expressão aumentada em tecidos) o *ACSL4* tem sido reportada como um possível biomarcador de fenótipo agressivo em tumores TN¹⁰⁵.

Da mesma maneira como verificado nos tecidos TN, as linhagens apresentam diminuição na expressão de *GLS2* e aumento de *GLS1*, indicando aumento no processamento de glutamina em glutamato; curiosamente, nas linhagens, 2 transportadores de glutamato estão com expressão aumentada podendo indicar que o *pool* de glutamato intracelular pode ter uma contribuição importante do aminoácido captado do meio.

Ainda sobre o metabolismo de glutamina, nas linhagens celulares, há aumento na expressão de 3 enzimas da via das hexosaminas. Nesta via verificamos a expressão aumentada da Glutamina-Fructose-6-fosfato transaminase (*GFPT2*), uma

enzima que controla o fluxo da glicose na via das hexosaminas e, juntamente com a GLS, converte glutamina em glutamato. Há também aumento na expressão de UDP-N-acetilglucosamina pirofosforilase 1 (*UAP1*), uma enzima que suprime uridina difosfato N-acetilglucosamina (UDP-GlcNAc) para biossíntese de glicoproteínas¹⁰⁶. A UDP-GlcNAc é encontrada em concentrações celulares altas¹⁰⁷, sendo utilizada não só para a O-GlcNAcilação, mas também para a biossíntese de glicanos extracelulares complexos. O processo de O-GlcNAcilação é uma modificação pós-traducional que ocorre no retículo endoplasmático; proteínas S-GlcNAcetiladas estão envolvidas na transcrição, tradução, montagem do citoesqueleto, transdução de sinal, e muitas outras funções celulares¹⁰⁸. O aumento na atividade da via das hexosaminas foi recentemente reportado em tumores de mama basais (com alta incidência de TN), tendo sido relatado sua importância para o aumento da glicólise e sobrevivência a estresse ¹⁰⁹.

Também em acordo como visto nos tecidos tumorais, verificamos em linhagens TN o aumento na expressão de proteínas *ALDH1A3* e *ALDH3A1* da via de oxidação (degradação) de aldeídos. Além destas, as linhagens também apresentaram o aumento na expressão de duas outras proteínas (*ALDH1A1* e *ALDH1B1*). O fato de ambas proteínas estarem aumentadas nos dois tipos de amostra as tornam interessantes alvos (Figura 17).

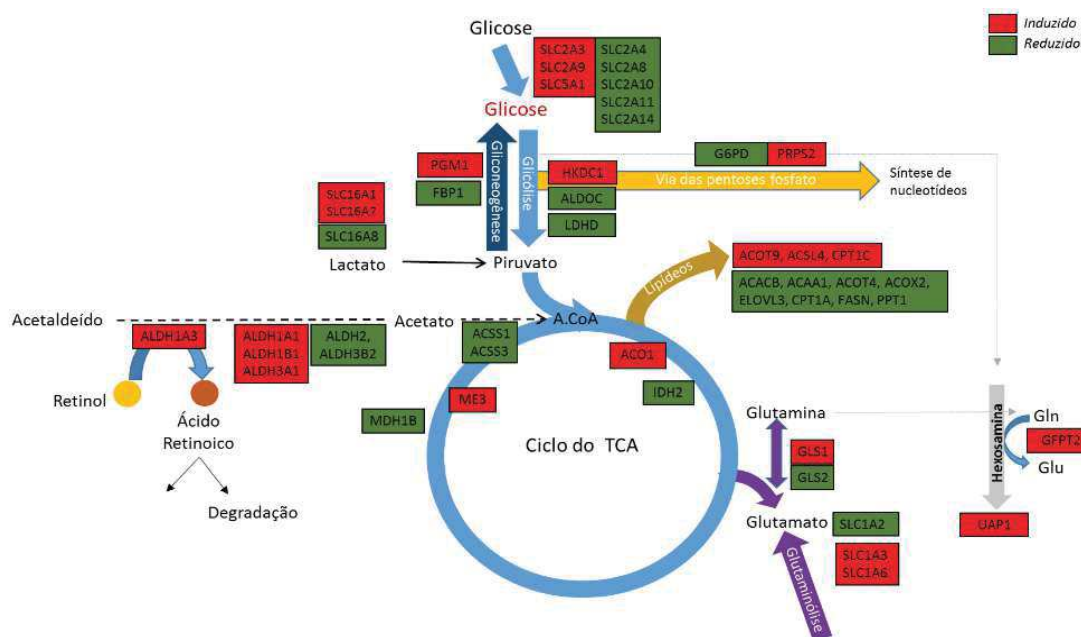


Figura 17. Genes diferencialmente expressos em linhagens tumorais TN em relação aos não-TN (dados foram obtidos *in-house* e de base de dados). A variação da expressão relativa de cada gene foi representada pela cor vermelha quando aumentada ($\log_2 \text{fold change} \geq +1$) e verde quando diminuída ($\log_2 \text{fold change} \leq -1$). SLC2A: transportador de glicose; SLC16A: transportador de lactato; SLC1A: transportador de glutamato; HKDC1 (*Hexokinase Domain Containing 1*); ALDOC: aldolase C; LDHD: Lactato desidrogenase D; FBP1: frutose bifosfatase 1; PGM1: fosfoglucomutase 1; G6PD: glicose-6-fosfato desidrogenase; PRPS2: fosforibosil pirofosfato sintetase 2; ALDH: Aldeído desidrogenase; ACSS: acil-CoA sintetase; ACO1: aconitase 1; IDH2: isocitrato desidrogenase 2; ME3: enzima málica 3; MDH1B: malato desidrogenase 1B; ACOT: acil-CoA tioesterases; ACSL: acil-CoA sintetase de cadeia longa; CPT1: carnitina palmitoiltransferases; ACACB: acetil CoA carboxilase beta; ACAA1: acetil CoA aciltransferase 1; ELOVL3: proteína de alongação dos ácidos graxos 3; ACOX2: acil CoA oxidase 2; FASN: sintase dos ácidos graxos; PPT1: palmitoil tioesterase 1; GLS1: Glutaminase1; GLS2: Glutaminase 2; GFPT2: glutamina-frutose-6-fosfato transaminase; UAP1: UDP-N-acetilglucosamina pirofosforilase 1.

3.3 CONCLUSÕES

- Análise realizada pelo programa Metacore de expressão gênica diferencial de tumores TN (em relação a não-TN), tanto de amostras de tecidos (disponíveis pelo banco público TCGA) quanto de linhagens celulares (sequenciadas por nós ou de dados obtidos de bancos públicos), nos mostrou que funções celulares ligados a processos de reestruturação de citoesqueleto, adesão celular, remodelamento de matriz e potencialmente migração/invasão, estão enriquecidas neste subtipo tumoral, coerente com já conhecida agressividade destes tumores. Há ainda ativação da via dos coactivadores *YAP/TAZ*, os quais regulam fatores de transcrição ligados a processos celulares tais como proliferação, sobrevivência, diferenciação, renovação de *stem-cells*, o que concorda com sua alta taxa de proliferação e incidência de *cancer stem-cells*. Já a análise de *metabolic networks* aponta para um cenário interessante de alta atividade ligada a sinalização/metabolismo de lipídeos que merece posterior investigação. Por fim, no cruzamento de informação de genes diferencialmente expressos e dados de metabolômica de linhagens celulares, mostrou que diferentes processos como o metabolismo de carboidratos, de lipídeos e aminoácidos, entre outros, estão alterados nas células tumorais TN em relação as não-TN.
- Uma análise manual realizada por nós com focos em vias metabólicas reconhecidamente alteradas em câncer revelou que tecidos tumorais TN potencialmente apresentam características que lhes garantem uma alta eficiência no uso de nutrientes, com alta captação e processamento de glicose, alto processamento de glutamina, re-aproveitamento de lactato e utilização de acetaldeídos como fonte energética e biossintética. No último caso, lhe garantindo, ainda, uma importante via de quimioresistência. Há ainda a indicação de ativação de vias de síntese/modificação de lipídeos que podem ter importante função sinalizadora nas células. Dos genes com expressão aumentada em tecidos TN (*HK3*, *PFKP*, *ENO1*, *LDHB*, *ACSL6*, *ELOVL3*, *ELOVL4*, *ALDH1A3*, *ALDH3A1*), alguns já tiveram sua relação com o subtipo

tumoral elucidada, enquanto que outros são de potencial interesse para estudos pois podem se revelar potenciais alvos de terapias.

- Da mesma maneira que os tecidos tumorais, as linhagens também apresentaram aumento na expressão de genes que podem indicar intensificação nas vias de metabolização de lipídeos e acetatos. Diferente do verificado em tecidos, o estudo das linhagens indicaram potencial aumento da via das hexosaminas e PPP, o que pode indicar que estas alterações são somente necessárias para células mantidas em cultura. Assim, por fim, abordar os processos de metabolização de lipídeos e acetato (e potencialmente as proteínas codificadas por *ALDH1A3*, *ALDH3A1*, *ACSL* e *SLC16A1*) podem revelar promissores alvos de tratamento dos tumores TN com alvo no metabolismo de nutrientes.

4 CAPÍTULO II: IDENTIFICAÇÃO E VALIDAÇÃO DE POTENCIAIS NOVOS ALVOS PARA O TRATAMENTO DE TUMORES DE MAMA TRIPLO NEGATIVO

4.1 METODOLOGIA

4.1.1 Definição de genes *up-regulated* em linhagens e tecidos tumorais de mama TN em relação a não-TN

Uma vez obtidas as listas de genes com expressão diferencial (como descrito no capítulo anterior, item 3.1.6) entre tecidos TN vs não-TN e linhagens celulares TN vs não-TN, realizamos uma intersecção e geramos uma lista em comum. Antes, entretanto, para a obtenção de uma listagem mais estridente (evitando erros do tipo I) dos dados de linhagens celulares, teste t utilizando a expressão gênica normalizada na escala log2 foram realizados para todos os genes expressos e o *p-value* obtido foi ajustado pelo método *false discovery ratio* (FDR)¹¹⁰. Somente genes cuja expressão diferencial obteve FDR menor que 0.05, tanto na análise do EBSeq quanto na análise do teste t, foram considerados. Com esta abordagem obtivemos 327 genes com expressão aumentada em comum entre pacientes e linhagens celulares (Tabela 3 do Anexo I).

4.1.2 Confirmação da expressão diferencial dos genes em dados de microarranjo de tecidos tumorais de mama

Os genes levantados usando as informações obtidas do cruzamento entre as listas obtidas dos dados do TCGA e de linhagens celulares foram então comparados com dados de outro grupo de pacientes do banco de microarranjos *ArrayExpress*¹¹¹. Quarenta e nove tumores de mama oriundos de biópsias, sendo 10 TN e 39 não-TN (E-GEOD-37126) foram obtidos, normalizados com o pacote *limma*¹¹² do R e empregados em duas correções das lâminas de microarranjo, uma eliminando o *background* das lâminas pelo método normal exponencial e a outra entre as lâminas pelo método dos quartis. As *probes* com replicatas na lâmina foram mergeadas em um único valor de intensidade e utilizada nas análises. Os valores de intensidade média das probes foram corrigidas para log2 *mean centering* baseados a partir da mediana dos resultados globais. O processamento dos dados foi realizado pelo aluno Douglas Adamoski. Definiu-se um corte de *p ajustado* ≤ 0.05 e um log2 *fold change*

de $\geq +1$, obtendo-se, assim, uma lista de 962 genes. Os 327 genes da Tabela 3 (Anexo I) foram cruzados com o desta lista, e, assim, 73 potenciais alvos foram identificados (Tabela 4, Anexo I).

4.1.3 Validação da expressão de genes alvos por qPCR em linhagens de tumor de mama

Da lista de potenciais alvos obtida (Tabela 4, Anexo I), três genes (*ARL4C*, *B3GNT5* e *GBP1*) foram avaliados quanto a expressão por qPCR empregando-se as condições descritas no item 3.1.7 e os *primers* descritos na Tabela 9. Estes genes foram escolhidos baseado na ausência de relação deles (a época de sua seleção) com câncer de mama triplo negativo. Para confirmar a expressão preferencial em linhagens TN dos 3 alvos selecionados, realizamos qPCR em amostras de cDNA das linhagens TN MDA-MB-231, MDA-MB-436, MDA-MB-157, MDA-MB-468, Hs578T e BT549 e da não-TN MCF7.

Tabela 9. Lista de *primers* utilizados para verificação da expressão gênica dos alvos em linhagens tumorais TN e não-TN.

GENE		PRIMERS	GENE		PRIMERS
rRNA 18S	F	ATTCCGATAACGAACGAG	ARL4C	F	TCTATGAGATGATCCTGAAACG
	R	AC		R	AGAGTAGGCCCGGTAAATCAG
		TCACAGACCTGTTATTGCT			
B3GNT5	F	AAATGAAGCTGGACTTGG	GBP1	F	ACTTCAGGAACAGGAGCAAC
	R	AG		R	TATGGTACATGCCTTTCGTC
		GCACGTCAAAATTCTTCCA			
		C			

4.1.4 Knock down de genes alvos em linhagens celulares

Os genes alvos selecionados foram silenciados pela expressão de sequências de shRNA. Para tanto, realizamos a clonagem das sequências em vetor lentiviral pLKO Puro (*Addgene*) modificado em nosso laboratório pelo aluno de doutorado Douglas Adamoski, de maneira a conter a proteína fluorescente mKO2 (*Monomeric Kusabira-Orange 2*) (Figura 18A-B). O vetor foi chamado de pLKO_mKO2_TRC. Esta abordagem permite, por microscopia de fluorescência, identificar células que

apresentaram a integração viral, e, assim, acompanhar o comportamento de crescimento destas células em específico sem haver mistura das informações entre células transduzidas e não-transduzidas.

Os oligos para a construção dos shRNAs foram desenhados a partir de sequências depositadas na plataforma TRC (*The RNAi Consortium*) e contém, no sentido 5' – 3', as extremidades coesivas para as enzimas Agel e EcoRI, região de reconhecimento do alvo, seis nucleotídeos para a formação de um *hairpin* e novamente a região de reconhecimento do alvo, complementar à anterior (Tabela 10). Sequência de shRNA para o gene da luciferase de *Firefly sp* foi utilizada como controle. Os oligos *forward* e *reverse* foram submetidos a uma reação de anelamento com utilização de 2,5 µM de cada oligo, tampão NEB 2 (*New English Biolabs*) 1X e água, em um volume final de 50 µL. A reação foi aquecida em termociclador a 95°C por 5 min e em seguida a 70°C por 10 min. Depois disso, a reação foi mantida em um decaimento de temperatura de 0.1°C/s até atingir a temperatura de 25°C.

Os oligos anelados foram subsequentemente ligados ao vetor modificado pLKO_mKO2_TRC previamente digerido com Agel e EcoRI (Figura 18C). Para a reação de ligação, foram utilizados 1 µL do oligo anelado (0,025µM), 1 µL de T4 DNA ligase (400000 U.mL⁻¹) (NEB), 1 µL de T4 polinucleotídeo quinase (10000 U.mL⁻¹) (NEB), 1 µL do tampão de T4 DNA ligase (NEB) 10X, 100 ng de pLKO digerido e água em um volume final de 10 µL. A reação foi incubada a 37°C por 3 horas em temperatura ambiente e depois transformada por choque térmico em bactérias *E. coli* DH5α quimicamente competentes previamente preparadas. A confirmação dos clones positivos foi realizada por PCR de colônia bacteriana.

As colônias positivas foram inoculadas em 2 mL de meio LB complementado de ampicilina, seguida pelo crescimento das bactérias sob agitação constante a 37°C por 16h. Após esse tempo, foi realizado o protocolo de minipreparação de DNA plasmidial caseiro adaptado de Sambrook¹¹³, e, posteriormente, realizou-se o sequenciamento para confirmar a presença e integridade do inserto com o *software SnapGene*. Após a confirmação dos clones positivos, os mesmos foram submetidos ao empacotamento viral para a formação das partículas lentivirais infectivas, procedimento realizado pelo laboratório LVV (Laboratório de Vetores Virais – LNBio).

Tabela 10. Sequências dos oligos para silenciamento para cada um dos alvos selecionados.

Nome da Sequência	Nome do gene	Sequência 5' - 3'
ARL4C_shRNA_1_F	Fator de ribosilação ADP tipo 4C	CCGGCCAAAGCGATGCTTCGAATTTCTCGAGAAATTCGAAGCATCGCTTTGGTTTTTG
ARL4C_shRNA_1_R		AATTCAAAAACCAAGCGATGCTTCGAATTTCTCGAGAAATTCGAAGCATCGCTTTGG
ARL4C_shRNA_2_F		CCGGGTCCCTGCATATCGTCATGTTCTCGAGAACATGACGATATGCAGGGACTTTTTG
ARL4C_shRNA_2_R		AATTCAAAAAGTCCCTGCATATCGTCATGTTCTCGAGAACATGACGATATGCAGGGAC
B3GNT5_shRNA_1_F	UDP-GlcNAc: betaGal beta-1,3-N-acetilglucosaminiltransferase 5	CCGGGCCTATGTAATCTCCGGTGATCTCGAGATCACCGGAGATTACATAGGCTTTTTG
B3GNT5_shRNA_1_R		AATTCAAAAAGCCTATGTAATCTCCGGTGATCTCGAGATCACCGGAGATTACATAGGC
B3GNT5_shRNA_2_F		CCGGATGATCGACGTTCCGGAATTACTCGAGTAATTCCGGAACGTCGATCATTTTTTG
B3GNT5_shRNA_2_R		AATTCAAAAATGATCGACGTTCCGGAATTACTCGAGTAATTCCGGAACGTCGATCAT
GBP1_shRNA_1_F	Proteína de ligação ao Guanilato 1	CCGGCGACGAAAGGCATGTACCATACTCGAGTATGGTACATGCCTTTCGTCGTTTTG
GBP1_shRNA_1_R		AATTCAAAAACGACGAAAGGCATGTACCATACTCGAGTATGGTACATGCCTTTCGTCG
GBP1_shRNA_2_F		CCGGTGAGACGACGAAAGGCATGTACTCGAGTACATGCCTTTCGTCGTCATTTTTG
GBP1_shRNA_2_R		AATTCAAAAATGAGACGACGAAAGGCATGTACTCGAGTACATGCCTTTCGTCGTCCTCA
LUC_shRNA_1_F	Luciferase	CCGGCTTACGCTGAGTACTTTCGACCTGACCCATCGAAGTACTCAGCGTAAGTTTTTG
LUC_shRNA_1_R		AATTCAAAACTTACGCTGAGTACTTTCGATGGGTCAGGTCGAAGTACTCAGCGTAAGCCGG

Para a obtenção do título viral, quantidades crescentes da preparação viral foram adicionadas com polibreno (1 µg/µL) em placas de 96 poços com células NIH3T3 cultivadas em DMEM baixa glicose suplementado com 10% soro fetal bovino, a 37°C e atmosfera 5% CO₂. Após 24 horas, o meio foi trocado, a placa fixada com solução de formaldeído 3,7% em PBS 1X e as células coradas com 1,5 µM do marcador nuclear DAPI (4',6-diamidino-2-fenilindole) em PBS 1X por 10 min. As imagens foram coletadas pelo microscópio leitor de placas Operetta (*Perkin Elmer*) e analisadas pelo *software* Columbus (*Perkin Elmer*). No protocolo de análise, o número total de células foi definido pela identificação de núcleos corados com DAPI, enquanto que o número de células efetivamente transduzidas (chamadas de positivas) foram definidas por intensidade de fluorescência de mKO2 (comprimento de onda de excitação 551nm e emissão 565nm) acima do limiar médio detectado em células não transduzidas (*background* de fluorescência).

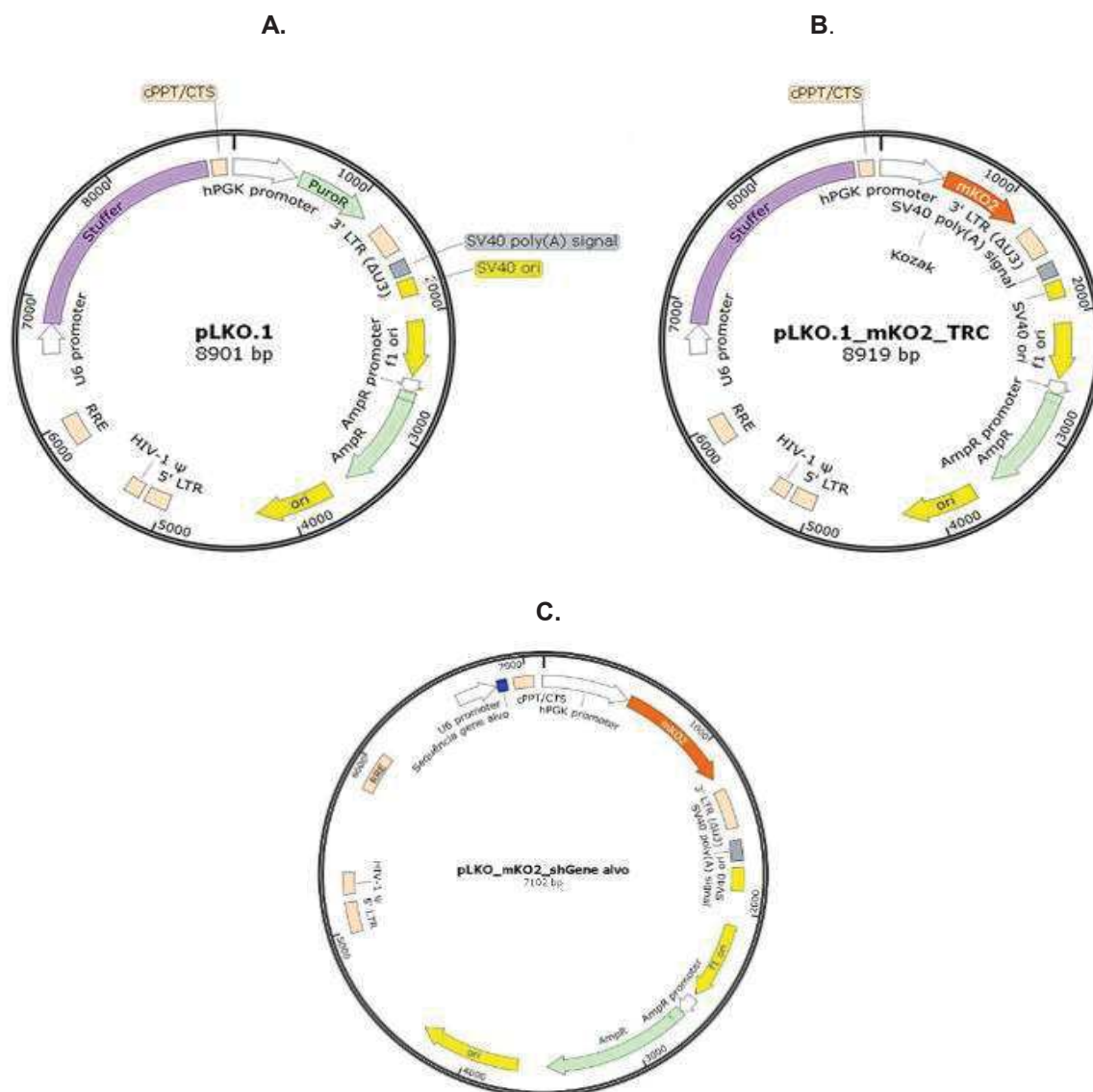


Figura 18. A. Vetor lentiviral original para silenciamento gênico pLKO.1, apresentando a expressão do shRNA através do promotor humano U6 e resistência ao agente de seleção puromicina com expressão através do promotor humano da fosfogliceroquinase. B. Vetor modificado para a remoção da resistência a puromicina e inserção de fluorescência pela expressão da proteína mKO2, permitindo identificação das células efetivamente transduzidas. C. Vetor modificado com a inserção da sequência de shRNA do gene alvo.

4.1.5 Avaliação da ação silenciamento gênico sobre proliferação celular por microscopia de fluorescência

As linhagens TN MDA-MB-231, MDA-MB-459, MDA-MB-578, MDA-MB-157, MDA-MB-436, MDA-MB-468, HCC30 e HCC1806 e as não-TN MCF-7, SKBR3 e T47D foram semeadas na densidade de 31,25 cells/mm² com polibreno (8 µg/mL) e vírus

em um MOI de 0,75, em triplicata. O meio foi trocado 24 horas após o plaqueamento e a cada 48 horas. Após 96 horas e 192 horas de transdução, as células foram fixadas com 3.7% formaldeído em 1X PBS por 20 min e coradas com 1.5 μ M de DAPI por 10 min. As imagens foram coletadas pelo microscópio leitor de placas Operetta (Perkin Elmer) e analisado pelo software Columbus (Perkin Elmer) com descrito no parágrafo acima. A proliferação celular (quando número de células no tempo 192hs > tempo 96 horas) e a morte celular (quando número de células no tempo 192hs < tempo 96 horas) foram calculados usando as seguintes equações: $100 * [(shGeneAlvo^{Tempo192} - shGeneAlvo^{Tempo96}) / (shLUC^{Tempo192} - shLUC^{Tempo96})]$ para proliferação e $100 * [(shGeneAlvo^{Tempo192} - shGeneAlvo^{Tempo96}) / (shGeneAlvo^{Tempo96})]$ para morte celular. A confirmação do *knock down* foi feita por qPCR segundo descrito no item 3.1.7.

Todo este procedimento, desde a clonagem até a realização dos experimentos de microscopia, foi realizado em colaboração com a aluna de doutorado Larissa Menezes.

4.2 RESULTADOS E DISCUSSÃO

4.2.1 Confirmação da expressão dos genes alvos selecionados por qPCR

Da lista obtida (Tabela 4, Anexo I) de genes *up-regulated* em TN de maneira consistente entre os *datasets* de tecidos tumorais do TCGA, linhagens celulares sequenciadas *in house* e de bancos públicos e de microarranjo de tecido tumorais (diferente grupo de pacientes), três genes o *ARL4C*, *B3GNT5* e *GBP1* (Figura 19) foram escolhidos baseado na ausência de relação deles (a época de sua seleção) com câncer de mama triplo negativo e por estarem com expressão aumentada tanto nos tecidos como nas linhagens TN em relação as Não-TN,

Para confirmar a expressão preferencial em linhagens TN dos 3 alvos selecionados, realizamos qPCR em amostras de cDNA das linhagens TN MDA-MB-231, MDA-MB-436, MDA-MB-468 e BT549 e das não-TN MCF7 e SKBR3. Os três genes selecionados apresentaram um nível de expressão consistentemente aumentado no grupo das TN (Figura 20).

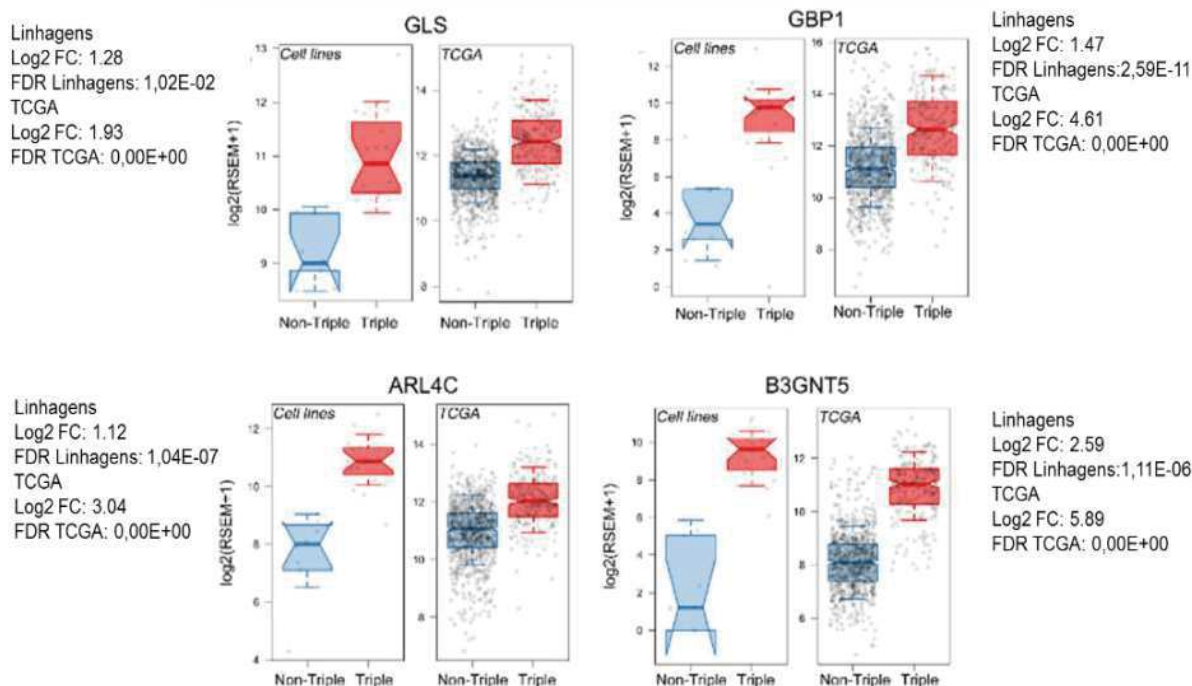


Figura 19. Genes *GBP1*, *ARL4C* e *B3GNT5* tem expressão aumentada em tecidos tumorais de mama TN (em relação aos não-TN) (dados obtidos do TCGA) e linhagens celulares TN (em relação aos não-TN) cujos RNAs foram sequenciados *in-house* ou obtidos de bancos públicos (*Cell lines*). *GLS* é um gene aumentado em tumores TN conforme já publicado por Gross et al 2014⁴¹.

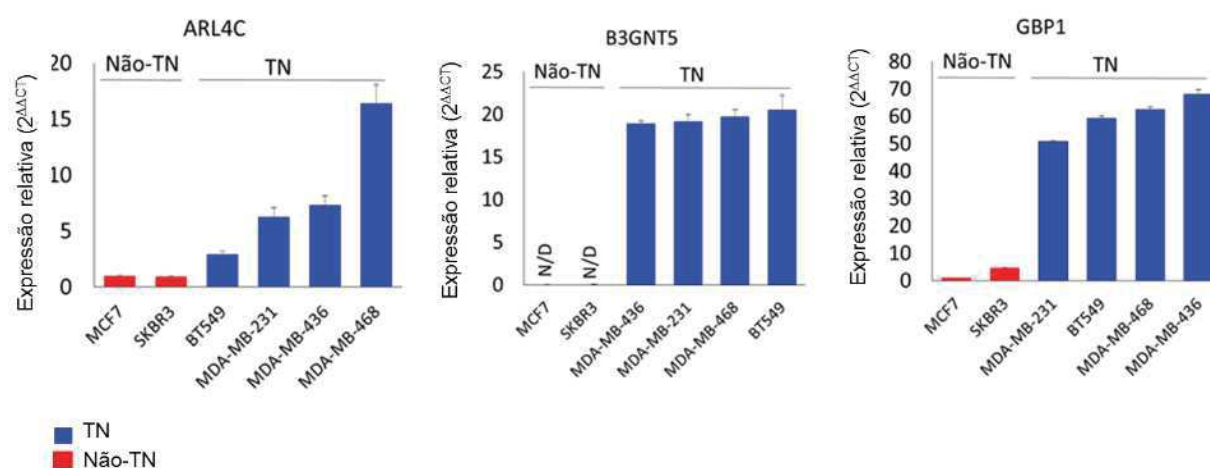


Figura 20. qPCR dos genes selecionados em linhagens de câncer de mama triplo-negativa (azul) e não-TN (vermelho). MCF-7 foi usada como calibrador. N/D: não detectado.

4.2.2 Análise da influência do *knock down* dos genes alvos sobre a proliferação celular

Inicialmente avaliamos a eficiência dos shRNAs em diminuir a expressão dos genes alvos (com 2 sequências de shRNA) na linhagem MDA-MB 231. Os resultados nos mostraram que houve uma redução de expressão de pelo menos 60% com o uso de todos os shRNAs, com exceção do shRNA.2 de *B3GNT5*, o qual não funcionou (Figura 21).

Posteriormente, avaliamos o efeito do *knock down* na proliferação celular de duas linhagens não-TN (MCF7 e SKBR3) e 4 linhagens TN (MDA-MB 231, MDA-MB 436, MDA-MB 468 e BT549). Os resultados mostraram que, dos 3 genes, a atenuação de *GBP1* com as duas sequências de shRNA afetou de maneira mais consistente e significativa as linhagens TN do que as não-TN, revelando-se, portanto, um alvo mais promissor (Figura 22). Vale ressaltar que o resultado foi inclusive mais proeminente, em termos de especificidade TN x não-TN e intensidade do efeito, do que o obtido com o *knock down* de GLS, um gene já definido como alvo em TN e para o qual já existe composto em estudos clínico Fase I ⁴¹.

Dado os resultados positivos obtidos com *GBP1*, expandimos as análises para outras linhagens TN e não-TN (Figura 23). Os resultados mostraram que, de fato, o *knock down* de *GBP1* afeta a proliferação de células de câncer de mama triplo negativo de maneira diferencial às não triplo-negativas, levando à morte nas linhagens TN (HCC38, HCC1806, MDA-MB468, MDA-MB436), a uma inibição na proliferação de 60% nas linhagens TN (BT549, HS578t, MDA-MB231, MDA-MB157) e uma redução de 20% nas linhagens não-TN (MCF7, SKBR3, T47D).

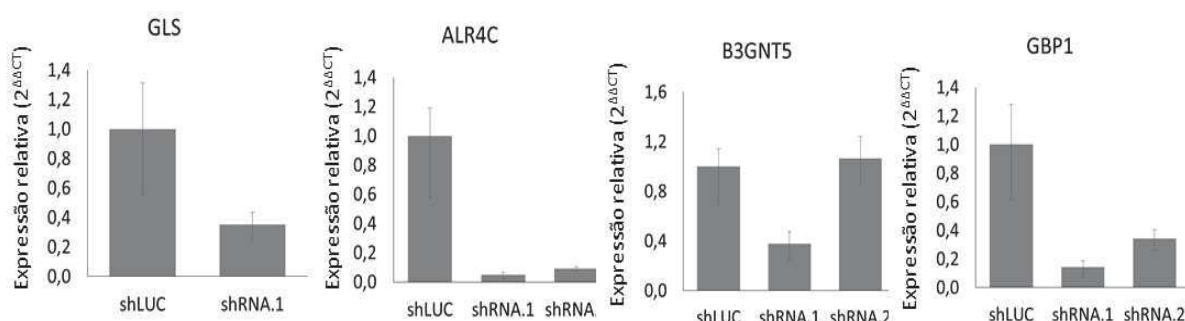


Figura 21. qPCR para verificar a atenuação da expressão dos genes ARL4C, B3GNT5 e GBP1 em MDA-MB 231. O *knock down* de GLS foi usado como controle. Dois conjuntos de shRNAs para cada um dos genes foram avaliados.

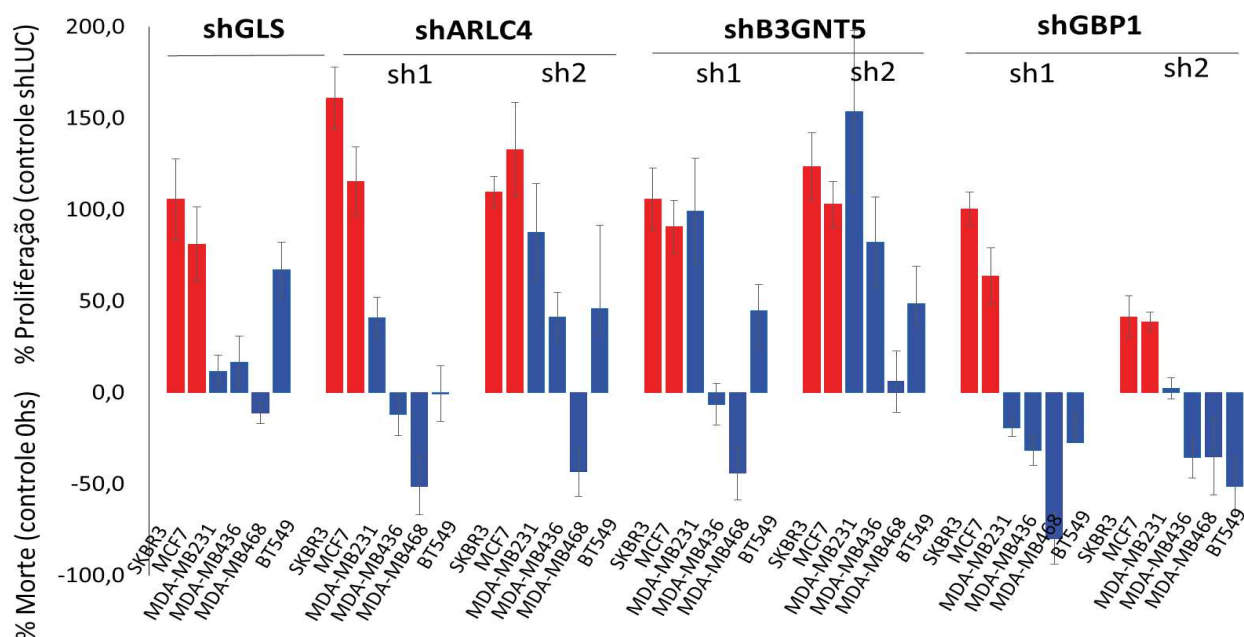


Figura 22. Porcentagem de proliferação (ou morte) de linhagens não-TN (barras em vermelho) e TN (barras em azul) após *knock down* dos alvos *ARL4C*, *B3GNT5* e *GBP1* com 2 sequências de shRNA distintas (sh1 e sh2). Sequência de shRNA contra o alvo não específico Luciferase foi usado como controle negativo. *GLS* é um alvo já elucidado em TN, estando um inibidor do mesmo em fase clínica I ⁴¹.

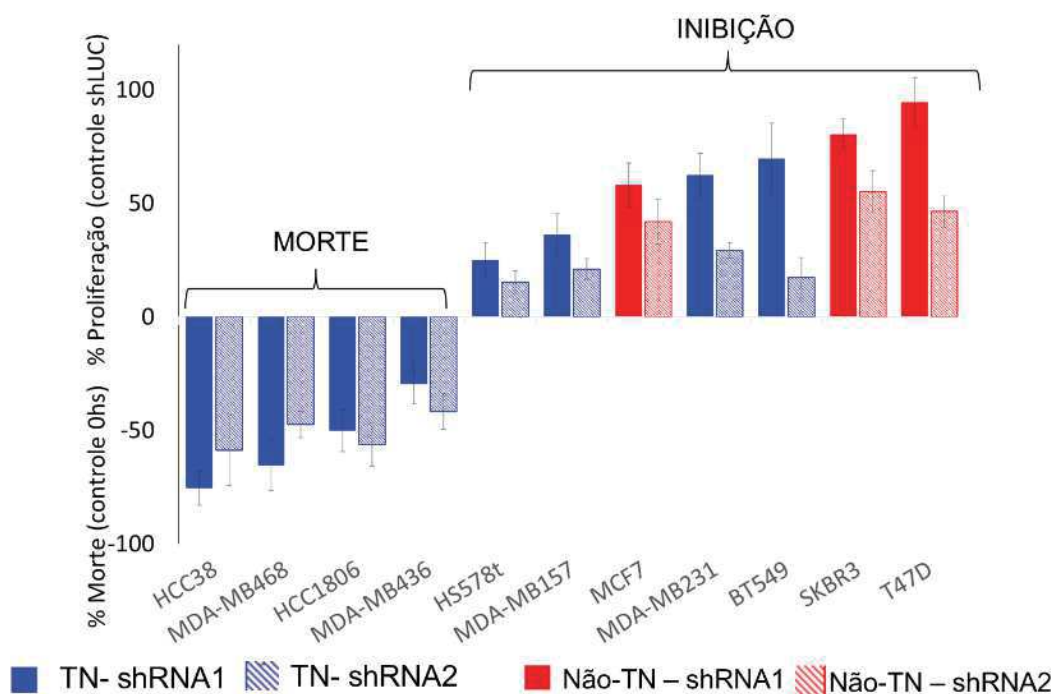


Figura 23. Porcentagem de proliferação (ou morte) de linhagens não-TN (barras em vermelho) e TN (barras em azul) após *knock down* do alvo *GBP1* com 2 sequências de shRNA distintas: sh1 (coluna sólida) e sh2 (coluna pálida). Sequência de shRNA contra o alvo não específico Luciferase foi usado como controle negativo.

A proteína *GBP1* (*Guanylate binding protein 1*), pertence à família das GTPases e sua expressão pode ser induzida pelos interferons *IL-1 β* , *TNF*, e *IFN- γ* *in vitro* ¹¹⁴, sendo detectado *in vivo* durante uma resposta inflamatória¹¹⁵. *GBP1* é expresso em vasos sanguíneos e células de origem monocítica ^{116,115}. A proteína se liga a nucleotídeos de guanina (GMP, GDP e GTP) e se distingue das *GTP-binding proteins* pela presença de 2 motivos de ligação (ao contrário de 3).

GBP1 tem ação anti-angiogênica já que inibe a proliferação e migração de células endoteliais através da indução da expressão da integrina *ITGA4*^{117,114,118} e da inibição de *MMP1* (metaloproteinase de matriz 1), uma collagenase necessária para a migração de células através da matriz extracelular¹¹⁴. *GBP1* foi identificado como supressor de tumor em células de câncer colo retal¹¹⁹ mas sua expressão está relacionada a resistência a droga paclitaxel e doxorubicina em diferentes linhagens celulares ¹²⁰. Um trabalho de 2013 mostra que *GBP1* apresenta um novo SNP em tumores TN, mas não há extensos relatos na literatura ligando este gene com este subtipo tumoral¹²¹. Em uma análise realizada na base de dados Oncomine, confirmamos em outros grupos de pacientes que tumores de mama do tipo carcinoma medular (que são frequentemente TN)¹²¹ (Figura 24A) e os TN (Figura 24B) apresentam níveis aumentados de expressão deste gene em relação a tecidos saudáveis ou outros subtipos tumorais de mama. Tal observação foi confirmada pela avaliação da expressão deste gene em tecidos normais de mama versus os tumorais, também disponíveis no banco de dados do TCGA (Figura 25).

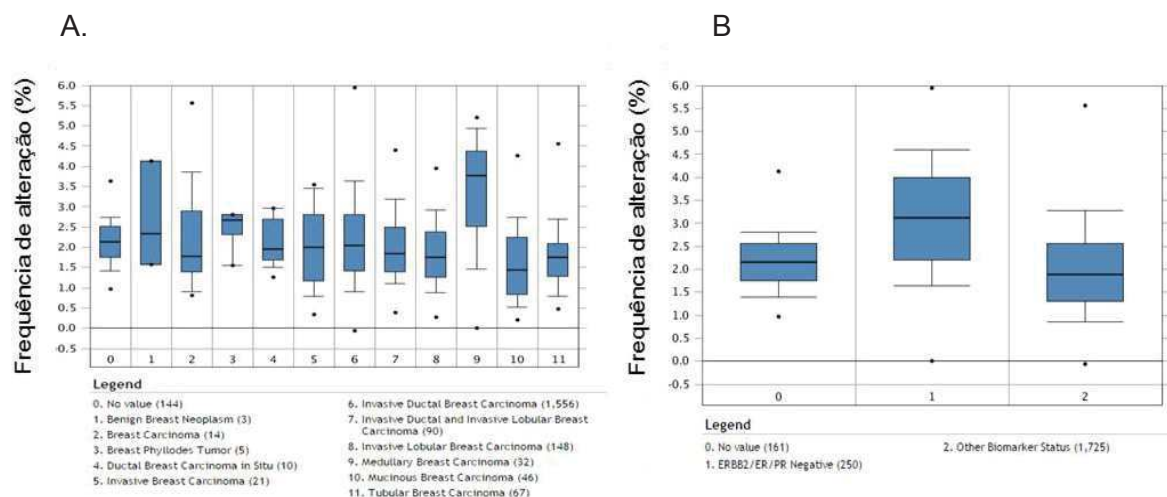


Figura 24. Expressão gênica de *GBP1* em câncer de mama conforme analisado pela base de dados Oncomine. A. Comparação entre diferentes subtipos tumorais de mama *versus* tecido normal. B. Expressão em função do *status* dos marcadores ER, PR e HER2 (*ERBB2*).

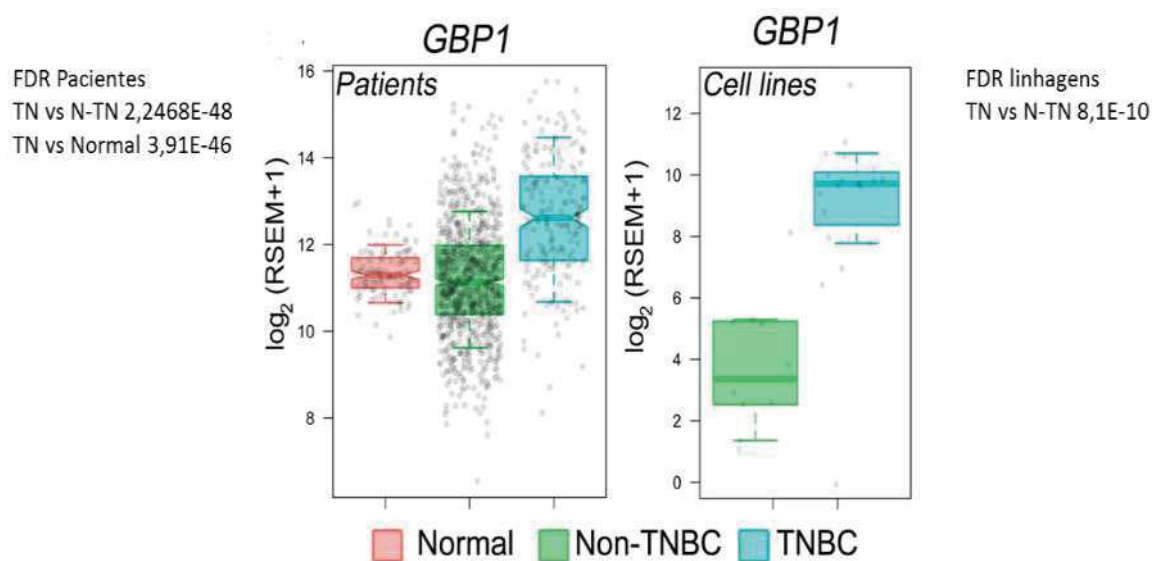


Figura 25. Expressão gênica de *GBP1* em tecidos de mama saudáveis e tecidos de câncer de mama separados entre não-TN e TN.

Por fim, realizamos uma análise da “*drugability*” deste alvo utilizando a base de dados canSAR (<https://cansar.icr.ac.uk>)¹²². Uma proteína é *druggable* se sua atividade pode ser modulada pela interação com pequenas moléculas alterando a função da

proteína com benefício terapêutico. Esta base utiliza três diferentes tipos de parâmetros *structure*, *ligand* and *network-based druggability*.

Structure Druggability examina a estrutura da proteína (quando disponível) e identifica cavidades na superfície disponíveis para interação com fármacos. GBP1 apresenta 9 diferentes estruturas cristalográficas (PDBs ID 1F5N_A, 1DG3_A, 2B8W_A, 2B8W_B, 2B92_A, 2B92_B, 2BC9_A, 2D4H_A e 2D4H_B), tendo sido identificado pelo menos 1 sítio de potencial interação. Em uma análise paralela utilizando o programa *pockdrug* 3.0, verificamos a presença de pelo menos 5 potenciais cavidades de ligação a pequenas moléculas (Figura 26).

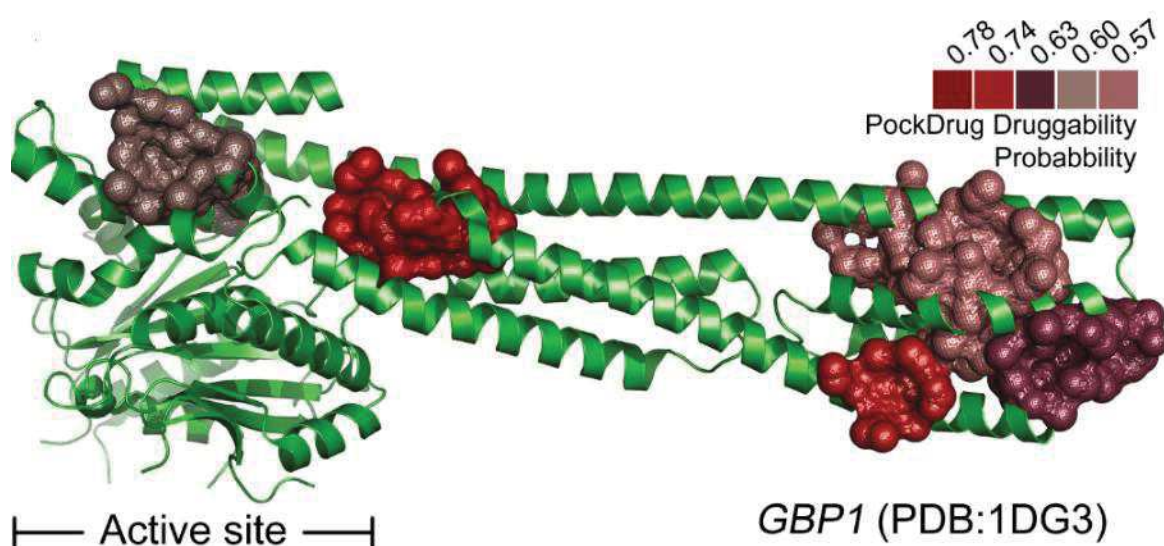


Figura 26. Cavidades na superfície da estrutura de GBP1 disponíveis para a interação com pequenas moléculas.

Ligand-based druggability compara a proteína alvo com todas as outras presentes no proteoma disponível na plataforma. Estima a probabilidade do alvo ser *druggable* com base nas propriedades químicas e de bioatividade de pequenas moléculas que foram testados contra a própria proteína e/ou seus homólogos. A lógica é que se o alvo se liga a compostos semelhantes a fármacos, é mais provável que seja *druggable* do que um alvo que não o faça. Neste sentido, *GBP1* obteve um índice de 90% (Figura 27).

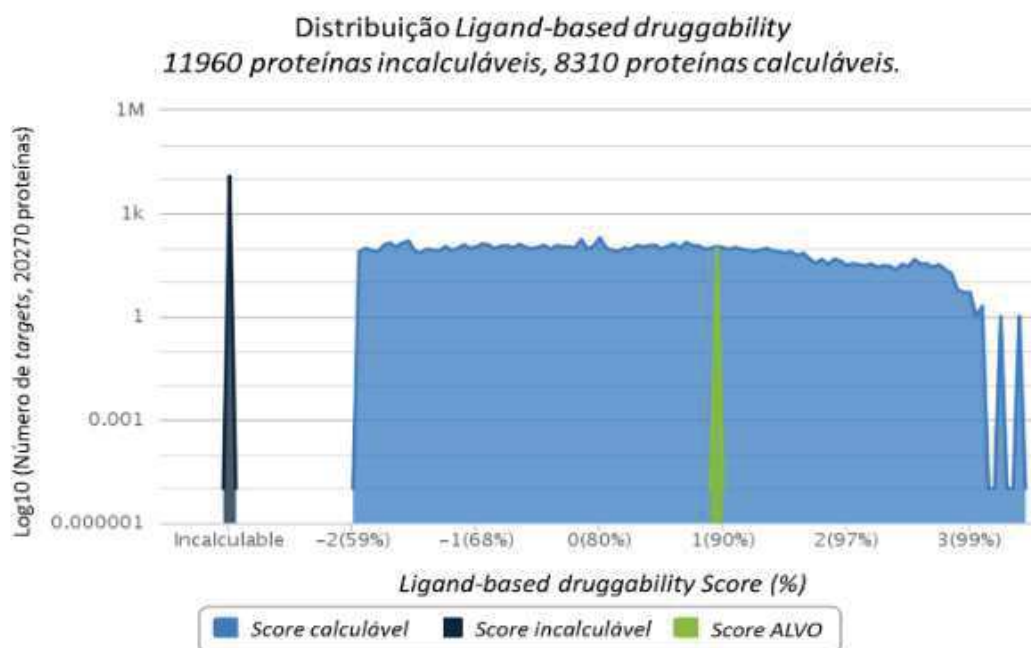
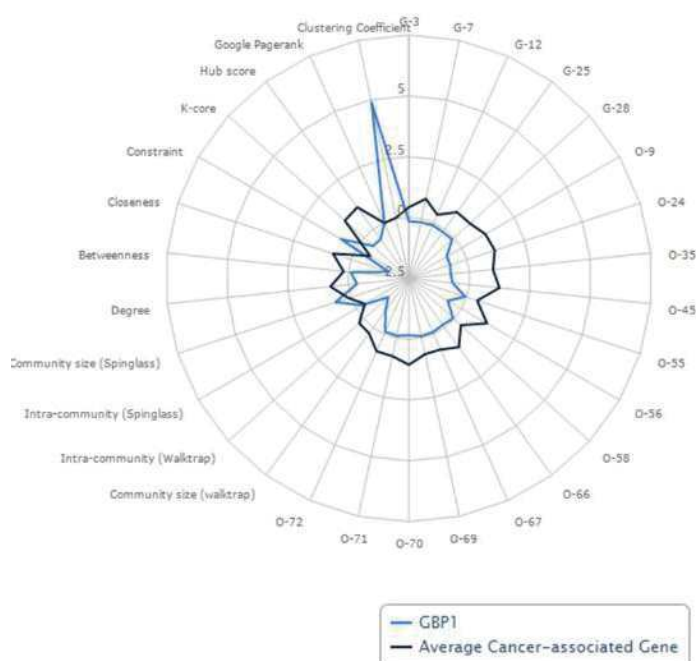


Figura 27. Distribuição *Ligand-based druggability* do alvo GBP1.

Por fim, a análise ***Network-based druggability*** examina as relações de interação proteína-proteína ao redor do alvo proposto¹²³. Neste sentido, avalia-se que uma proteína envolvida em muitos eventos de sinalização e de interação proteína-proteína seria um bom alvo, dado que a inibição de sua função tem o potencial de interromper diversos processos dentro da célula. No sentido contrário, uma proteína com poucas interações não será um bom alvo pois sua inibição afetaria poucos processos na célula. Para *GBP1*, o *Network-based druggability* foi de 84,80% para diversas doenças e 51,20% para câncer (Figura 28).



Druggability Percentile: 84.80

Figura 28. Network-based Druggability

Em conjunto, os dados apresentados acima indicam que esta proteína pode ser um potencial novo alvo para o tratamento deste subtipo tumoral. Os resultados apresentados neste capítulo estão sendo sumarizados em um artigo que está em elaboração.

4.3 CONCLUSÕES

- Geramos uma lista de 73 genes com expressão aumentada em tumores de mama TN (em relação a não-TN) em tecidos tumorais (obtidos por 2 técnicas diferentes e de 2 grupos de pacientes) e linhagens celulares sequenciadas *in-house* ou obtidas de bancos públicos. Desta lista, 3 genes (*ARL4C*, *B3GNT5* e *GBP1*) foram avaliados quanto a expressão em linhagens TN e não-TN, mostrando estar consistentemente mais expressos em linhagens TN. Dos três genes, um gene, relativo a proteína *GBP1*, uma *guanylate-binding protein*, mostrou diminuir seletivamente a proliferação celular de linhagens TN quando silenciado. *GBP1* é uma proteína induzida por processo inflamatório, normalmente relacionada a processos de defesa do organismo, e relacionada a inibição de angiogênese. Em tumores de próstata é descrita como supressor tumoral. Sua função para o estabelecimento/progressão/sobrevivência do subtipo TN precisa ser avaliada de maneira a confirmá-la como potencial alvo para o tratamento desta doença. Em especial, é preciso realizar o *knock down* da mesma em linhagens não tumorigênicas para verificar a toxicidade de sua inibição para células não cancerosas.

5 REFERENCIAS

1. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100**, 57–70 (2000).
2. Hanahan, D. & Weinberg, R. a. Hallmarks of cancer: the next generation. *Cell* **144**, 646–74 (2011).
3. ji lu, Nicole, L., Solimin, S. & Elledge, J. Principles of Cancer Therapy: Oncogene and Non-oncogene Addiction. *Cell* **136**, p823–837 (2009).
4. Inca. *Estimativa Incidência de câncer no Brasil - 2014*. (2014). doi:978-85-7318-194-4
5. Group, C., Factors, H. & Cancer, B. Menarche , menopause , and breast cancer risk : individual participant meta-analysis , including 118 964 women with breast cancer from 117 epidemiological studies Collaborative Group on Hormonal Factors in Breast Cancer Summary Background. 1–11 (2015).
6. Meier-Abt, F. *et al.* Parity induces differentiation and reduces Wnt/Notch signaling ratio and proliferation potential of basal stem/progenitor cells isolated from mouse mammary epithelium. *Breast Cancer Res.* **15**, R36 (2013).
7. Greiser, C. M., Greiser, E. M. & Dören, M. Menopausal hormone therapy and risk of breast cancer: A meta-analysis of epidemiological studies and randomized controlled trials. *Hum. Reprod. Update* **11**, 561–573 (2005).
8. Carey, A. & Carey, L. Understanding and Treating Triple-Negative Breast Cancer. *Oncology* **22**, 1233–1243 (2008).
9. Stein, R. a *et al.* Estrogen-related receptor alpha is critical for the growth of estrogen receptor-negative breast cancer. *Cancer Res.* **68**, 8805–12 (2008).
10. Kariagina, A., Aupperlee, M. D. & Haslam, S. Z. Progesterone receptor isoform functions in normal breast development and breast cancer. *Crit. Rev. Eukaryot. Gene Expr.* **18**, 11–33 (2008).
11. Turner, N. *et al.* Integrative molecular profiling of triple negative breast cancers identifies amplicon drivers and potential therapeutic targets. *Oncogene* **29**, 2013–23 (2010).
12. Cowley, S. M., Hoare, S., Mosselman, S. & Parker, M. G. Estrogen receptors alpha and beta form heterodimers on DNA. *J. Biol. Chem.* **272**, 19858–19862 (1997).
13. Danielsen, A. J. & Maihle, N. J. The EGF/ErbB receptor family and apoptosis. *Growth Factors* **20**, 1–15 (2002).
14. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).

15. Sorlie, T. *et al.* Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U. S. A.* **98**, 10869–10874 (2001).
16. Anders, C. K. & Carey, L. a. Biology, metastatic patterns, and treatment of patients with triple-negative breast cancer. *Clin. Breast Cancer* **9 Suppl 2**, S73–81 (2009).
17. Hudis, C. a & Gianni, L. Triple-negative breast cancer: an unmet medical need. *Oncologist* **16 Suppl 1**, 1–11 (2011).
18. Schneider, B. P. *et al.* Triple-negative breast cancer: Risk factors to potential targets. *Clin. Cancer Res.* **14**, 8010–8018 (2008).
19. Fong, P. C. AZD2281 (KU-0059436), a PARP (poly ADP-ribose polymerase) inhibitor with single agent anticancer activity in patients with BRCA deficient ovarian cancer: results from a phase I study. *J. Clin. Oncol.* **26**, 5510 (2008).
20. Lehmann, B. D. *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
21. Tchou, J. *et al.* Degree of tumor FDG uptake correlates with proliferation index in triple negative breast cancer. *Mol. Imaging Biol.* **12**, 657–662 (2010).
22. Herschkowitz, J., Simin, K., Weigman, V., Mikaelian, et al. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol.* **8**, R76 (2007).
23. Peddi, P. F., Ellis, M. J. & Ma, C. Molecular basis of triple negative breast cancer and implications for therapy. *Int. J. Breast Cancer* **2012**, 217185 (2012).
24. Griffiths, C. L. & Olin, J. L. Triple negative breast cancer: a brief review of its characteristics and treatment options. *J. Pharm. Pract.* **25**, 319–23 (2012).
25. DeBerardinis, Ralph J Sayed, Nabil, Ditsworth, Dara, Thompson, C. B. Brick by brick: metabolism and tumor cell growth. *Curr. Opin. Genet. Dev.* **18**, 54–61 (2008).
26. Vander-Heiden, M. Metabolic changes in cancer. 52 (2012). at <http://webinar.sciencemag.org/webinar/archive/metabolic-changes-cancer>
27. Warburg, O. On the Origin of Cancer Cells. *Science* **123**, 309–314 (1956).
28. Gatenby, R. a & Gillies, R. J. Why do cancers have high aerobic glycolysis? *Nat. Rev. Cancer* **4**, 891–9 (2004).
29. Wise, D. R. *et al.* Myc regulates a transcriptional program that stimulates mitochondrial glutaminolysis and leads to glutamine addiction. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 18782–18787 (2008).
30. Le, A. *et al.* Glucose-independent glutamine metabolism via TCA cycling for proliferation and survival in B cells. *Cell Metab.* **15**, 110–21 (2012).

31. Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029–33 (2009).
32. Vera Miranda Gonçalves, F. B. and R. M. R. in *'Molecular Considerations and Evolving Surgical Management Issues in the Treatment of Patients with a Brain Tumor'* (ed. Lichtor, T.) (2015). doi:10.5772/59606
33. Semenza, G. L. Tumor metabolism: Cancer cells give and take lactate. *J. Clin. Invest.* **118**, 3835–3837 (2008).
34. Vizán, P., Mazurek, S. & Cascante, M. Robust metabolic adaptation underlying tumor progression. *Metabolomics* **4**, 1–12 (2008).
35. DeBerardinis, R. J. *et al.* Beyond aerobic glycolysis: transformed cells can engage in glutamine metabolism that exceeds the requirement for protein and nucleotide synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 19345–19350 (2007).
36. Collier, H. a. Is cancer a metabolic disease? *Am. J. Pathol.* **184**, 4–17 (2014).
37. DeBerardinis, Ralph J. sayed, Nabil. Ditsworth, D. and C. T. brick by brick and tumor cell growth. *Curr. Opin. Genet. Dev.* **18**, 54–61 (2008).
38. Curi, R., Newsholme, P. & Newsholme, E. A. Metabolism of pyruvate by isolated rat mesenteric lymphocytes , lymphocyte mitochondria and isolated mouse macrophages. **250**, 383–388 (1988).
39. DeBerardinis, R. J. & Cheng, T. Q's next: the diverse functions of glutamine in metabolism, cell biology and cancer. *Oncogene* **29**, 313–24 (2010).
40. Robinson, M. M. *et al.* Novel mechanism of inhibition of rat kidney-type glutaminase by bis-2-(5-phenylacetamido-1,2,4-thiadiazol-2-yl)ethyl sulfide (BPTES). *Biochem. J.* **406**, 407–414 (2007).
41. Gross, M. I. *et al.* Antitumor activity of the glutaminase inhibitor CB-839 in triple-negative breast cancer. *Mol. Cancer Ther.* **13**, 890–901 (2014).
42. Szeliga, M. *et al.* Silencing of GLS and overexpression of GLS2 genes cooperate in decreasing the proliferation and viability of glioblastoma cells. *Tumor Biol.* **35**, 1855–1862 (2014).
43. Seltzer, M. J. *et al.* Inhibition of glutaminase preferentially slows growth of glioma cells with mutant IDH1. *Cancer Res.* **70**, 8981–8987 (2010).
44. Lee, S. Y. *et al.* Dlx-2 and glutaminase upregulate epithelial-mesenchymal transition and glycolytic switch. *Oncotarget* **7**, 7925–39 (2016).
45. Gordan, J. D., Thompson, C. B. & Simon, M. C. HIF and c-Myc: sibling rivals for control of cancer cell metabolism and proliferation. *Cancer Cell* **12**, 108–13 (2007).
46. Gao, P. *et al.* c-Myc suppression of miR-23a/b enhances mitochondrial glutaminase expression and glutamine metabolism. *Nature* **458**, 762–765 (2009).

47. Owen, O. E., Kalhan, S. C. & Hanson, R. W. The key role of anaplerosis and cataplerosis for citric acid cycle function. *J. Biol. Chem.* **277**, 30409–30412 (2002).
48. McClelland, M. L. *et al.* An Integrated Genomic Screen Identifies LDHB as an Essential Gene for Triple-Negative Breast Cancer. *Cancer Res.* **72**, 5812–5823 (2012).
49. Farabegoli, F. *et al.* Galloflavin, a new lactate dehydrogenase inhibitor, induces the death of human breast cancer cells with different glycolytic attitude by affecting distinct signaling pathways. *Eur. J. Pharm. Sci.* **47**, 729–738 (2012).
50. Timmerman, L. A. *et al.* Glutamine sensitivity analysis identifies the xCT antiporter as a common triple-negative breast tumor therapeutic target. *Cancer Cell* **24**, 450–65 (2013).
51. Roessner, U. & Bowne, J. What is metabolomics all about? *Biotechniques* **46**, 363–365 (2009).
52. Çakir, T., Hendriks, M. M. W. B., Westerhuis, J. a. & Smilde, A. K. Metabolic network discovery through reverse engineering of metabolome data. *Metabolomics* **5**, 318–329 (2009).
53. Morozova, O., Hirst, M. & Marra, M. a. Applications of new sequencing technologies for transcriptome analysis. *Annu. Rev. Genomics Hum. Genet.* **10**, 135–151 (2009).
54. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq : a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
55. Claudino, W. M. *et al.* Metabolomics: available results, current research projects in breast cancer, and future applications. *J. Clin. Oncol.* **25**, 2840–6 (2007).
56. Canevarolo, R. R. Metabolômica da resistência ao metotrexato na leucemia linfóide aguda. (2012).
57. Matsuoka, Y. & Shimizu, K. in *chemical biology* (ed. Ekinici, D.) (2012). doi:10.5772/35121
58. Fan, T. W. M. *et al.* Altered regulation of metabolic pathways in human lung cancer discerned by (13)C stable isotope-resolved metabolomics (SIRM). *Mol. Cancer* **8**, 41 (2009).
59. Yang, C., Richardson, A. D., Osterman, A. & Smith, J. W. Profiling of central metabolism in human cancer cells by two-dimensional NMR, GC-MS analysis, and isotopomer modeling. *Metabolomics* **4**, 13–29 (2007).
60. Metallo, C. M. *et al.* Evaluation of 13 C isotopic tracers for metabolic flux analysis in mammalian cells. *Biotechniques* **144**, 167–174 (2009).
61. Meadows, A. L. *et al.* Metabolic and morphological differences between rapidly proliferating cancerous and normal breast epithelial cells. *Biotechnol. Prog.* **24**, 334–341 (2008).

62. Zamboni, N. ¹³C metabolic flux analysis in complex systems. *Curr. Opin. Biotechnol.* **22**, 103–8 (2011).
63. Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally. *Cancer Cell* **10**, 515–527 (2009).
64. Elstrodt, F. *et al.* BRCA1 mutation analysis of 41 human breast cancer cell lines reveals three new deleterious mutants. *Cancer Res.* **66**, 41–45 (2006).
65. Wishart, D. S. *et al.* HMDB: A knowledgebase for the human metabolome. *Nucleic Acids Res.* **37**, 603–610 (2009).
66. Martin, J. a & Wang, Z. Next-generation transcriptome assembly. *Nat. Rev. Genet.* **12**, 671–82 (2011).
67. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
68. Varley, K. E. *et al.* Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res. Treat.* **146**, 287–297 (2014).
69. Daemen, A. *et al.* Modeling precision treatment of breast cancer. *Genome Biol.* **14**, R110 (2013).
70. Kodama, Y., Shumway, M. & Leinonen, R. The sequence read archive: Explosive growth of sequencing data. *Nucleic Acids Res.* **40**, 2011–2013 (2012).
71. Leinonen, R., Sugawara, H., Shumway, M. & International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res.* **39**, D19–21 (2011).
72. Andrews, S. FastQC: A quality control tool for high throughput sequence data. (2014).
73. Jiang, H., Lei, R., Ding, S.-W. & Zhu, S. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
74. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
75. Bullard, J. H., Purdom, E., Hansen, K. D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94 (2010).
76. Fraley, C., Raftery, A. E., Murphy, T. B. & Scrucca, L. *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation. Technical Report 597* (2012).
77. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–52 (2012).
78. Leng, N. *et al.* EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* **29**, 1035–1043 (2013).

79. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
80. Kolde, R. pheatmap: Pretty Heatmaps. R package. (2012).
81. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
82. Vu, V. Q. ggbiplot: A ggplot2 based biplot. (2011).
83. Ewing, B. *et al.* Base-Calling of Automated Sequencer Traces Using. *Genome Res.* 175–185 (2005). doi:10.1101/gr.8.3.175
84. Shi, P., Feng, J. & Chen, C. Hippo pathway in mammary gland development and breast cancer. *Acta Biochim. Biophys. Sin. (Shanghai)*. **47**, 53–9 (2015).
85. Prat, A. & Perou, C. M. Deconstructing the molecular portraits of breast cancer. *Mol. Oncol.* **5**, 5–23 (2011).
86. Ridgway, N. D. The role of phosphatidylcholine and choline metabolites to cell proliferation and survival. *Crit. Rev. Biochem. Mol. Biol.* **48**, 20–38 (2013).
87. Strub, G. M., Maceyka, M., Hait, N. C., Milstien, S. & Spiegel, S. Extracellular and intracellular actions of sphingosine-1-phosphate. *Adv. Exp. Med. Biol.* **688**, 141–155 (2010).
88. Reis-Filho, J. S. & Tutt, a N. J. Triple-negative tumours: a critical review. *Histopathology* **52**, 108–118 (2008).
89. Wilson, J. E. Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. *J. Exp. Biol.* **206**, 2049–2057 (2003).
90. Radojkovic, J. & Ureta, T. Hexokinase isoenzymes from the Novikoff hepatoma. *Biochemistry* **242**, 895–903 (1987).
91. Wyatt, E. *et al.* Regulation and cytoprotective role of hexokinase III. *PLoS One* **5**, (2010).
92. Gao, H., Luo, X.-G., Chen, X. & Wang, J. Identification of key genes affecting disease free survival time of pediatric acute lymphoblastic leukemia based on bioinformatic analysis. *Blood Cells, Mol. Dis.* **54**, 38–43 (2015).
93. Sánchez-Martínez, C. & J. Aragón, J. Analysis of phosphofructokinase subunits and isozymes in ascites tumor cells and its original tissue, murine mammary gland. *FEBS Lett.* **409**, 86–90 (1997).
94. Moon, J.-S. *et al.* Krüppel-like factor 4 (KLF4) activates the transcription of the gene for the platelet isoform of phosphofructokinase (PFKP) in breast cancer. *J. Biol. Chem.* **286**, 23808–16 (2011).
95. Louie, S. M. *et al.* GSTP1 Is a Driver of Triple-Negative Breast Cancer Cell Metabolism and Pathogenicity. *Cell Chem. Biol.* **23**, 567–578 (2016).

96. Capello, M., Ferri-Borgogno, S., Cappello, P. & Novelli, F. α -Enolase: a promising therapeutic and diagnostic tumor target. *FEBS J.* **278**, 1064–74 (2011).
97. Sonveaux, P. *et al.* Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. *J. Clin. Invest.* **118**, 3930–3942 (2008).
98. Santos, C. R. & Schulze, A. Lipid metabolism in cancer. *FEBS J.* **279**, 2610–2623 (2012).
99. Soupene, E., Dinh, N. P., Siliakus, M. & Kuypers, F. A. Activity of the acyl-CoA synthetase ACSL6 isoforms: role of the fatty acid Gate-domains. *BMC Biochem.* **11**, 18 (2010).
100. Marcato, P. *et al.* Aldehyde dehydrogenase 1A3 influences breast cancer progression via differential retinoic acid signaling. *Mol. Oncol.* **9**, 17–31 (2015).
101. Opdenaker, L. M. *et al.* Immunohistochemical analysis of aldehyde dehydrogenase isoforms and their association with estrogen-receptor status and disease progression in breast cancer. *Breast Cancer Targets Ther.* **6**, 205–209 (2014).
102. Edenberg, H. J. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res. Health* **30**, 5–13 (2007).
103. Sládek, N. E., Kollander, R., Sreerama, L. & Kiang, D. T. Cellular levels of aldehyde dehydrogenases (ALDH1A1 and ALDH3A1) as predictors of therapeutic responses to cyclophosphamide-based chemotherapy of breast cancer: A retrospective study. *Cancer Chemother. Pharmacol.* **49**, 309–321 (2002).
104. Li, G.-H. & Huang, J.-F. Inferring therapeutic targets from heterogeneous data: HKDC1 is a novel potential therapeutic target for cancer. *Bioinformatics* **30**, 748–52 (2014).
105. Wu, X. *et al.* Long Chain Fatty Acyl-CoA Synthetase 4 Is a Biomarker for and Mediator of Hormone Resistance in Human Breast Cancer. *PLoS One* **8**, 1–20 (2013).
106. Ying, H. *et al.* Oncogenic Kras Maintains Pancreatic Tumors through Regulation of Anabolic Glucose Metabolism. *Cell* **149**, 656–670 (2012).
107. Yun, J. *et al.* Glucose Deprivation Contributes to the Development of KRAS Pathway Mutations in Tumor Cells. *Science (80-.)*. **325**, 1555–1559 (2009).
108. Slawson, C., Copeland, R. J. & Hart, G. W. O-GlcNAc signaling: A metabolic link between diabetes and cancer? *Trends Biochem. Sci.* **35**, 547–555 (2010).
109. Ferrer, C. M. *et al.* O-GlcNAcylation regulates cancer metabolism and survival stress signaling via regulation of HIF-1 pathway. *Mol. cel* **54**, 820–831 (2014).
110. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statistical Soc. Ser. B* **57**, 289–300 (1995).

111. Carraro, D. M. *et al.* Comprehensive Analysis of BRCA1, BRCA2 and TP53 Germline Mutation and Tumor Characterization: A Portrait of Early-Onset Breast Cancer in Brazil. *PLoS One* **8**, 7–9 (2013).
112. Smyth, G. K. & Speed, T. Normalization of cDNA microarray data. *Methods* **31**, 265–273 (2003).
113. Sambrook, J. R. D. *Molecular Cloning. Molecular Cloning: a laboratory manual* (cold spring harbor laboratory press, 2001).
114. Guenzi, E. *et al.* The guanylate binding protein-1 GTPase controls the invasive and angiogenic capability of endothelial cells through inhibition of MMP-1 expression. *EMBO J.* **22**, 3772–3782 (2003).
115. Lubeseder-Martellato, C. *et al.* Guanylate-binding protein-1 expression is selectively induced by inflammatory cytokines and is an activation marker of endothelial cells during inflammatory diseases. *Am. J. Pathol.* **161**, 1749–1759 (2002).
116. Naschberger, E. *et al.* Nuclear factor-kappaB motif and interferon-alpha-stimulated response element co-operate in the activation of guanylate-binding protein-1 expression by inflammatory cytokines in endothelial cells. *Biochem. J.* **379**, 409–420 (2004).
117. Guenzi, E. *et al.* The helical domain of GBP-1 mediates the inhibition of endothelial cell proliferation by inflammatory cytokines. *EMBO J.* **20**, 5568–5577 (2001).
118. Weinländer, K. *et al.* Guanylate binding protein-1 inhibits spreading and migration of endothelial cells through induction of integrin alpha4 expression. *FASEB J.* **22**, 4168–4178 (2008).
119. Britzen-Laurent, N. *et al.* GBP-1 acts as a tumor suppressor in colorectal cancer cells. *Carcinogenesis* **34**, 153–62 (2013).
120. Duan, Z., Foster, R., Brakora, K. a, Yusuf, R. Z. & Seiden, M. V. GBP1 overexpression is associated with a paclitaxel resistance phenotype. *Cancer Chemother. Pharmacol.* **57**, 25–33 (2006).
121. Horvath, A. *et al.* Novel insights into breast cancer genetic variance through RNA sequencing. *Sci. Rep.* **3**, 2256 (2013).
122. Tym, J. E. *et al.* CanSAR: Updated cancer research and drug discovery knowledgebase. *Nucleic Acids Res.* **44**, D938–D943 (2016).
123. Halling-Brown, M. D., Bulusu, K. C., Patel, M., Tym, J. E. & Al-Lazikani, B. canSAR: An integrated cancer public translational research and drug discovery resource. *Nucleic Acids Res.* **40**, 947–956 (2012).

6 ANEXO I

Gráficos

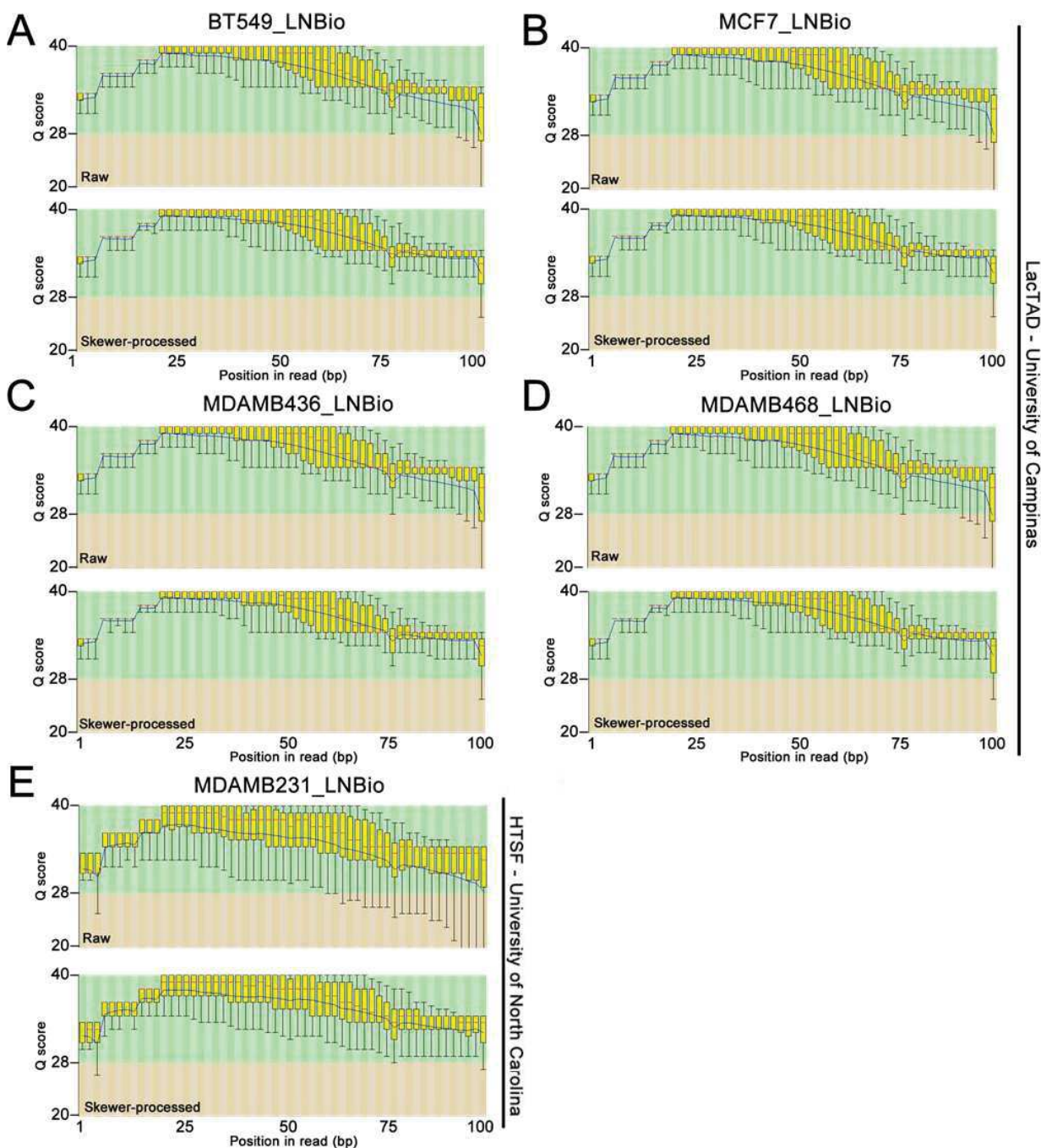


Figura 1. Avaliação da qualidade dos *reads* antes (RAW) e após (*skewer-processed*) o *trimming* para as linhagens (A) BT549, (B) MCF7, (C) MDA-MB436 e (D) MDAM-B468 sequenciadas no LACTAD-UNICAMP, (E) MDA-MB231, sequenciada no HTSF. *Box-plot* dos valores de PHRED (Q score) para cada posição de base para todas as sequências lidas tanto antes (acima) quanto depois (abaixo) do *trimming*. Sem o *trimming* (RAW) evidencia-se uma redução da qualidade no final das posições dos *reads* por pb, típica do processo de sequenciamento. Dados com *trimming* apresentam melhora significativa nos valores dos scores.

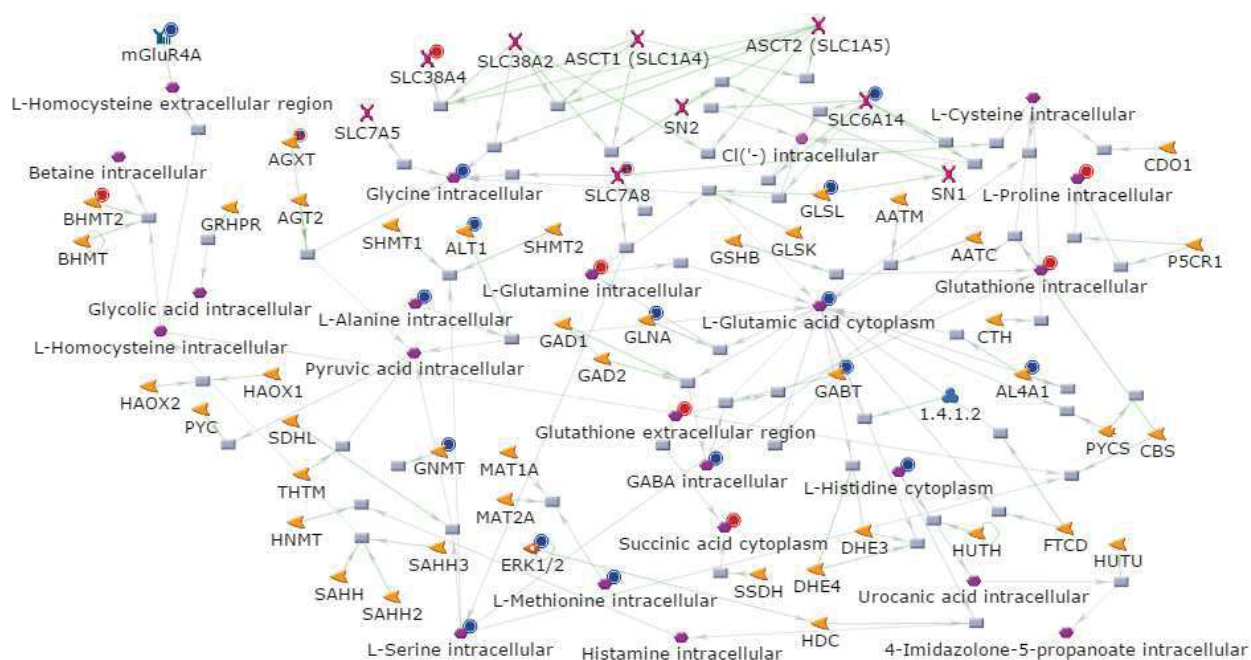


Figura 2. Rede metabólica relacionada com o metabolismo de aminoácidos, conforme análise do Metacore. Genes com expressão diferencial com $\log_2 \text{fold change} \pm 1$, p valor $< 0,05$. Metabólitos com média da concentração normalizada por massa proteica. Círculo vermelho indica indução e círculo azul diminuição de expressão em TN. ◀ Enzima, ▶ proteína fosfatase, ● Metabólito, ⚡ generic binding protein, ⚡ proteína G alfa., ★ fator de transcrição. ⬇ regulador de Proteína-G. ✕ Transportador.

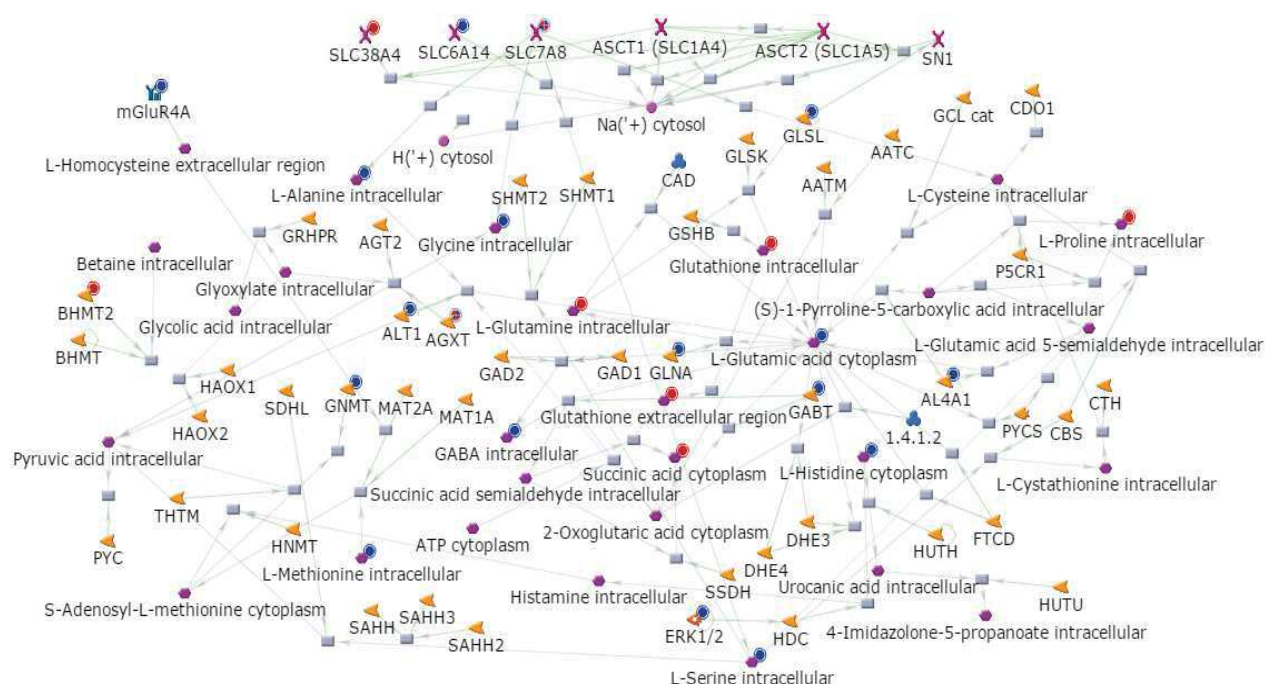


Figura 3. Rede metabólica relacionada com o Metabolismo de Lipídeos_ Glicosfingolipídios, conforme análise do Metacore. Genes com expressão diferencial com $\log_2 \text{fold change} \pm 1$, p valor $<0,05$. Metabólitos com média da concentração normalizada por massa proteica. Círculo vermelho indica indução e círculo azul diminuição de expressão em TN. ◀ Enzima, ▶ proteína fosfatase, ● Metabólito, ⚡ *generic binding protein*, ⚡ proteína G alfa., ★ fator de transcrição. ↓ regulador de Proteína-G. ✕ Transportador.

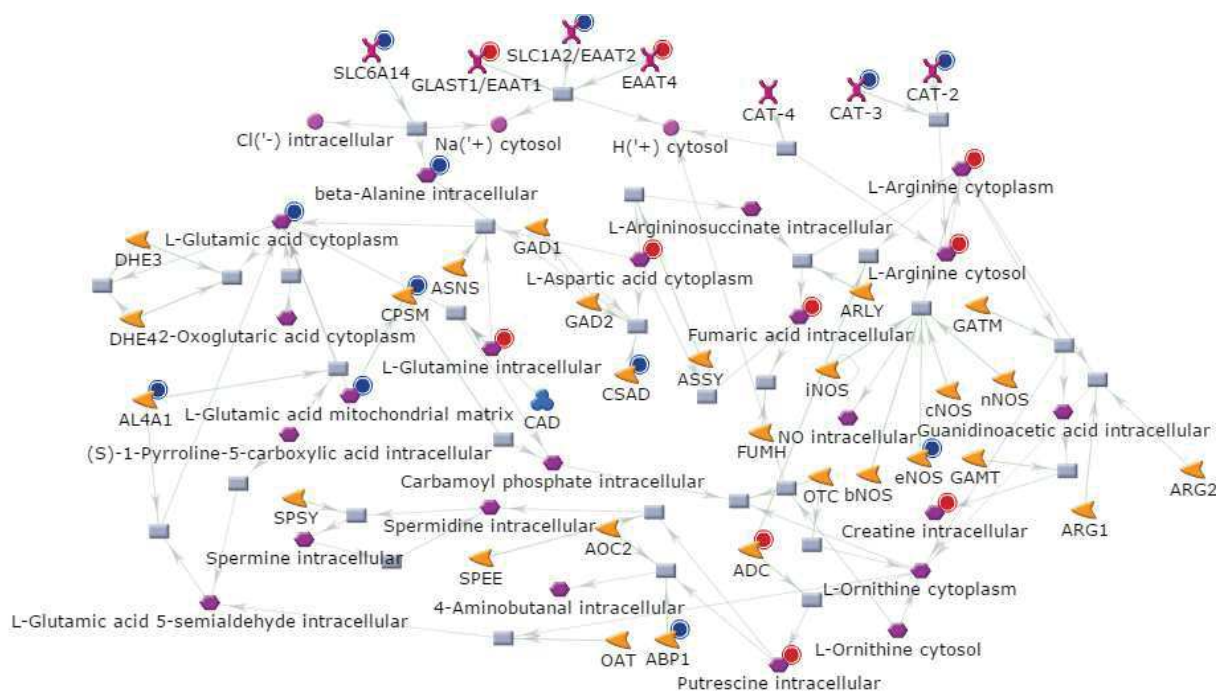


Figura 4. Rede metabólica relacionada com o metabolismo da Arginina conforme análise do Metacore. Genes com expressão diferencial com $\log_2 \text{fold change} \pm 1$, p valor $< 0,05$. Metabólitos com média da concentração normalizada por massa proteica. Círculo vermelho indica indução e círculo azul diminuição de expressão em TN. ◀ Enzima, ▶ proteína fosfatase, ● Metabólito, Z generic binding protein, ★ proteína G alfa., ★ fator de transcrição. ↓ regulador de Proteína-G. X Transportador.

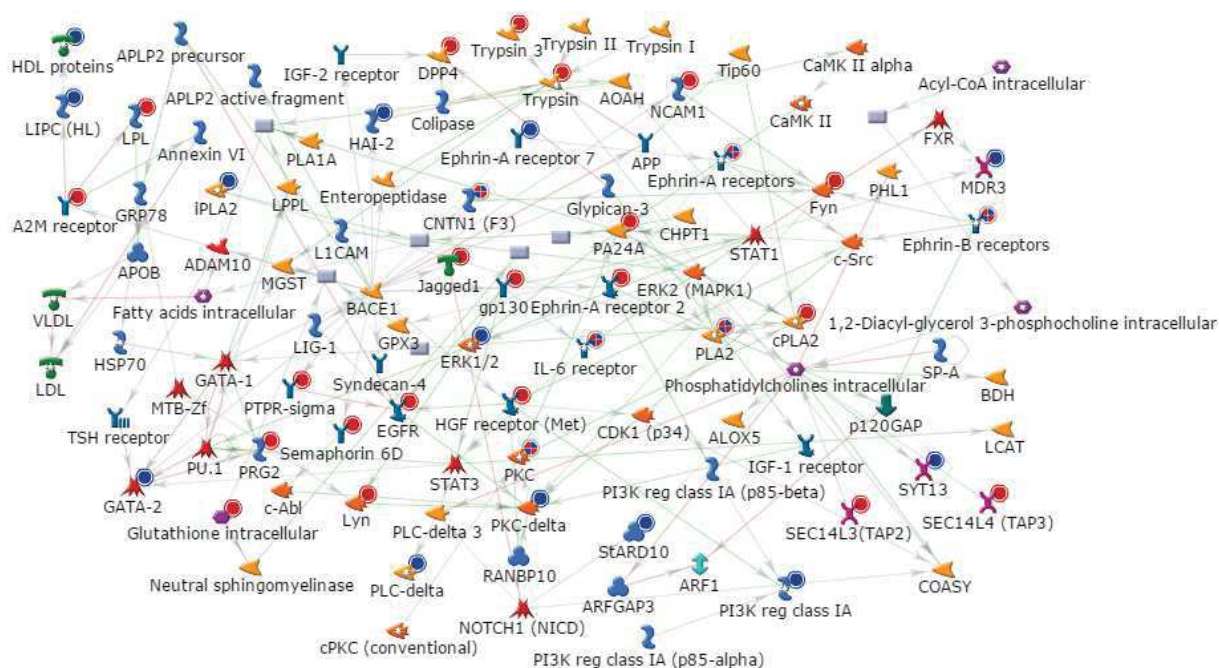


Figura 5. Rede metabólica relacionada com a Via da fosfatidilcolina conforme análise do Metacore. Genes com expressão diferencial com $\log_2 \text{fold change} \pm 1$, p valor $< 0,05$. Metabólitos com média da concentração normalizada por massa proteica. Círculo vermelho indica indução e círculo azul diminuição de expressão em TN. ◀ Enzima, ▶ proteína fosfatase, ● Metabólito, ◌ *generic binding protein*, ◌ proteína G alfa., ★ fator de transcrição. ⬇ regulador de Proteína-G. ✕ Transportador.

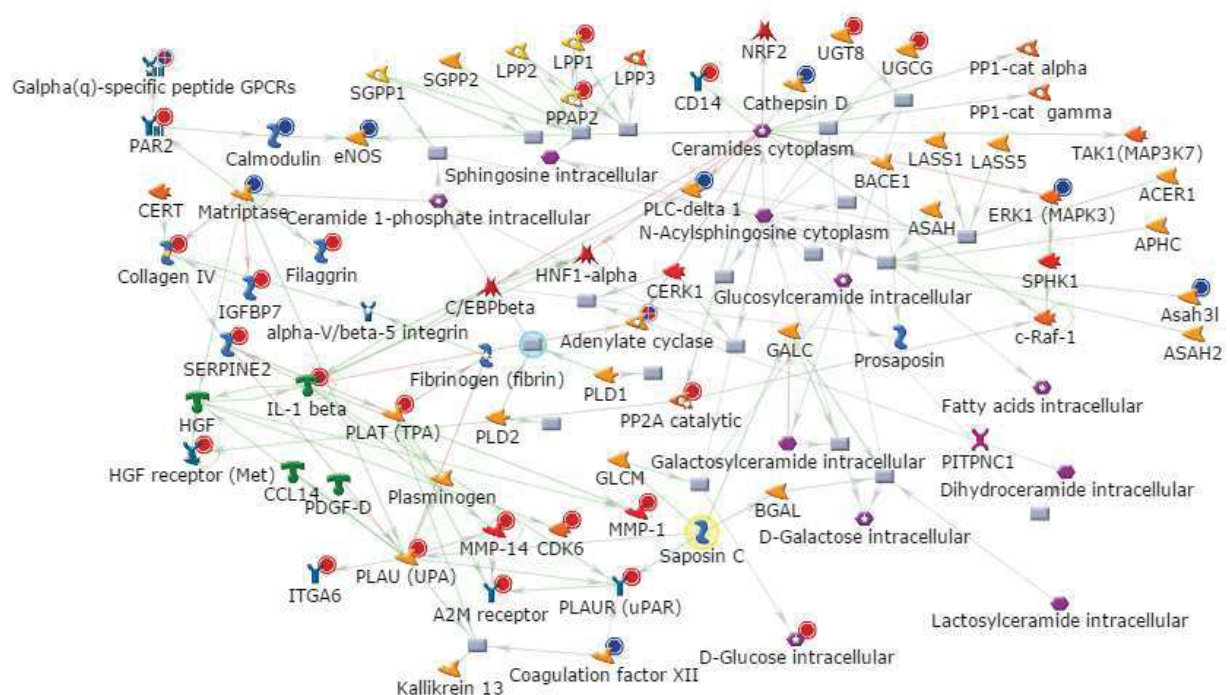


Figura 6. Rede metabólica relacionada com a Via da n-acil esfingosina fosfato conforme análise do Metacore. Genes com expressão diferencial com $\log_2 \text{fold change} \pm 1$, p valor $< 0,05$. Metabólitos com média da concentração normalizada por massa proteica. Círculo vermelho indica indução e círculo azul diminuição de expressão em TN. ◀ Enzima, ▶ proteína fosfatase, ● Metabólito, ● generic binding protein, ● proteína G alfa., ★ fator de transcrição, ● regulador de Proteína-G, ✕ Transportador.

Tabela 1. Concentração de metabólitos intracelular ($\mu\text{Mol/mg}$ de proteína). Média dos valores da triplicata de cada amostra.

	Triplo Negativo (TN)							Não Triplo Negativo (N-TN)		
	(ER -; PR -; HER2 -)							(ER +; PR +; HER2 -)	(ER +; PR -; HER2 +)	
Metabólitos	MDA-MB157	MDA-MB231	MDA-MB436	MDA-MB468	HCC1806	BT459	HS578t	SKBR3	MCF7	ZR751
4_Aminobutirato	0,000E+00	1,333E-02	0,000E+00	0,000E+00	0,000E+00	1,709E-02	1,795E-02	7,007E-03	1,336E-02	0,000E+00
Acetato	6,575E-02	2,238E-02	6,853E-03	1,004E-02	8,145E-02	4,895E-02	5,884E-02	4,841E-03	8,930E-02	3,778E-03
Alanina	4,576E-03	1,430E-02	1,328E-01	1,129E-01	2,658E-02	4,432E-02	1,255E-01	2,846E-02	1,028E-01	4,166E-03
Arginina	0,000E+00	0,000E+00	0,000E+00	2,981E-02	0,000E+00	2,536E-02	0,000E+00	ND	1,080E-01	0,000E+00
Asparagina	0,000E+00	3,493E-02	7,886E-02	5,644E-02	3,548E-02	4,090E-02	1,433E-02	3,495E-02	5,503E-02	1,066E-02
Aspartato	3,045E-02	2,753E-02	1,507E-01	1,202E-02	2,623E-02	2,536E-02	4,571E-02	2,295E-02	1,305E-01	6,976E-03
Colina	7,908E-04	1,070E-03	1,476E-03	1,888E-03	1,768E-03	3,528E-03	6,186E-03	2,614E-03	3,340E-03	1,356E-03
Creatina	7,513E-03	3,892E-03	4,407E-02	0,000E+00	3,324E-02	2,789E-02	0,000E+00	3,485E-02	4,545E-02	1,056E-02
Creatina fosfato	1,051E-02	7,783E-03	3,237E-02	0,000E+00	6,017E-02	5,083E-02	0,000E+00	1,075E-01	8,668E-02	1,540E-02
Dimetilamina	0,000E+00	0,000E+00	0,000E+00	0,000E+00	1,297E-03	0,000E+00	2,112E-03	0,000E+00	3,775E-03	0,000E+00
Etanol	5,084E-02	1,751E-02	4,691E-02	5,753E-02	5,558E-02	6,681E-02	7,876E-02	5,683E-02	4,327E-02	1,444E-02
Formato	8,191E-03	2,238E-02	3,479E-02	2,186E-02	1,798E-02	1,852E-02	2,339E-02	3,224E-02	2,207E-02	0,000E+00
Fumarato	0,000E+00	3,697E-03	2,319E-03	1,391E-03	2,534E-03	2,536E-03	3,319E-03	3,389E-03	5,082E-03	1,356E-03
Galactitol	0,000E+00	0,000E+00	0,000E+00	8,575E-02	0,000E+00	0,000E+00	0,000E+00	0,000E+00	ND	0,000E+00
Galactose	0,000E+00	0,000E+00	4,966E-02	4,859E-02	0,000E+00	0,000E+00	0,000E+00	0,000E+00	ND	0,000E+00
Glucitol	4,383E-02	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	ND	0,000E+00
Glicose	2,265E-02	5,643E-02	0,000E+00	0,000E+00	0,000E+00	0,000E+00	3,440E-02	6,845E-02	2,614E-02	2,722E-02
Glutamato	9,654E-02	2,656E-01	2,360E-01	1,236E-01	2,227E-01	3,857E-01	1,109E+00	9,091E-02	4,148E-01	2,626E-02
Glutamina	2,565E-02	9,077E-02	5,282E-02	7,045E-02	9,135E-02	1,968E-01	5,267E-01	1,175E-01	3,258E-01	3,633E-02
Glutationa	1,401E-02	7,589E-02	4,133E-02	4,869E-02	9,807E-02	8,015E-02	3,049E-01	8,278E-02	2,014E-01	1,095E-02
Glicerol	0,000E+00	2,092E-02	5,893E-02	2,733E-02	0,000E+00	5,435E-02	5,311E-02	7,097E-02	3,194E-02	1,841E-02
Glicine	1,604E-02	8,562E-03	1,443E-01	1,541E-01	2,363E-02	1,164E-01	3,330E-01	1,207E-01	1,172E-01	2,771E-02
Histidina	0,000E+00	0,000E+00	3,384E-02	2,663E-02	0,000E+00	0,000E+00	2,957E-02	0,000E+00	2,672E-02	0,000E+00

Isoleucina	0,000E+00	1,498E-02	2,172E-02	2,057E-02	1,473E-02	2,161E-02	8,841E-02	1,017E-02	4,400E-02	4,263E-03
Lactato	1,520E-02	1,672E-01	2,161E-02	5,296E-02	1,307E-01	1,012E-01	3,158E-01	1,010E-01	3,546E-01	8,981E-02
Leucina	0,000E+00	1,936E-02	2,119E-02	2,345E-02	1,756E-02	2,205E-02	1,135E-01	1,171E-02	4,516E-02	3,875E-03
Malato	0,000E+00	4,719E-02	2,941E-02	2,037E-02	2,469E-02	3,649E-02	5,356E-02	4,192E-02	7,928E-02	7,751E-03
Metionina	0,000E+00	6,616E-03	1,023E-02	1,192E-02	1,108E-02	5,623E-03	3,742E-02	4,647E-03	1,612E-02	1,259E-03
N-Dimetilglicina	0,000E+00	2,335E-03	9,488E-04	0,000E+00	0,000E+00	3,087E-03	3,470E-03	1,840E-03	6,098E-03	1,066E-03
N_Acetilaspato	0,000E+00	0,000E+00	4,122E-02	0,000E+00	0,000E+00	4,035E-02	0,000E+00	0,000E+00	7,086E-02	0,000E+00
N-Acetilglutamato	0,000E+00	1,119E-02	0,000E+00	0,000E+00	0,000E+00	0,000E+00	1,448E-02	0,000E+00	2,759E-03	0,000E+00
NAD+	0,000E+00	1,605E-02	1,307E-02	8,545E-03	1,049E-02	1,885E-02	2,851E-02	1,259E-02	3,383E-02	0,000E+00
O-Fosfocolina	1,384E-02	2,852E-01	6,062E-02	6,518E-02	5,021E-02	1,899E-01	1,423E-01	1,641E-01	3,872E-01	1,139E-01
O-Fosfoetanolamina	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	0,000E+00	2,293E-02	8,045E-02	1,725E-01	0,000E+00
Pantotenato	0,000E+00	1,022E-02	0,000E+00	8,347E-03	0,000E+00	7,387E-03	2,504E-02	2,120E-02	1,321E-02	0,000E+00
Fenilalanina	0,000E+00	0,000E+00	0,000E+00	2,216E-02	0,000E+00	0,000E+00	6,608E-02	5,906E-03	3,122E-02	0,000E+00
Prolina	1,440E-02	3,240E-02	0,000E+00	1,481E-02	2,693E-02	3,749E-02	6,427E-02	2,527E-02	4,690E-02	1,279E-02
Propionato	0,000E+00	0,000E+00	0,000E+00	0,000E+00	3,124E-03	3,638E-03	0,000E+00	2,033E-03	4,792E-03	0,000E+00
Putrescina	0,000E+00	2,656E-02	0,000E+00	1,510E-02	0,000E+00	2,789E-02	0,000E+00	7,745E-03	3,964E-02	9,494E-03
Serina	0,000E+00	0,000E+00	0,000E+00	7,572E-02	0,000E+00	3,330E-02	1,101E-01	3,563E-02	0,000E+00	0,000E+00
Succinato	2,316E-03	3,113E-03	3,057E-03	3,577E-03	4,538E-03	2,756E-03	7,091E-03	1,743E-03	1,045E-02	1,841E-03
Taurina	2,960E-02	1,686E-01	3,627E-02	7,055E-03	5,628E-02	7,751E-02	2,093E-01	2,788E-02	4,977E-01	7,751E-03
Treonina	4,688E-03	1,635E-02	5,703E-02	8,764E-02	1,945E-02	2,613E-02	2,115E-01	1,472E-02	1,243E-01	4,747E-03
Tirosina	0,000E+00	1,080E-02	4,881E-02	4,382E-02	1,179E-02	1,191E-02	1,038E-01	0,000E+00	4,545E-02	0,000E+00
Valina	0,000E+00	9,535E-03	1,877E-02	2,087E-02	1,067E-02	9,702E-03	1,056E-01	6,099E-03	4,980E-02	2,616E-03
Mio-Inositol	2,361E-02	7,112E-02	4,712E-02	1,053E-02	3,342E-02	4,575E-02	2,793E-01	3,205E-02	2,643E-02	7,751E-03
Sn-Glicerolo-3-fosfocolina	8,417E-03	1,878E-02	1,265E-01	1,550E-02	2,959E-02	7,354E-02	3,455E-02	5,131E-03	1,616E-01	1,792E-02
Beta alanina	0,000E+00	1,537E-02	0,000E+00	0,000E+00	3,477E-03	1,114E-02	3,183E-02	3,873E-03	9,583E-03	0,000E+00

Tabela 2. Genes associados às rotas metabólicas diferencialmente expressos nas amostras de pacientes provenientes do TCGA TN em relação aos N-TN. Genes diferenciais com *fold change* $\geq +1$ para genes induzidos (*up*) e ≤ -1 para genes reduzidos (*Down*), dados com *p-valor* < 0.05 em negrito.

Gene	Tecidos (TCGA)			Linhagens celulares		
	p-valor	Log2FC	Status	p-valor	Log2FC	Status
Glicólise						
HK1	1,25E-06	-0,32	NO DIFERENCIAL	8,89E-01	-0,43	NO DIFERENCIAL
HK2	8,66E-01	-0,17	NO DIFERENCIAL	9,64E-01	0,16	NO DIFERENCIAL
HK3	0,00E+00	1,17	UP	0,00E+00	7,07	UP
GPI	0,00E+00	0,73	NO DIFERENCIAL	3,47E-01	-0,51	NO DIFERENCIAL
PFKL	1,98E-11	0,45	NO DIFERENCIAL	4,63E-02	-0,81	NO DIFERENCIAL
PFKM	3,85E-01	-0,14	NO DIFERENCIAL	9,00E-01	-0,23	NO DIFERENCIAL
PFKP	0,00E+00	1,64	UP	7,15E-01	0,70	NO DIFERENCIAL
PFKFB1	0,00E+00	-0,86	NO DIFERENCIAL	Não detectado		
PFKFB2	1,28E-05	0,44	NO DIFERENCIAL	7,62E-01	-0,30	NO DIFERENCIAL
PFKFB3	6,12E-05	-0,45	NO DIFERENCIAL	7,27E-01	-1,18	DOWN
PFKFB4	9,06E-01	0,07	NO DIFERENCIAL	9,08E-01	-0,29	NO DIFERENCIAL
TALDO1	2,31E-12	0,50	NO DIFERENCIAL	9,48E-01	0,00	NO DIFERENCIAL
ALDOA	1,02E-07	-0,40	NO DIFERENCIAL	2,13E-03	-0,79	NO DIFERENCIAL
ALDOC	1,16E-09	0,58	NO DIFERENCIAL	1,20E-10	-1,94	DOWN
TPI1	0,00E+00	0,84	NO DIFERENCIAL	9,80E-01	-0,04	NO DIFERENCIAL
GAPDH	0,00E+00	0,93	NO DIFERENCIAL	9,63E-01	-0,38	NO DIFERENCIAL
BPGM	4,28E-03	0,24	NO DIFERENCIAL	5,16E-01	0,93	NO DIFERENCIAL
PGK1	6,33E-10	0,45	NO DIFERENCIAL	9,82E-01	0,12	NO DIFERENCIAL
PGAM1	3,18E-08	0,31	NO DIFERENCIAL	9,46E-01	0,30	NO DIFERENCIAL
PGAM2	8,46E-07	-0,47	NO DIFERENCIAL	2,51E-01	-0,81	NO DIFERENCIAL
ENO1	0,00E+00	1,19	UP	9,70E-01	0,69	NO DIFERENCIAL
ENO2	1,05E-02	-0,32	NO DIFERENCIAL	6,87E-01	0,06	NO DIFERENCIAL
ENO3	3,11E-01	-0,23	NO DIFERENCIAL	5,02E-01	-1,03	DOWN
ENO4	Não detectado			8,42E-01	-0,74	NO DIFERENCIAL
PKM2	8,17E-07	0,40	NO DIFERENCIAL	9,64E-01	0,55	NO DIFERENCIAL
LDHA	1,13E-02	0,26	NO DIFERENCIAL	9,79E-01	0,56	NO DIFERENCIAL
LDHB	0,00E+00	2,08	UP	2,99E-01	2,30	UP
LDHC	2,72E-07	0,52	NO DIFERENCIAL	Não detectado		
LDHD	2,10E-10	-0,67	NO DIFERENCIAL	1,27E-13	-3,47	DOWN
GALM	4,11E-01	0,23	NO DIFERENCIAL	8,17E-01	-0,49	NO DIFERENCIAL
Gliconeogênese						
PGM1	0,00E+00	0,96	NO DIFERENCIAL	3,69E-07	1,52	UP
PGM2	2,73E-08	0,36	NO DIFERENCIAL	8,43E-01	0,60	NO DIFERENCIAL
PGM3	1,88E-02	0,21	NO DIFERENCIAL	3,58E-01	0,79	NO DIFERENCIAL
G6PC3	0,00E+00	-0,89	NO DIFERENCIAL	7,70E-01	-0,81	NO DIFERENCIAL
GK	8,58E-01	-0,10	NO DIFERENCIAL	9,17E-01	0,01	NO DIFERENCIAL
ADPGK	3,18E-07	0,24	NO DIFERENCIAL	9,48E-01	-0,18	NO DIFERENCIAL
FBP1	0,00E+00	-2,40	DOWN	3,36E-13	-2,94	DOWN

PCK1	0,00E+00	-1,74	DOWN		Não detectado	
PCK2	1,66E-04	-0,38	NO DIFERENCIAL	7,14E-01	-0,97	NO DIFERENCIAL
PC	6,95E-01	0,18	NO DIFERENCIAL	7,68E-01	-1,15	DOWN

Metabolismo do etanol

ACSS1	2,76E-04	-0,42	NO DIFERENCIAL	1,35E-02	-1,56	DOWN
ACSS2	7,72E-01	0,12	NO DIFERENCIAL	9,46E-01	0,00	NO DIFERENCIAL
ACSS3	5,22E-15	-0,74	NO DIFERENCIAL	3,76E-07	-2,74	DOWN
ALDH1A1	9,04E-01	-0,17	NO DIFERENCIAL	3,03E-08	5,01	UP
ALDH1A2	4,44E-09	-0,43	NO DIFERENCIAL		Não detectado	
ALDH1A3	0,00E+00	1,43	UP	8,80E-05	3,37	UP
ALDH1B1	5,50E-10	0,58	NO DIFERENCIAL	5,68E-03	1,31	UP
ALDH2	6,02E-01	-0,22	NO DIFERENCIAL	2,43E-05	-2,18	DOWN
ALDH3A1	0,00E+00	1,22	UP	1,30E-10	5,02	UP
ALDH3B2	0,00E+00	-0,86	NO DIFERENCIAL	7,06E-13	-2,74	DOWN
ALDH7A1	5,35E-01	0,14	NO DIFERENCIAL	9,57E-01	0,11	NO DIFERENCIAL
ALDH9A1	8,99E-01	0,08	NO DIFERENCIAL	9,20E-01	0,04	NO DIFERENCIAL
ADH5	1,15E-05	0,19	NO DIFERENCIAL	1,58E-01	0,63	NO DIFERENCIAL
AKR1A1	9,64E-01	0,04	NO DIFERENCIAL	1,94E-02	-0,67	NO DIFERENCIAL

Glutaminólise e TCA

PDHA1	0,00E+00	0,77	NO DIFERENCIAL	9,41E-01	-0,06	NO DIFERENCIAL
PDHB	0,00E+00	-0,42	NO DIFERENCIAL	9,42E-01	-0,27	NO DIFERENCIAL
PDHX	4,77E-01	0,15	NO DIFERENCIAL	9,46E-01	0,25	NO DIFERENCIAL
DLD	9,16E-08	0,28	NO DIFERENCIAL	9,15E-01	-0,36	NO DIFERENCIAL
DLAT	2,11E-14	0,41	NO DIFERENCIAL	9,57E-01	-0,04	NO DIFERENCIAL
CS	8,46E-01	-0,07	NO DIFERENCIAL	2,84E-02	-0,51	NO DIFERENCIAL
ACO1	9,39E-01	-0,04	NO DIFERENCIAL	3,14E-03	1,75	UP
ACO2	5,98E-01	0,11	NO DIFERENCIAL	4,74E-01	-0,22	NO DIFERENCIAL
IDH1	9,36E-01	0,11	NO DIFERENCIAL	9,57E-01	0,18	NO DIFERENCIAL
IDH2	1,59E-06	0,51	NO DIFERENCIAL	9,34E-05	-1,51	DOWN
IDH3A	9,06E-01	-0,09	NO DIFERENCIAL	3,74E-01	-0,49	NO DIFERENCIAL
IDH3B	8,08E-01	0,12	NO DIFERENCIAL	9,63E-01	0,12	NO DIFERENCIAL
IDH3G	1,25E-06	0,34	NO DIFERENCIAL	8,42E-01	-0,45	NO DIFERENCIAL
OGDH	9,43E-01	0,03	NO DIFERENCIAL	9,71E-01	0,07	NO DIFERENCIAL
OGDHL	0,00E+00	1,43	UP	4,32E-01	-0,71	NO DIFERENCIAL
DLST	5,05E-01	-0,07	NO DIFERENCIAL	5,11E-03	-0,60	NO DIFERENCIAL
SUCLG1	7,12E-10	0,26	NO DIFERENCIAL	8,99E-01	-0,43	NO DIFERENCIAL
SUCLG2	0,00E+00	-0,56	NO DIFERENCIAL	8,99E-01	-0,60	NO DIFERENCIAL
SDHA	5,73E-02	0,20	NO DIFERENCIAL	9,01E-01	-0,10	NO DIFERENCIAL
SDHB	3,93E-07	0,25	NO DIFERENCIAL	9,48E-01	0,00	NO DIFERENCIAL
SDHC	9,67E-01	0,01	NO DIFERENCIAL	9,51E-01	-0,20	NO DIFERENCIAL
SDHD	2,77E-13	0,34	NO DIFERENCIAL	9,47E-01	0,20	NO DIFERENCIAL
FH	3,53E-11	0,39	NO DIFERENCIAL	6,23E-01	0,53	NO DIFERENCIAL
MDH1	0,00E+00	0,58	NO DIFERENCIAL	9,61E-01	0,19	NO DIFERENCIAL
MDH1B	0,00E+00	-0,94	NO DIFERENCIAL	3,71E-03	-1,47	DOWN
MDH2	9,94E-14	0,43	NO DIFERENCIAL	2,37E-02	-0,86	NO DIFERENCIAL

ME1	8,77E-15	0,79	NO DIFERENCIAL	9,05E-01	0,60	NO DIFERENCIAL
ME2	0,00E+00	0,90	NO DIFERENCIAL	7,54E-01	0,66	NO DIFERENCIAL
ME3	4,26E-10	-0,60	NO DIFERENCIAL	1,64E-12	4,01	UP
ACLY	1,93E-03	-0,22	NO DIFERENCIAL	3,29E-01	-0,62	NO DIFERENCIAL
GLUL	0,00E+00	-1,29	DOWN	8,81E-01	-0,56	NO DIFERENCIAL
GLS2	0,00E+00	-1,84	DOWN	3,89E-06	-1,79	DOWN
GLS	0,00E+00	1,26	UP	1,02E-02	1,90	UP
GLUD1	0,00E+00	-0,74	NO DIFERENCIAL	8,84E-01	0,37	NO DIFERENCIAL
GLUD2	0,00E+00	-0,74	NO DIFERENCIAL	1,14E-01	-1,51	DOWN
GOT1	3,32E-03	0,24	NO DIFERENCIAL	9,31E-01	-0,43	NO DIFERENCIAL
GOT2	9,74E-10	-0,34	NO DIFERENCIAL	9,67E-01	0,01	NO DIFERENCIAL

Via das Pentose fosfato

G6PD	1,02E-01	0,34	NO DIFERENCIAL	3,02E-04	-1,74	DOWN
H6PD	9,60E-01	-0,01	NO DIFERENCIAL	8,78E-01	0,48	NO DIFERENCIAL
PGLS	8,78E-01	0,12	NO DIFERENCIAL	6,25E-01	0,60	NO DIFERENCIAL
PGD	1,08E-07	0,38	NO DIFERENCIAL	9,74E-01	0,08	NO DIFERENCIAL
RPIA	0,00E+00	0,82	NO DIFERENCIAL	6,22E-01	0,57	NO DIFERENCIAL
RBKS	0,00E+00	-0,92	NO DIFERENCIAL	7,79E-01	-0,62	NO DIFERENCIAL
PRPS1	5,49E-01	0,12	NO DIFERENCIAL	8,96E-01	0,78	NO DIFERENCIAL
PRPS2	2,22E-10	0,53	NO DIFERENCIAL	5,54E-05	1,32	UP
TKT	6,05E-12	0,70	NO DIFERENCIAL	5,65E-01	-0,42	NO DIFERENCIAL
DERA	2,25E-13	0,61	NO DIFERENCIAL	9,23E-01	-0,76	NO DIFERENCIAL
RPE	2,93E-11	0,30	NO DIFERENCIAL	9,49E-01	0,16	NO DIFERENCIAL
TALDO1	2,31E-12	0,50	NO DIFERENCIAL	9,48E-01	0,00	NO DIFERENCIAL

Via da hexosamina

GFPT1	9,69E-01	-0,01	NO DIFERENCIAL	9,56E-01	0,18	NO DIFERENCIAL
GFPT2	8,52E-05	0,44	NO DIFERENCIAL	0,00E+00	4,69	UP
GNPDA1	2,48E-01	-0,09	NO DIFERENCIAL	7,39E-01	0,43	NO DIFERENCIAL
GNPDA2	0,00E+00	-0,54	NO DIFERENCIAL	3,83E-02	0,80	NO DIFERENCIAL
GNPNAT1	2,84E-09	-0,40	NO DIFERENCIAL	8,37E-01	-0,40	NO DIFERENCIAL
AMDHD1	3,53E-01	-0,25	NO DIFERENCIAL	5,92E-09	5,33	UP
AMDHD2	3,46E-05	-0,32	NO DIFERENCIAL	1,97E-01	-0,81	NO DIFERENCIAL
PGM3	1,88E-02	0,21	NO DIFERENCIAL	3,58E-01	0,79	NO DIFERENCIAL
UAP1	1,74E-07	0,37	NO DIFERENCIAL	2,37E-02	2,06	UP
UAP1L1	5,57E-02	-0,18	NO DIFERENCIAL	9,75E-02	-1,56	DOWN

Metabolismo de lipídeos

ACAA1	7,79E-01	-0,07	NO DIFERENCIAL	1,03E-08	-2,18	DOWN
ACAA2	2,60E-09	0,46	NO DIFERENCIAL	9,42E-01	0,16	NO DIFERENCIAL
ACACA	1,76E-05	-0,43	NO DIFERENCIAL	8,68E-01	0,24	NO DIFERENCIAL
ACACB	0,00E+00	-1,29	DOWN	5,20E-06	-2,64	DOWN
ACADM	7,89E-01	0,11	NO DIFERENCIAL	9,30E-01	0,46	NO DIFERENCIAL
ACADS	1,34E-13	-0,62	NO DIFERENCIAL	7,40E-01	0,59	NO DIFERENCIAL
ACADSB	0,00E+00	-2,47	DOWN	7,99E-01	-0,86	NO DIFERENCIAL
ACADV1	6,58E-01	-0,07	NO DIFERENCIAL	2,53E-01	-0,47	NO DIFERENCIAL
ACAT1	5,80E-01	-0,17	NO DIFERENCIAL	9,57E-01	-0,03	NO DIFERENCIAL

ACAT2	2,90E-10	0,58	NO DIFERENCIAL	2,33E-01	0,99	NO DIFERENCIAL
ACOT1	0,00E+00	-0,86	NO DIFERENCIAL	4,75E-01	-1,12	DOWN
ACOT2	0,00E+00	-0,84	NO DIFERENCIAL	5,07E-01	-0,62	NO DIFERENCIAL
ACOT4	0,00E+00	-1,60	DOWN	1,95E-14	-3,06	DOWN
ACOT7	3,29E-09	0,53	NO DIFERENCIAL	9,52E-01	0,66	NO DIFERENCIAL
ACOT8	9,25E-01	0,04	NO DIFERENCIAL	1,04E-03	-0,86	NO DIFERENCIAL
ACOT9	0,00E+00	0,89	NO DIFERENCIAL	2,45E-05	1,72	UP
ACOX1	9,45E-01	0,03	NO DIFERENCIAL	8,07E-01	-0,79	NO DIFERENCIAL
ACOX2	0,00E+00	-2,84	DOWN	1,41E-02	-2,12	DOWN
ACOX3	1,33E-15	-0,54	NO DIFERENCIAL	8,75E-01	-0,38	NO DIFERENCIAL
ACSL1	1,75E-05	0,42	NO DIFERENCIAL	9,57E-01	0,14	NO DIFERENCIAL
ACSL3	9,67E-01	-0,09	NO DIFERENCIAL	9,69E-01	-0,09	NO DIFERENCIAL
ACSL4	2,66E-15	0,78	NO DIFERENCIAL	0,00E+00	5,65	UP
ACSL5	1,79E-08	0,45	NO DIFERENCIAL	4,99E-01	0,73	NO DIFERENCIAL
ACSL6	0,00E+00	2,25	UP	8,93E-01	-0,79	NO DIFERENCIAL
CPT1A	0,00E+00	-0,89	NO DIFERENCIAL	1,22E-13	-1,43	DOWN
CPT1B	8,47E-04	-0,32	NO DIFERENCIAL	9,15E-01	-0,27	NO DIFERENCIAL
CPT1C	2,16E-02	0,37	NO DIFERENCIAL	5,51E-03	1,42	UP
CPT2	1,75E-05	-0,20	NO DIFERENCIAL	1,45E-02	-0,62	NO DIFERENCIAL
ECHS1	3,81E-01	-0,09	NO DIFERENCIAL	7,67E-01	-0,56	NO DIFERENCIAL
EHHADH	5,95E-01	0,12	NO DIFERENCIAL	9,10E-01	-0,27	NO DIFERENCIAL
ELOVL1	0,00E+00	0,57	NO DIFERENCIAL	6,84E-01	-0,34	NO DIFERENCIAL
ELOVL2	0,00E+00	-3,06	DOWN	3,81E-01	-1,12	DOWN
ELOVL3	0,00E+00	1,09	UP	4,71E-02	-1,94	DOWN
ELOVL4	0,00E+00	1,69	UP	8,38E-01	-0,07	NO DIFERENCIAL
ELOVL5	0,00E+00	-1,60	DOWN	4,87E-01	0,61	NO DIFERENCIAL
ELOVL6	8,17E-01	0,19	NO DIFERENCIAL	9,47E-01	-0,27	NO DIFERENCIAL
ELOVL7	1,19E-12	0,63	NO DIFERENCIAL	8,51E-01	-0,81	NO DIFERENCIAL
FAS	1,51E-11	0,68	NO DIFERENCIAL	5,02E-02	1,04	UP
FASN	0,00E+00	-1,00	NO DIFERENCIAL	5,33E-09	-2,18	DOWN
GCDH	9,30E-01	0,03	NO DIFERENCIAL	9,30E-01	0,33	NO DIFERENCIAL
HADH	9,55E-01	0,00	NO DIFERENCIAL	2,68E-01	-0,86	NO DIFERENCIAL
HADHA	5,03E-01	0,10	NO DIFERENCIAL	9,60E-01	-0,17	NO DIFERENCIAL
HADHB	2,31E-02	-0,10	NO DIFERENCIAL	9,09E-01	-0,36	NO DIFERENCIAL
HSD17B12	4,33E-10	0,39	NO DIFERENCIAL	9,46E-01	0,40	NO DIFERENCIAL
MCAT	1,41E-01	0,18	NO DIFERENCIAL	6,51E-01	-0,07	NO DIFERENCIAL
MECR	3,46E-09	0,40	NO DIFERENCIAL	9,37E-01	0,03	NO DIFERENCIAL
OXSM	9,32E-01	0,01	NO DIFERENCIAL	4,39E-02	-0,69	NO DIFERENCIAL
PPT1	0,00E+00	-0,56	NO DIFERENCIAL	2,52E-04	-1,64	DOWN
PTPLB	2,71E-05	0,38	NO DIFERENCIAL	9,41E-01	0,19	NO DIFERENCIAL
TECR	2,46E-01	-0,15	NO DIFERENCIAL	9,52E-01	-0,22	NO DIFERENCIAL

Tabela 3. Intersecção da lista de genes *up-regulated* em amostras de tecidos TN x não-TN (TCGA) com a lista de genes *up-regulated* em linhagens TN x não-TN.

Gene	Linhagens celulares			TCGA			Gene	Linhagens celulares			TCGA		
	Log2FC	Ttest	p-valor	Log2FC	Ttest	p-valor		Log2FC	Ttest	p-valor	Log2FC	Ttest	p-valor
ADD2	8,08	1,06E-04	0,00E+00	2,22	4,33E-22	0, +00	KRT75	3,39	7,28E-03	1,50E-05	3,06	1,98E-15	0,00E+00
ADM	2,83	7,22E-04	1,92E-07	1,56	1,02E-24	0,00E+00	L3MBTL4	5,27	1,04E-03	1,11032E-10	2,39	5,36E-55	0,00E+00
ADORA2B	3,14	4,40E-06	6,90E-12	1,74	1,24E-35	0,00E+00	LAMA1	7,87	3,05E-04	0,00E+00	1,44	3,90E-06	0,00E+00
AFAP1-AS1	8,14	7,71634E-07	0,00E+00	3,71	0,00	0,00E+00	LAMB3	4,39	3,37E-04	9,94E-13	1,34	7,61E-09	0,00E+00
AGPAT4	3,21	1,41E-03	6,20E-06	1,05	3,69E-29	0,00E+00	LAMC2	4,90	4,77E-04	2,89E-15	1,61	1,34E-21	0,00E+00
AKT3	3,87	3,83E-05	8,88E-16	1,02	5,33E-05	0,00E+00	LANCL3	3,60	5,27E-06	8,45E-09	2,15	8,47E-42	0,00E+00
ALDH3A1	4,95	1,50E-03	1,30E-10	1,22	5,49E-09	0,00E+00	LGR5	5,11	1,16E-03	1,50E-11	3,39	1,79E-11	0,00E+00
ANKLE1	2,69	1,22E-03	4,36E-05	2,34	8,38E-46	0,00E+00	LIF	3,68	3,19E-03	8,16E-08	1,39	4,93E-10	0,00E+00
ANKRD1	8,01	4,94E-04	0,00E+00	2,84	1,30E-18	0,00E+00	LIPG	3,86	9,31E-04	2,34E-06	2,32	9,98E-31	0,00E+00
ANKRD36BP2	1,68	1,12E-02	8,26E-03	1,42	2,62E-17	0,00E+00	LPAR3	3,02	4,75E-03	2,55E-04	1,49	1,03E-18	0,00E+00
ANP32E	1,44	1,35E-03	6,89E-03	1,58	6,37E-40	0,00E+00	LRAT	3,01	1,01E-02	7,16E-07	1,05	3,44E-12	0,00E+00
ANXA1	6,11	1,51E-07	0,00E+00	1,16	2,83E-25	0,00E+00	LY6K	5,61	1,07E-05	4,26E-12	1,72	2,47E-17	0,00E+00
ANXA8	4,09	1,74E-03	1,17E-08	2,70	4,35E-32	0,00E+00	LYN	3,92	3,29E-04	1,25E-14	1,34	2,04E-40	0,00E+00
ANXA8L1	5,27	8,17E-04	2,97E-11	2,72	1,51E-28	0,00E+00	MAML2	2,71	6,65E-05	0,00E+00	1,38	3,22E-34	0,00E+00
APCDD1L	5,39	5,04E-05	8,94E-13	2,18	1,22E-17	0,00E+00	MAMLD1	3,38	8,57E-04	1,09E-06	1,72	1,81E-29	0,00E+00
ARHGAP22	6,35	6,43E-07	0,00E+00	1,44	1,15E-35	0,00E+00	MAP7D3	2,46	5,13E-03	2,61E-09	1,28	2,68E-36	0,00E+00
ARL4C	3,03	2,21E-05	1,04E-07	1,10	2,09E-38	0,00E+00	MARCO	3,25	2,60E-03	4,07E-06	2,36	2,44E-52	0,00E+00
ARNTL2	4,00	1,66E-04	1,20E-09	1,79	2,26E-36	0,00E+00	MB21D1	3,57	2,60E-04	2,85E-09	1,26	1,92E-29	0,00E+00
B3GNT5	5,93	1,11E-06	5,55E-16	2,57	2,54E-95	0,00E+00	MELK	1,48	4,33E-03	1,28E-05	1,44	8,67E-61	0,00E+00
BAG2	3,08	9,42E-05	5,35E-09	1,62	1,35E-33	0,00E+00	MET	2,76	4,72E-05	4,92E-10	1,88	7,83E-33	0,00E+00
BATF3	5,22	9,52E-13	0,00E+00	1,11	9,97E-29	0,00E+00	MFGE8	2,59	4,70E-04	6,84E-06	2,21	2,61E-34	0,00E+00
BCAN	2,12	8,52E-03	3,96E-03	1,92	3,14E-01	0,00E+00	MFI2	2,71	4,07E-03	1,99E-06	2,55	1,00E-38	0,00E+00
BCAT1	7,67	1,13E-05	0,00E+00	1,04	4,31E-15	0,00E+00	MIR155HG	2,72	6,57E-04	4,01E-05	1,06	1,74E-21	0,00E+00
BIN1	1,93	3,83E-03	1,84E-03	1,03	1,61E-18	0,00E+00	MIR31HG	4,95	2,63E-05	1,11E-15	1,17	1,91E-01	0,00E+00
BMP2	4,88	1,27E-03	1,63E-09	2,12	1,76E-19	0,00E+00	MMP1	9,75	2,30E-04	0,00E+00	1,71	6,45E-25	0,00E+00
BMP6	3,53	4,67E-04	1,32E-05	1,53	2,15E-09	0,00E+00	MMP7	8,71	7,98E-04	0,00E+00	3,33	6,37E-38	0,00E+00
BNC1	5,83	3,70E-03	1,22E-12	3,09	5,68E-08	0,00E+00	MOXD1	5,25	8,21E-05	0,00E+00	1,28	5,42E-07	0,00E+00
BTG3	1,96	4,15E-03	4,76E-09	1,56	3,29E-73	0,00E+00	MRAP2	3,10	2,47E-03	2,87E-07	1,57	8,79E-21	0,00E+00
C17orf51	4,19	7,57E-04	2,56E-09	1,17	1,56E-33	0,00E+00	MSLN	5,45	9,15E-05	1,11E-16	2,68	1,51E-35	0,00E+00
C1orf61	3,36	9,66E-03	7,13E-04	2,36	4,47E-22	0,00E+00	MSN	7,72	1,70E-06	0,00E+00	1,16	1,09E-48	0,00E+00
C22orf23	2,13	5,92E-04	9,25E-05	1,18	1,19E-35	0,00E+00	MT1X	2,77	5,14E-04	1,78E-05	1,07	3,17E-24	0,00E+00
C6orf147	1,81	1,89E-03	3,39E-05	1,12	1,08E-13	0,00E+00	MTHFD1L	1,53	2,84E-05	9,19E-07	1,66	9,71E-79	0,00E+00
C6orf223	6,66	5,97E-04	1,11E-16	2,04	1,94E-29	0,00E+00	MURC	2,65	1,32E-04	2,87E-06	1,35	1,04E-40	0,00E+00

C9orf40	1,22	6,05E-03	3,24E-03	1,28	1,02E-50	0,00E+00	MYH15	6,10	5,86E-08	7,11E-15	1,07	4,90E-16	0,00E+00
CA13	2,15	9,22E-03	4,28E-08	1,42	6,06E-31	0,00E+00	MYO10	2,78	5,49E-11	4,13E-09	1,43	3,33E-52	0,00E+00
CA9	5,03	3,87E-04	1,76E-11	3,08	1,12E-44	0,00E+00	NALCN	4,84	3,34E-07	3,78E-11	1,35	5,61E-04	0,00E+00
CACNA1A	3,88	2,27E-03	8,27E-09	1,30	1,52E-12	0,00E+00	NCAM1	7,13	4,20E-03	0,00E+00	1,03	1,38E-05	0,00E+00
CALB2	4,04	1,28E-02	2,24E-08	2,62	2,54E-19	0,00E+00	NCAPG	1,58	5,12E-03	4,74E-05	1,03	4,24E-52	0,00E+00
CCDC18	1,97	2,40E-04	8,46E-04	1,13	1,71E-46	0,00E+00	NCOA7	2,32	7,44E-04	4,57E-02	1,38	2,51E-46	0,00E+00
CCDC71L	2,36	3,60E-07	1,12E-04	1,19	7,18E-56	0,00E+00	NEURL3	2,66	3,57E-03	2,40E-02	1,22	1,00E-24	0,00E+00
CCDC82	2,98	7,44E-07	7,02E-09	1,37	8,68E-78	0,00E+00	NFIL3	1,37	1,36E-04	4,09E-02	1,08	2,72E-73	0,00E+00
CCDC88A	5,20	9,97E-09	0,00E+00	1,21	3,94E-25	0,00E+00	NGF	6,69	5,69E-07	0,00E+00	1,03	1,94E-02	0,00E+00
CCL5	4,98	7,15E-03	7,25E-13	1,00	1,34E-13	0,00E+00	NMT2	1,56	4,96E-04	2,77E-03	1,09	5,50E-45	0,00E+00
CD109	4,58	4,75E-04	2,22E-16	1,01	9,76E-08	0,00E+00	NWD1	4,36	3,23E-03	7,37E-10	1,77	2,70E-17	0,00E+00
CD14	4,67	3,55E-04	8,75E-10	1,12	1,99E-13	0,00E+00	NXN	2,29	1,00E-02	4,78E-04	1,36	3,41E-48	0,00E+00
CD70	5,97	1,05E-03	2,22E-16	1,54	4,60E-18	0,00E+00	OGFRL1	2,30	1,45E-04	8,98E-05	1,49	4,92E-28	0,00E+00
CDC20	1,27	8,53E-03	2,61E-03	1,79	3,18E-86	0,00E+00	OSR1	5,94	1,22E-05	3,91E-13	1,88	4,20E-17	0,00E+00
CDC7	1,30	5,58E-03	5,87E-03	1,11	1,76E-38	0,00E+00	OTX1	1,71	8,09E-03	6,57E-03	2,19	2,36E-35	0,00E+00
CDCA7	2,26	7,17E-03	2,02E-11	2,38	3,15E-85	0,00E+00	P3H2	6,48	7,52E-07	2,54E-13	1,50	1,29E-09	0,00E+00
CDK6	2,53	8,62E-04	1,71E-02	1,84	3,61E-39	0,00E+00	PAX6	1,75	4,44E-03	2,86E-03	1,74	2,83E-08	0,00E+00
CENPV	4,11	5,89E-05	1,79E-10	1,35	7,76E-16	0,00E+00	PCOLCE2	5,94	2,85E-09	0,00E+00	1,57	6,06E-21	0,00E+00
CHI3L2	5,46	7,10E-04	3,53E-13	2,69	2,35E-36	0,00E+00	PCSK9	8,97	1,54E-05	0,00E+00	1,83	1,55E-09	0,00E+00
CHODL	4,89	1,16E-03	6,88E-12	3,67	7,97E-61	0,00E+00	PDE1C	7,30	3,59E-06	0,00E+00	1,65	2,09E-09	0,00E+00
CHST2	6,73	6,60E-07	0,00E+00	1,92	1,36E-39	0,00E+00	PDZK1IP1	6,64	5,37E-03	3,33E-16	1,44	2,89E-17	0,00E+00
CHST3	2,51	7,19E-04	1,70E-04	1,58	8,65E-46	0,00E+00	PELI1	1,98	2,68E-04	4,41E-05	1,26	8,51E-61	0,00E+00
CHST6	3,83	7,61E-03	2,37E-09	2,09	5,06E-28	0,00E+00	PI3	5,50	2,13E-03	2,22E-16	4,26	3,31E-37	0,00E+00
CLCN4	1,93	6,20E-04	1,03E-05	2,34	2,57E-74	0,00E+00	PIM1	2,23	5,93E-05	1,08E-12	1,39	1,45E-49	0,00E+00
CMTM7	3,06	2,47E-04	6,35E-07	1,61	2,03E-49	0,00E+00	PLAC8	4,07	4,46E-03	1,88E-09	1,12	2,08E-11	0,00E+00
CNKSR2	2,07	1,31E-02	4,82E-04	2,10	1,08E-21	0,00E+00	PLAG1	3,35	5,33E-03	4,48E-07	1,52	4,90E-10	0,00E+00
CNTNAP3	5,04	8,54E-06	2,11E-15	1,90	1,60E-26	0,00E+00	PLAGL1	3,73	1,13E-02	5,20E-10	1,46	1,43E-31	0,00E+00
COTL1	2,70	6,11E-06	3,28E-06	1,46	1,32E-47	0,00E+00	PLAUR	3,82	5,76E-06	6,00E-11	1,04	3,41E-41	0,00E+00
CPNE8	3,49	9,68E-03	3,13425E-09	1,23	6,78E-21	0,00E+00	PLB1	2,59	1,17E-04	2,97E-04	1,59	1,25E-12	0,00E+00
CREB5	4,20	6,94E-05	2,60902E-14	1,02	8,39E-15	0,00E+00	PLIN2	2,67	8,62E-03	5,10E-05	1,07	1,55E-47	0,00E+00
CSPG4	3,67	5,39E-03	3,13E-09	1,77	4,03E-08	0,00E+00	PM20D2	2,26	2,67E-03	2,75E-04	1,97	1,01E-66	0,00E+00
CSRP2	3,52	5,04E-04	2,61E-14	1,78	1,19E-52	0,00E+00	POLR3G	1,73	1,13E-03	1,26E-02	1,48	8,94E-35	0,00E+00
CTSC	2,43	1,73E-05	4,19E-08	1,22	1,13E-31	0,00E+00	POPDC3	4,04	6,38E-06	5,57E-09	2,45	9,92E-29	0,00E+00
CXCL1	9,44	1,75E-09	2,32E-09	2,84	3,87E-43	0,00E+00	POU5F1	2,33	5,96E-03	3,46E-02	2,86	7,90E-53	0,00E+00
CXCL11	3,12	1,14E-03	1,47E-11	1,48	2,57E-14	0,00E+00	PRG2	3,82	1,63E-06	6,63E-08	2,00	1,37E-27	0,00E+00
CXCL3	7,40	4,30E-10	8,48E-05	2,00	7,27E-41	0,00E+00	PRKCDBP	8,96	1,73E-11	0,00E+00	1,00	1,10E-13	0,00E+00
CXCL5	10,16	8,80E-06	0,00E+00	3,96	1,20E-44	0,00E+00	PRKCK	5,24	6,73E-05	4,02E-14	1,45	1,10E-20	0,00E+00
CXCL8	8,77	1,72E-11	0,00E+00	1,83	3,91E-26	0,00E+00	PRKD3	1,80	6,52E-06	4,90E-07	1,16	1,73E-49	0,00E+00

CYP26B1	1,80	1,65E-03	0,00E+00	1,46	5,57E-21	0,00E+00	PRNP	2,26	2,90E-04	1,19E-03	1,09	4,86E-54	0,00E+00
CYP27C1	6,56	1,42E-07	6,38E-04	3,29	9,15E-41	0,00E+00	PRSS12	5,50	8,05E-07	0,00E+00	3,00	3,75E-26	0,00E+00
CYP7B1	4,34	2,08E-04	0,00E+00	1,47	6,35E-36	0,00E+00	PRSS3	7,23	4,31E-04	7,29E-13	2,10	5,29E-32	0,00E+00
CYYR1	4,96	3,62E-04	3,81E-11	1,02	3,96E-07	0,00E+00	PRTFDC1	4,35	1,22E-03	4,07E-14	2,03	3,14E-33	0,00E+00
DAND5	2,17	3,39E-04	4,60E-11	1,05	3,48E-07	0,00E+00	PSAT1	1,85	1,23E-03	1,33E-04	3,12	1,15E-114	0,00E+00
DCHS2	4,50	1,30E-04	1,02E-03	2,67	1,73E-22	0,00E+00	PTGFR	7,51	1,63E-07	0,00E+00	1,85	1,99E-16	0,00E+00
DEPDC1	1,37	3,23E-03	2,26E-11	1,42	2,44E-59	0,00E+00	PTGS2	8,77	3,19E-07	0,00E+00	2,70	8,55E-20	0,00E+00
DNAH11	3,01	9,12E-03	1,79E-02	1,13	2,38E-17	0,00E+00	PTK7	1,59	6,21E-03	1,16E-02	1,02	6,69E-52	0,00E+00
DNER	6,32	5,15E-05	1,75E-06	3,15	2,65E-35	0,00E+00	PTPRZ1	7,05	2,22E-04	0,00E+00	2,58	3,03E-15	0,00E+00
DSC3	7,02	7,79E-05	1,11E-16	2,58	6,99E-26	0,00E+00	PTX3	7,87	2,51E-08	0,00E+00	3,05	5,22E-41	0,00E+00
DSG3	9,38	3,95E-04	0,00E+00	3,25	1,82E-13	0,00E+00	QRFRP	5,16	1,12E-04	1,48E-10	2,96	2,59E-08	0,00E+00
DZIP1	8,69	1,44E-13	0,00E+00	1,43	4,82E-38	0,00E+00	RARRES1	9,40	6,53E-04	0,00E+00	2,67	1,49E-42	0,00E+00
EFNA5	1,96	1,68E-05	0,00E+00	1,52	1,73E-37	0,00E+00	RASSF4	2,26	1,35E-03	4,09E-06	1,49	7,11E-37	0,00E+00
EGFR	3,72	1,44E-04	2,96E-03	2,59	4,23E-49	0,00E+00	RCN3	3,85	2,73E-03	1,04E-07	1,15	4,72E-09	0,00E+00
ELAVL2	5,03	6,43E-04	1,14E-05	1,55	8,29E-05	0,00E+00	RCSD1	4,20	1,26E-02	2,52E-10	1,05	8,89E-12	0,00E+00
EN1	6,64	1,92E-08	4,66E-15	4,24	1,83E-54	0,00E+00	RELN	4,62	3,29E-03	5,06E-09	1,20	6,56E-01	0,00E+00
EPHA2	2,37	1,35E-04	0,00E+00	1,02	9,78E-28	0,00E+00	RGS2	3,67	1,16E-04	1,72E-11	1,59	8,22E-22	0,00E+00
EPHB2	2,89	1,10E-03	3,49E-06	1,04	2,68E-15	0,00E+00	RIOK1	1,29	6,92E-06	2,39E-06	1,07	8,27E-39	0,00E+00
ERICH5	3,64	6,26E-03	5,73E-07	2,14	7,70E-43	0,00E+00	RNF182	6,41	8,57E-07	0,00E+00	2,51	4,03E-26	0,00E+00
ETV4	2,84	1,09E-04	1,20E-04	1,30	2,86E-19	0,00E+00	ROR1	6,94	1,15E-10	0,00E+00	1,24	2,78E-12	0,00E+00
ETV5	5,04	5,52E-06	6,88E-15	1,25	4,66E-20	0,00E+00	RSU1	1,32	2,56E-04	4,33E-03	1,14	2,61E-50	0,00E+00
ETV6	1,49	2,33E-03	4,23E-02	1,18	3,29E-36	0,00E+00	RTPA	6,06	7,14E-04	7,23E-14	1,19	6,88E-10	0,00E+00
FABP5	4,31	2,60E-03	1,18E-13	1,90	2,68E-45	0,00E+00	RUNX3	3,97	4,42E-03	1,67E-08	1,58	1,25E-39	0,00E+00
FAM135A	2,13	4,60E-03	1,19E-04	1,06	2,09E-42	0,00E+00	RYR3	2,52	1,30E-03	1,44E-04	1,99	1,63E-17	0,00E+00
FAM171A1	6,63	5,91E-07	0,00E+00	2,39	3,07E-70	0,00E+00	S100A2	8,87	1,12E-07	0,00E+00	3,66	5,22E-29	0,00E+00
FAM26F	2,46	3,72E-03	2,40E-03	1,29	1,14E-19	0,00E+00	SCHIP1	2,08	2,42E-03	3,99E-03	1,69	7,78E-42	0,00E+00
FAM81A	2,28	2,16E-03	4,03E-06	1,06	9,46E-08	0,00E+00	SEMA5B	5,82	1,21E-02	1,54E-12	1,62	2,26E-13	0,00E+00
FBN2	6,94	2,99E-03	0,00E+00	1,67	1,40E-04	0,00E+00	SERPINB2	6,54	1,07E-04	0,00E+00	2,69	3,78E-13	0,00E+00
FBXL13	3,45	3,77E-04	1,88E-09	2,09	1,88E-03	0,00E+00	SERPINB5	7,20	5,29E-03	0,00E+00	2,76	4,10E-38	0,00E+00
FERMT1	3,27	1,22E-03	2,33E-08	2,96	8,71E-49	0,00E+00	SERPINB7	5,35	1,70E-07	4,10E-13	3,14	4,17E-20	0,00E+00
FGFBP1	6,52	5,29E-05	0,00E+00	2,77	3,05E-42	0,00E+00	SERPINE2	7,72	5,16E-07	0,00E+00	1,60	3,34E-21	0,00E+00
FIGN	1,60	1,55E-03	4,07E-09	1,75	1,12E-24	0,00E+00	SFRP1	7,99	8,69E-08	0,00E+00	2,47	9,75E-59	0,00E+00
FMNL2	3,83	5,27E-11	1,38E-12	1,40	1,10E-59	0,00E+00	SH3TC2	3,76	1,12E-04	7,41E-11	1,55	4,55E-18	0,00E+00
FOSL1	6,76	2,24E-10	0,00E+00	2,17	1,17E-41	0,00E+00	SHC3	5,24	1,68E-04	1,11E-16	1,82	9,00E-15	0,00E+00
FOXC1	3,12	3,21E-03	1,78E-07	3,42	9,55E-77	0,00E+00	SIM1	6,34	1,20E-06	0,00E+00	2,03	7,69E-07	0,00E+00
FOXD3	1,91	1,08E-02	2,71E-03	1,57	7,45E-10	0,00E+00	SIRPA	2,70	2,31E-03	4,60E-04	1,11	1,71E-30	0,00E+00
FOXL1	5,78	3,92E-05	1,11E-15	2,16	4,90E-30	0,00E+00	SKP2	1,30	5,78E-04	4,63E-02	1,34	1,37E-36	0,00E+00
FOXQ1	7,56	2,87E-08	0,00E+00	2,13	9,89E-14	0,00E+00	SLAMF7	6,01	9,08E-04	3,33E-16	1,11	3,03E-09	0,00E+00

FRMD5	2,21	3,99E-03	1,23E-03	1,55	2,56E-09	0,00E+00	SLC16A1	3,92	2,41E-04	1,91E-11	1,60	5,15E-45	0,00E+00
FSCN1	4,95	6,08E-07	6,99E-15	2,00	8,37E-49	0,00E+00	SLC25A37	1,54	3,17E-06	3,73E-04	1,38	7,26E-44	0,00E+00
FXYD5	2,06	4,77E-03	9,87E-03	1,11	2,23E-26	0,00E+00	SLC35F3	4,83	8,09E-04	0,00E+00	1,81	6,71E-15	0,00E+00
FZD7	2,55	2,00E-03	1,53E-06	1,23	5,28E-28	0,00E+00	SLC43A3	3,91	1,11E-03	7,77E-16	1,81	1,04E-51	0,00E+00
GALNT13	5,03	4,13E-04	1,53E-14	2,43	1,16E-16	0,00E+00	SLC9A6	2,10	6,50E-06	7,61E-06	1,07	6,97E-64	0,00E+00
GATA6	1,68	1,65E-03	3,42E-02	1,33	1,08E-18	0,00E+00	SMOX	3,01	1,32E-06	8,32E-06	1,39	6,24E-49	0,00E+00
GBP1	4,83	7,91E-06	2,59E-11	1,46	1,35E-30	0,00E+00	SOD2	3,46	1,23E-04	5,72E-06	1,22	8,91E-36	0,00E+00
GBP5	6,18	5,32E-06	3,64E-13	1,75	7,58E-18	0,00E+00	SOX15	3,29	1,03E-02	2,99E-03	1,42	4,60E-04	0,00E+00
GBP6	6,42	2,42E-03	3,93E-14	1,06	3,42E-05	0,00E+00	SOX9	2,00	9,76E-04	2,40E-02	1,01	2,15E-34	0,00E+00
GIN54	1,32	8,09E-03	1,13E-03	1,25	4,72E-40	0,00E+00	SPEG	2,24	2,13E-03	3,06E-04	1,47	1,35E-21	0,00E+00
GLIPR2	2,51	7,94E-05	1,74E-05	1,62	2,14E-53	0,00E+00	SRSF12	2,91	1,74E-03	8,51E-06	3,32	2,96E-71	0,00E+00
GLS	1,92	3,51E-07	1,02E-02	1,26	2,45E-38	0,00E+00	ST3GAL6	2,77	9,85E-04	5,43E-12	1,76	4,72E-47	0,00E+00
GPR161	3,15	2,57E-05	1,65E-11	1,70	6,54E-46	0,00E+00	STAC	3,13	4,87E-04	5,08E-04	2,95	9,99E-59	0,00E+00
GPR3	3,28	8,38E-06	5,85E-08	1,58	2,92E-35	0,00E+00	STEAP1B	4,39	6,89E-04	3,52E-06	1,74	1,86E-19	0,00E+00
GPX7	4,05	2,06E-04	1,25E-10	1,33	1,54E-36	0,00E+00	STIL	1,03	5,47E-03	2,12E-02	1,20	3,60E-44	0,00E+00
GRIN2A	5,28	2,80E-04	1,06E-11	1,71	5,96E-01	0,00E+00	SUSD5	4,12	3,94E-04	1,69E-13	1,21	1,92E-02	0,00E+00
GSTP1	8,30	2,08E-08	0,00E+00	1,65	7,98E-76	0,00E+00	TAP2	1,78	6,28E-05	4,98E-06	1,26	8,42E-34	0,00E+00
HACD1	2,81	9,10E-03	1,23E-08	2,07	1,08E-36	0,00E+00	TBX19	1,47	3,00E-03	4,64E-02	1,75	1,65E-47	0,00E+00
HAPLN3	2,25	8,78E-04	2,68E-03	2,58	3,08E-68	0,00E+00	TFAP2E	1,41	7,86E-03	2,45E-03	1,42	2,12E-15	0,00E+00
HAS2-AS1	4,48	8,90E-05	9,21E-11	1,66	5,89E-08	0,00E+00	TFCP2L1	4,16	1,56E-03	5,34E-10	1,84	1,00E-25	0,00E+00
HAS3	4,42	5,03E-05	1,16E-11	1,03	6,47E-03	0,00E+00	TGFB1	5,05	5,01E-05	1,06E-11	1,18	5,71E-10	0,00E+00
HERC5	5,42	1,22E-05	0,00E+00	1,06	5,05E-20	0,00E+00	TIAM2	3,11	2,35E-06	4,35E-08	1,57	8,19E-25	0,00E+00
HMGA2	7,53	1,34E-08	0,00E+00	2,99	3,89E-05	0,00E+00	TLR6	5,35	5,78E-08	4,94E-14	1,06	3,18E-19	0,00E+00
HOMER3	1,43	1,59E-03	3,72E-02	1,08	1,71E-32	0,00E+00	TM4SF1	3,76	1,04E-04	2,73E-10	1,29	7,72E-34	0,00E+00
HOXA1	4,35	5,99E-05	6,19E-13	1,31	2,90E-21	0,00E+00	TMCC2	2,39	4,83E-03	8,91E-07	2,60	1,20E-49	0,00E+00
HOXA10	2,79	1,19E-02	7,36E-07	1,00	2,57E-02	0,00E+00	TMEFF1	5,36	9,16E-06	1,39E-12	1,64	1,22E-40	0,00E+00
HRCT1	3,64	1,13E-03	6,17E-09	3,15	5,21E-38	0,00E+00	TMEM158	7,12	4,77E-07	0,00E+00	2,34	3,11E-46	0,00E+00
ICAM1	3,14	3,23E-03	1,70E-09	1,28	6,00E-24	0,00E+00	TMEM171	2,88	3,23E-04	3,75E-08	2,44	2,09E-43	0,00E+00
IFRD1	1,21	2,00E-03	7,43E-03	1,58	7,98E-73	0,00E+00	TMEM217	2,58	6,29E-05	2,72E-06	1,25	5,19E-31	0,00E+00
IGF2BP2	5,36	1,36E-05	0,00E+00	3,25	4,75E-44	0,00E+00	TMEM45A	4,28	4,38E-05	4,06E-09	1,24	1,58E-19	0,00E+00
IGF2BP3	6,10	1,55E-09	0,00E+00	3,74	6,18E-39	0,00E+00	TMEM71	6,47	8,06E-04	0,00E+00	2,02	2,59E-31	0,00E+00
IGFBPL1	2,43	1,20E-02	8,98E-06	1,18	8,63E-08	0,00E+00	TNFRSF10D	4,34	2,67E-04	5,95E-14	1,37	1,51E-10	0,00E+00
IL15RA	2,43	2,28E-03	7,90E-05	1,10	1,92E-20	0,00E+00	TNFRSF21	2,72	8,64E-09	9,24E-10	1,44	4,73E-42	0,00E+00
IL1A	8,42	2,33E-04	0,00E+00	2,41	7,25E-16	0,00E+00	TNFRSF9	5,62	7,79E-06	2,85E-12	1,12	1,54E-16	0,00E+00
IL1RL2	2,64	4,05E-03	8,14E-04	1,61	4,51E-38	0,00E+00	TNFSF9	2,59	1,21E-06	4,52E-06	1,03	1,71E-16	0,00E+00
IL22RA1	2,56	1,84E-03	6,69E-06	1,22	1,25E-09	0,00E+00	TNIP3	4,94	8,57E-06	1,06E-11	2,12	1,82E-23	0,00E+00
IL32	5,39	1,90E-04	4,43E-13	1,40	2,31E-25	0,00E+00	TPX2	1,15	8,33E-03	2,54E-05	1,26	4,78E-65	0,00E+00
ITGA6	4,01	8,99E-04	5,55E-16	1,18	1,42E-13	0,00E+00	TRDMT1	2,00	6,66E-05	1,86E-05	1,14	4,35E-37	0,00E+00

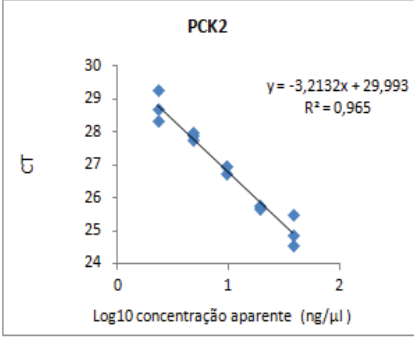
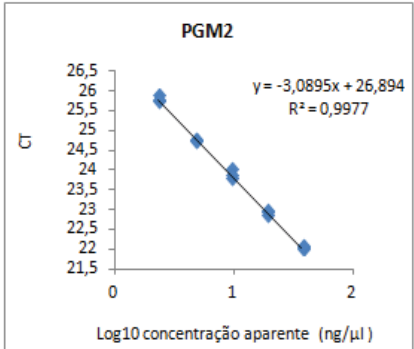
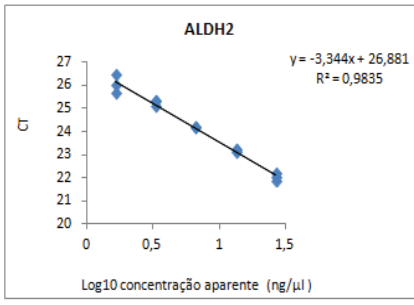
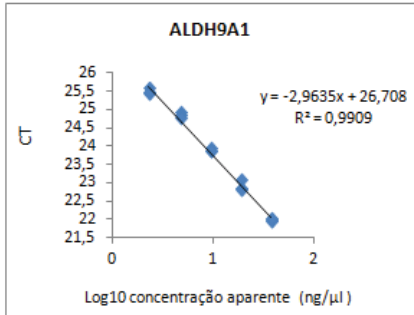
ITGB8	4,62	1,63E-04	1,45E-14	1,41	1,21E-33	0,00E+00	TTK	1,33	5,99E-03	1,20E-03	1,58	1,02E-61	0,00E+00
ITM2C	1,86	1,44E-03	3,60E-04	1,48	1,49E-42	0,00E+00	TUBB6	1,94	8,22E-05	1,18E-04	1,18	4,55E-42	0,00E+00
JRKL	1,56	3,19E-05	4,63E-09	1,19	2,09E-64	0,00E+00	UBASH3B	5,52	7,56E-08	0,00E+00	1,99	2,61E-41	0,00E+00
KCNS1	4,33	9,68E-05	3,98E-10	2,74	3,31E-44	0,00E+00	UBE2E3	1,27	3,05E-04	1,46E-03	1,25	4,43E-112	0,00E+00
KHDC1	1,81	2,57E-03	3,95E-04	1,08	7,83E-20	0,00E+00	UCHL1	5,66	1,19E-02	2,11E-15	2,78	1,28E-27	0,00E+00
KHDRBS3	2,70	5,92E-03	4,37E-06	1,76	4,47E-25	0,00E+00	UPP1	3,85	2,55E-08	1,94E-14	1,40	3,03E-38	0,00E+00
KIAA0020	1,35	1,64E-03	6,58E-03	1,17	2,51E-42	0,00E+00	USP44	3,48	2,06E-05	1,92E-08	1,18	4,82E-05	0,00E+00
KIAA1644	8,58	1,68E-07	0,00E+00	1,20	1,98E-01	0,00E+00	VANGL2	6,70	3,35E-04	0,00E+00	1,12	7,20E-22	0,00E+00
KIRREL2	4,21	3,23E-03	5,32E-10	1,34	2,23E-20	0,00E+00	VNN1	5,97	7,40E-08	7,33E-15	2,13	3,54E-14	0,00E+00
KLF5	2,54	1,78E-03	7,54E-04	1,84	7,30E-43	0,00E+00	VSNL1	5,37	9,96E-04	3,18E-14	2,35	3,46E-25	0,00E+00
KLK5	7,75	1,14E-02	0,00E+00	3,78	1,56E-22	0,00E+00	WNK3	3,65	5,98E-03	8,73E-12	1,41	1,10E-20	0,00E+00
KRT14	11,61	1,09E-05	0,00E+00	2,65	1,95E-10	0,00E+00	WNT10A	6,93	2,01E-05	0,00E+00	2,27	3,04E-32	0,00E+00
KRT17	7,31	1,51E-03	0,00E+00	2,54	1,73E-28	0,00E+00	WNT5B	5,72	1,23E-06	0,00E+00	1,31	4,23E-13	0,00E+00
KRT5	10,80	1,76E-04	0,00E+00	2,39	6,60E-27	0,00E+00	WWTR1	2,21	1,16E-07	4,23E-11	1,23	1,30E-35	0,00E+00
KRT6A	8,81	1,40E-03	0,00E+00	3,39	8,46E-35	0,00E+00	XDH	5,31	6,32E-05	1,26E-11	2,15	1,85E-07	0,00E+00
KRT6B	9,61	3,47E-03	0,00E+00	3,01	4,48E-47	0,00E+00	ZC3H12C	3,29	2,16E-03	8,80E-10	1,14	1,19E-33	0,00E+00
							ZNF280B	4,47	5,15E-03	1,48E-10	2,14	1,71E-32	0,00E+00

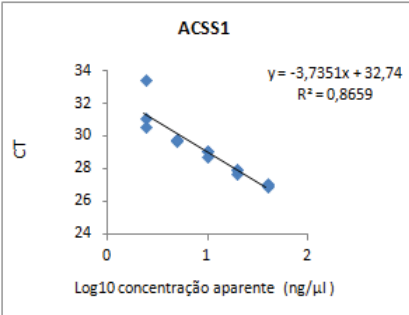
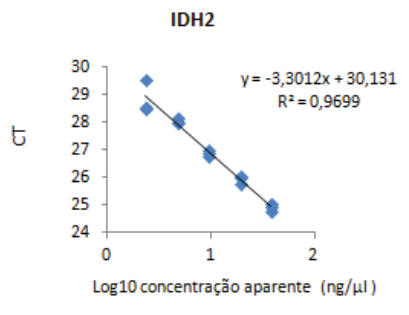
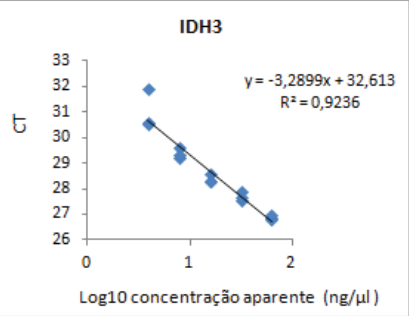
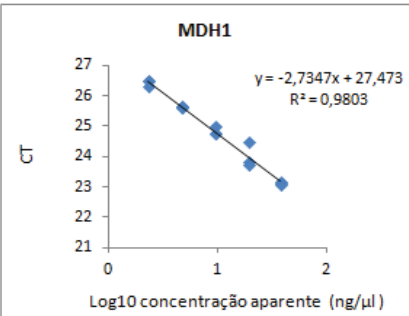
Tabela 4. Intersecção da lista de genes *up-regulated* em amostras de tecidos TN x não-TN (TCGA) sequenciadas por RNA-Seq, com a de amostras de tecidos TN x não-TN analisadas por microarranjo (*ArrayExpress*) e a lista de genes *up-regulated* em linhagens TN x não-TN (RNA-Seq).

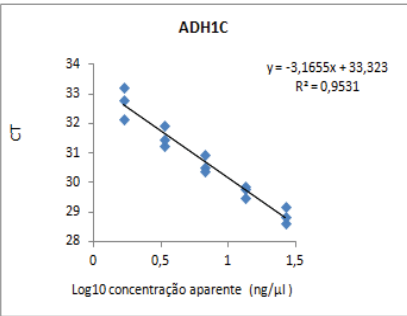
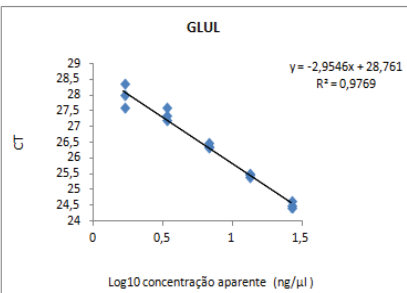
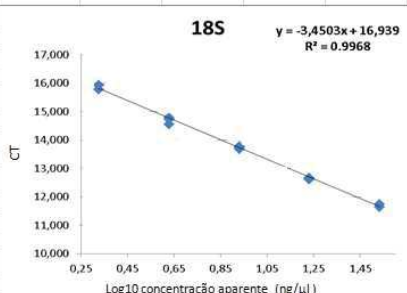
Log2FC				Log2FC			
Gene	Linhagens	TCGA	Microarranjo	Gene	Linhagens	TCGA	Microarranjo
1 ADORA2B	3,14	1,74	1,05	38 MARCO	3,25	2,36	2,86
2 ANP32E	1,44	1,58	1,14	39 MET	2,76	1,88	1,82
3 ARHGAP22	6,35	1,44	1,21	40 MFGE8	2,59	2,21	1,19
4 ARL4C	3,03	1,10	1,03	41 MFI2	2,71	2,55	1,03
5 ARNTL2	4,00	1,79	1,10	42 MMP7	8,71	3,33	2,68
6 B3GNT5	5,93	2,57	1,73	43 MOXD1	5,25	1,28	1,43
7 BAG2	3,08	1,62	1,12	44 MSLN	5,45	2,68	1,02
8 BTG3	1,96	1,56	1,14	45 MYO10	2,78	1,43	1,52
9 CALB2	4,04	2,62	1,56	46 NCOA7	2,32	1,38	1,21
10 CDCA7	2,26	2,38	1,46	47 NFIL3	1,37	1,08	1,03
11 CDK6	2,53	1,84	1,12	48 OGFRL1	2,30	1,49	1,47
12 CHI3L2	5,46	2,69	3,05	49 PCOLCE2	5,94	1,57	1,23
13 CHODL	4,89	3,67	2,09	50 PDZK1IP1	6,64	1,44	2,90
14 CSRP2	3,52	1,78	1,24	51 PRNP	2,26	1,09	1,07
15 DSG3	9,38	3,25	1,05	52 PRTFDC1	4,35	2,03	1,47
16 EGFR	3,72	2,59	1,21	53 PSAT1	1,85	3,12	2,43
17 FAM26F	2,46	1,29	1,23	54 PTK7	1,59	1,02	1,02
18 FBXL13	3,45	2,09	1,04	55 PTX3	7,87	3,05	1,06
19 FMNL2	3,83	1,40	1,18	56 RARRES1	9,40	2,67	1,95
20 FOXC1	3,12	3,42	2,56	57 RASSF4	2,26	1,49	1,01
21 FOXL1	5,78	2,16	1,15	58 RTP4	6,06	1,19	1,21
22 FXYD5	2,06	1,11	1,14	59 SERPINE2	7,72	1,60	1,29
23 GBP1	4,83	1,46	1,29	60 SFRP1	7,99	2,47	3,13
24 GBP5	6,18	1,75	1,34	61 SLC16A1	3,92	1,60	1,05
25 GSTP1	8,30	1,65	1,42	62 SLC25A37	1,54	1,38	1,10
26 HAPLN3	2,25	2,58	1,47	63 SLC43A3	3,91	1,81	1,74
27 ICAM1	3,14	1,28	1,13	64 SOD2	3,46	1,22	1,47
28 IFRD1	1,21	1,58	1,19	65 SOX9	2,00	1,01	1,20
29 IGF2BP3	6,10	3,74	1,32	66 SPEG	2,24	1,47	1,14
30 IL32	5,39	1,40	1,50	67 TIAM2	3,11	1,57	1,03
31 ITM2C	1,86	1,48	1,62	68 TNFRSF21	2,72	1,44	1,47
32 KHDRBS3	2,70	1,76	1,94	69 TPX2	1,15	1,26	1,40
33 KRT6A	8,81	3,39	1,37	70 TTK	1,33	1,58	1,09
34 KRT6B	9,61	3,01	1,85	71 UBE2E3	1,27	1,25	1,04
35 LAMC2	4,90	1,61	1,15	72 UCHL1	5,66	2,78	1,32
36 LY6K	5,61	1,72	2,01	73 WWTR1	2,21	1,23	1,04
37 LYN	3,92	1,34	1,06				

Tabela 5. Lista de *primers* utilizados para as confirmações de expressão gênica diferencial.

GENE		PRIMERS	TM °C	GC %	Tamanho do amplificado (pb)	Eficiência %	
PFKFB2	F R	CAAGCCCTGACTATCCTGAAAG GCCACGTTTATCACCTTGA	62.3 62.4	50 50	146	112	<p>PFKFB2</p> <p>$y = -3,0501x + 29,899$ $R^2 = 0,9918$</p> <p>Ct</p> <p>Log10 concentração aparente (ng/μl)</p>
ENO2	F R	CATCAAGGACAAATACGGCAAG CGATCTTTCCGTGTAGCCAG	61.8 62.7	45.5 52.4	137	110	<p>ENO2</p> <p>$y = -3,1022x + 29,813$ $R^2 = 0,9945$</p> <p>Ct</p> <p>Log10 concentração aparente (ng/μl)</p>
ENO3	F R	AATGTGGGTGATGAAGGTGG TGCGATAGAACTCAGATGCTG	62.3 61.9	50 47.6	145	103	<p>ENO3</p> <p>$y = -3,2473x + 30,768$ $R^2 = 0,9807$</p> <p>Ct</p> <p>Log10 concentração aparente (ng/μl)</p>
FBP1	F R	TGAGAAGGATGCTCTGCAAC ACCAAAATGAACTCCCCGATG	62.1 62.8	50 46,7	144	75	<p>FBP1</p> <p>$y = -4,0901x + 34,561$ $R^2 = 0,9912$</p> <p>Ct</p> <p>Log10 concentração aparente (ng/μl)</p>

PCK2	F R	GAGCTGTGGATGAGAGGTTTC CACATAGGCTGAGTCAGTGAG	62.2 61.7	52.4 52.4	128	104	 <p>PCK2</p> <p>$y = -3,2132x + 29,993$ $R^2 = 0,965$</p>
PGM2	F R	TGCTGATAGACTTGCTGTGG TGTACGTGTCTTTGAGAGCAC	61.9 62.1	50 47.6	149	110	 <p>PGM2</p> <p>$y = -3,0895x + 26,894$ $R^2 = 0,9977$</p>
ALDH2	F R	GACCGTGGTTACTTCATCCAG CTCCAACAACCTCCTCTATG	62.2 61.5	52.4 52.4	131	69	 <p>ALDH2</p> <p>$y = -3,344x + 26,881$ $R^2 = 0,9835$</p>
ALDH9 A1	F R	CTACCCCTTTTCAGATTGCCTC AGCCTCACTGTAGATTTTCAGC	62.1 61.9	52.4 47.6	127	117	 <p>ALDH9A1</p> <p>$y = -2,9635x + 26,708$ $R^2 = 0,9909$</p>

ACSS1	F R	CTCTTCATGCTGTACACCTCAG ACAGCCAAAGATGTCACCTG	62.2 62.4	50 50	135	85	 <p>ACSS1</p> <p>$y = -3,7351x + 32,74$ $R^2 = 0,8659$</p>
IDH2	F R	AAAACATCCCACGCCTAGTC GACCATTTTGAAAGTGCCGG	62.2 61.9	50 50	122	100	 <p>IDH2</p> <p>$y = -3,3012x + 30,131$ $R^2 = 0,9699$</p>
IDH3B	F R	GGGTGTGATTGAGTGTTTGAAG ACCCATCCCCAAGTTTCATG	61.7 62.3	45.5 50	143	101	 <p>IDH3</p> <p>$y = -3,2899x + 32,613$ $R^2 = 0,9236$</p>
MDH1	F R	CAGCCTTAGATAAATACGCCAAG GATCCAAACGAGTCAAGCAAC	61.2 62.4	43.5 47.6	138	132	 <p>MDH1</p> <p>$y = -2,7347x + 27,473$ $R^2 = 0,9803$</p>

ADH1C	F R	CACCTTCTCCAGTACACAG CGAAAATCCACAGCCAATGAG	61.3 61.7	56 55	100	107	 <p>ADH1C</p> <p>$y = -3,1655x + 33,323$ $R^2 = 0,9531$</p>
GLUL	F R	CCCCTTTTCGGTGACAGAAG GGCTTAGGATCAAAGGTTGC	62 61	55 50.0	141	117	 <p>GLUL</p> <p>$y = -2,9546x + 28,761$ $R^2 = 0,9769$</p>
rRNA 18S	F R	ATTCCGATAACGAACGAGAC TCACAGACCTGTTATTGCTC	55,43 55,09	45 45	129	94,90	 <p>18S</p> <p>$y = -3,4503x + 16,939$ $R^2 = 0,9968$</p>



COORDENADORIA DE PÓS-GRADUAÇÃO
INSTITUTO DE BIOLOGIA
Universidade Estadual de Campinas
Caixa Postal 6109. 13083-970, Campinas, SP, Brasil
Fone (19) 3521-6378. email: cpgib@unicamp.br



DECLARAÇÃO

Em observância ao §5º do Artigo 1º da Informação CCPG-UNICAMP/001/15, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Tese de Doutorado, intitulada ***"Transcriptômica e metabolômica como ferramentas para o entendimento do processo de adaptação metabólica tumoral dos tumores de mama triplo-negativo"***, desenvolvida no Programa de Pós-Graduação em Biociências e Tecnologia de Produtos Bioativos do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura: Melissa Quintero Escobar
Nome do(a) aluno(a): Melissa Quintero Escobar

Assinatura: Sandra Martha Gomes Dias
Nome do(a) orientador(a): Sandra Martha Gomes Dias

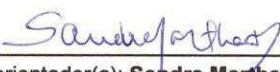
Data: 03/10/2016

Profa. Dra. Rachel Meneguello
Presidente
Comissão Central de Pós-Graduação
Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Transcriptômica e metabolômica como ferramentas para o entendimento do processo de adaptação metabólica tumoral dos tumores de mama triplo-negativo**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 03/10/2016

Assinatura : 
Nome do(a) autor(a): **Melissa Quintero Escobar**
RG n.º V745779V

Assinatura : 
Nome do(a) orientador(a): **Sandra Martha Gomes Dias**
RG n.º 177054749