

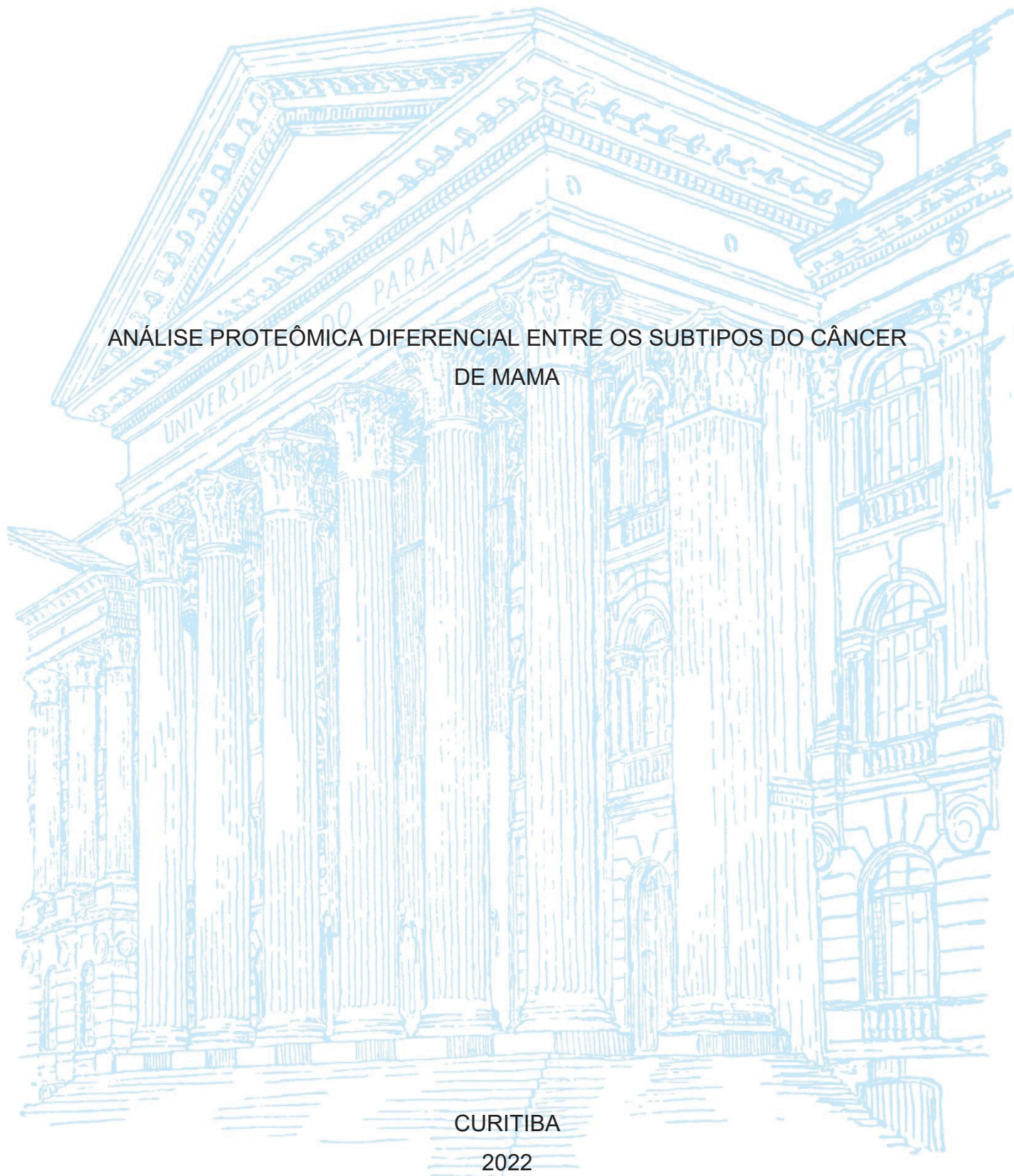
UNIVERSIDADE FEDERAL DO PARANÁ

ALEXANDRE LUIZ KORTE DE AZEVEDO

ANÁLISE PROTEÔMICA DIFERENCIAL ENTRE OS SUBTIPOS DO CÂNCER
DE MAMA

CURITIBA

2022



ALEXANDRE LUIZ KORTE DE AZEVEDO

ANÁLISE PROTEÔMICA DIFERENCIAL ENTRE OS SUBTIPOS DO CÂNCER
DE MAMA

Dissertação apresentada ao curso de Pós-Graduação em Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Mestre em Genética.

Orientador: Prof. Dr. Iglénir João Cavalli

Co-orientadora: Prof^a. Dr^a. Enilze Maria de Souza Fonseca Ribeiro

Co-orientadora: Dr^a. Talita Helen

Bombardelli Gomig

Curitiba,

2022

DADOS INTERNACIONAIS DE CATALOGAÇÃO NA PUBLICAÇÃO (CIP) UNIVERSIDADE
FEDERAL DO PARANÁ
SISTEMA DE BIBLIOTECAS – BIBLIOTECA DE CIÊNCIAS BIOLÓGICAS

Azevedo, Alexandre Luiz Korte de

Análise proteômica diferencial entre os subtipos do câncer de mama / Alexandre Luiz Korte de Azevedo. – Curitiba, 2022. 1 recurso on-line : PDF.

Dissertação (Mestrado) – Universidade Federal do Paraná, Setor de Ciências Biológicas, Programa de Pós-Graduação em Genética.

Orientador: Prof. Dr. Iglénir João Cavalli.

Coorientadoras: Prof^ª. Dr^ª. Enilze Maria de Souza Fonseca Ribeiro e Dr^ª. Talita Helen Bombardelli Gomig

1. Mama - Câncer. 2. Proteômica. 3. Genética. I. Cavalli, Iglénir João. II. Ribeiro, Enilze Maria de Souza Fonseca, 1958-. III. Gomig, Talita Helen Bombardelli, 1989-. IV. Universidade Federal do Paraná. Setor de Ciências Biológicas. Programa de PósGraduação em Genética. IV. Título.

Bibliotecária: Giana Mara Seniski Silva. CRB-9/1406



MINISTÉRIO DA EDUCAÇÃO
SETOR DE CIÊNCIAS BIOLÓGICAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO GENÉTICA -
40001016006P1

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação GENÉTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **ALEXANDRE LUIZ KORTE DE AZEVEDO** intitulada: **ANÁLISE PROTEÔMICA DIFERENCIAL ENTRE OS SUBTIPOS DO CÂNCER DE MAMA**, sob orientação do Prof. Dr. IGLENIR JOÃO CAVALLI, que após terem inquirido o aluno e realizada a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

CURITIBA, 18 de Março de 2022.

Assinatura Eletrônica

29/03/2022 16:14:36.0

IGLENIR JOÃO CAVALLI

Presidente da Banca Examinadora

Assinatura Eletrônica

18/03/2022 15:58:12.0

ANA CLAUDIA BONATTO

Avaliador Interno (UNIVERSIDADE FEDERAL DO PARANÁ)

Assinatura Eletrônica

20/03/2022 14:00:32.0

ADRIANA FROHLICH MERCADANTE

Avaliador Externo (UNIVERSIDADE FEDERAL DO PARANÁ)

AGRADECIMENTOS

Aos meus orientadores (as), que durante todo este processo estiveram ao meu lado, agradeço a oportunidade, que nasce de nosso convívio, de crescer como pessoa e como pesquisador, e por toda a dedicação e solicitude em ensinar e contribuir para minha formação.

A todos os professores e técnicos que contribuíram de diversas formas com o planejamento e execução deste trabalho.

Aos membros do Laboratório de Citogenética Humana e Oncogenética, por me receberem tão bem, lá atrás, quando iniciei minha jornada no laboratório como aluno de iniciação científica, e que me acompanharam até hoje, contribuindo com este e outros projetos nos quais tomei parte. Agradeço o companheirismo e apoio demonstrado durante estes anos.

À minha família, que me apoiou e passou confiança durante os períodos de incertezas e dúvidas. Pelo carinho e amor demonstrados no dia a dia, e pela paciência de me ouvir falar sobre genética durante os momentos mais variados.

À agência de fomento CNPq e CAPES pelo apoio financeiro.

A todos os demais que, em algum momento, contribuíram para a execução deste projeto de mestrado.

E, acima de tudo, a Deus, que nunca me abandona.

RESUMO

O câncer de mama é uma doença com altos índices de incidência, mortalidade e impactos sociais e econômicos. Os tumores mamários são classificados em subtipos, visando direcionar o tratamento e prever o prognóstico dos pacientes, porém, a classificação atual, baseada em quatro marcadores imunoistoquímicos, não é suficiente para uma classificação satisfatória destes tumores. Além disso, as características biológicas que geram as diferenças entre tumores não são completamente conhecidas. Neste sentido, as tecnologias proteômicas e interatômicas têm grande relevância, tanto na investigação das diferenças biológicas entre subtipos através da identificação de proteínas diferencialmente expressas, as quais influenciam vias e processos relacionados com as diferenças clínicas e biológicas dos subtipos, quanto na identificação de marcadores proteicos que possam complementar e otimizar a classificação e discriminação entre subtipos. Neste estudo, mais especificamente no capítulo 1, métodos proteômicos baseados em espectrometria de massas LC-MS/MS livre de marcação foram utilizados para obtenção e comparação do proteoma total de tumores mamários de diferentes subtipos (luminal A, Luminal B, HER2+ enriquecido e triplo negativo) e, através da aplicação de análises computacionais de enriquecimento funcional e de predição de interatoma, a relevância biológica de proteínas diferencialmente expressas foi evidenciada. Além disso, a tecnologia de machine learning baseada em supported vector machines foi aplicada para a obtenção de painéis proteicos com potencial para servirem como painéis de discriminação e de classificação de tumores mamários em subtipos. Proteínas diferencialmente expressas (DEPs) puderam ser identificadas entre todos os subtipos, identificando oncogenes e supressores de tumor que possuem padrões de expressão que variam segundo a subdivisão entre subtipos. Através das análises de enriquecimento, os contextos biológicos nos quais se incluem estas DEPs puderam ser explorados, sendo identificados processos e vias biológicas que possuem níveis de atuação e ativação/inativação que diferem segundo o subtipo do tumor, incluindo alterações, por exemplo, nas vias de endocitose mediada por clatrina, metabolismo de glicose e carboidratos, citoproteção por HMOX1 e regulação da estabilidade e atividade de PTEN. Os painéis proteicos aqui propostos obtiveram bom desempenho, alcançando taxas de falso-positivos sempre menores que 10%, e sensibilidade e especificidade sempre maiores que 75%. O capítulo 2 desta dissertação, por sua vez, traz análises sobre as proteínas ribossômicas (RPs) no contexto do câncer de mama, trazendo novas informações acerca do padrão de expressão proteômico e transcriptômico das RPs, bem como de seu potencial para predição de prognóstico e do panorama funcional e mutacional das RPs no câncer de mama.

Palavras-chave: Câncer de mama; subtipos; proteômica; painéis proteicos.

ABSTRACT

Breast cancer is a disease with high rates of incidence, mortality, and social and economic impacts. The mammary tumors can be classified into subtypes, aiming to manage patient treatment and predict patient prognostics, however, the actual classification method, based in four immunohistochemical markers, is not sufficient to a satisfactory classification of the tumors. Furthermore, the biological characteristics involved in the differences between subtypes are not fully known. In this context, proteomic and interatomic technologies have great relevance both in the investigation of the biological differences between subtypes through the identification of differentially expressed proteins, which influence biological pathways and processes related to clinical and biological differences among the subtypes, and in the identification of proteic biomarkers that can be used to complement and optimize the classification and discrimination between subtypes. In this study, more specifically in chapter 1, label free mass spectrometry (LC-ESI-MS/MS) based proteomics was applied to obtain and compare total proteomes of breast tumors of different subtypes and, through application of computational enrichment analysis and interatomic predictions, the biological relevance of the differentially expressed proteins (DEPs) was investigated. Further, machine learning-based technologies were applied to obtain proteic panels with the potential to discriminate and classify breast tumors in subtypes. We found DEPs in the comparisons of all subtypes, identifying oncogenes and tumor suppressor genes with expression patterns that vary accordingly with the subtype classification. Through the enrichment analysis, the biological context in which the DEPs take part was explored, being that processes and pathways that shows differential behavior among the subtypes were found, including, for example, clathrin-mediated endocytosis, glucose and carbohydrates metabolism, *HMOX1* mediated cytoprotection and PTEN stability and activity regulation. The proteic panels proposed achieved great performances, both in the cross-validation and in the independent cohort validation, reaching rates of false-positives under 10%, and sensibility and specificity over 75%. The chapter 2, in turn, describes the analysis about the ribosomal proteins (RPs) in breast cancer, contributing with new data about the RPs proteomic and transcriptomic expression pattern, as well about its potential as prognostic markers and mutational and functional landscape in breast cancer.

Key-words: Breast cancer; subtypes; proteomics; proteic panels.

LISTA DE FIGURAS

FIGURA 1 - ESQUEMA DA ESTRUTURA ANATÔMICA DA MAMA FEMININA	20
FIGURA 2 - ESQUEMA DAS ETAPAS DA PROGRESSÃO DO CARCINOMA MAMÁRIO E CÉLULAS ENVOLVIDAS	24
FIGURA 3 - <i>HALLMARKS</i> DO CÂNCER. CARACTERÍSTICAS ADQUIRIDAS PELAS CÉLULAS TUMORAIS AO LONGO DA TUMORIGÊNESE.....	25
FIGURA 4 - IMUNOISTOQUÍMICA DOS PRINCIPAIS TIPOS HISTOLÓGICOS DO CÂNCER DE MAMA	29
FIGURA 5 - PRINCIPAIS LINHAS DE PESQUISA, DENTRE AS NOVAS CIÊNCIAS ÔMICAS, UTILIZADAS PARA A INVESTIGAÇÃO DO CÂNCER DE MAMA	35
FIGURA 6 - POSSIBILIDADES DE ABORDAGENS E APLICAÇÕES DA PROTEÔMICA BASEADA EM MS NA PESQUISA DO CÂNCER	37
FIGURA 7 - <i>PIPELINE</i> CLÁSSICA PARA ANÁLISES BIOINFORMÁTICAS DE DADOS PROVENIENTES DE ESTUDOS DE PROTEÔMICA BASEADA EM ESPECTROMETRIA DE MASSAS	39
FIGURA 8 - ESQUEMA E RESULTADOS DA QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA DE CADA AMOSTRA BIOLÓGICA	66
FIGURA 9 - <i>CLUSTERIZAÇÕES</i> PARA DETERMINAÇÃO DA SIMILARIDADE ENTRE AMOSTRAS, CONSIDERANDO O PROTEOMA TOTAL	68
FIGURA 10 - <i>UPSET PLOT</i> INDICANDO A QUANTIDADE DE DEPS DE CADA COMPARAÇÃO	69
FIGURA 11 - <i>CLUSTERIZAÇÃO</i> BASEADA EM CORRELAÇÃO DE PEARSON, CONSIDERANDO APENAS AS 122 DEPS PROVENIENTES DA ANÁLISE DE ANOVA	71

FIGURA 12 - <i>CLUSTERIZAÇÃO</i> HIERARQUIZADA COM BASE NAS 122 DEPS PROVENIENTES DA ANÁLISE DE ANOVA	72
FIGURA 13 - <i>VULCANO PLOTS</i> REPRESENTANDO O RESULTADO DAS ANÁLISES DA ABORDAGEM UM-CONTRA-TODOS	73
FIGURA 14 - ONCOGENES E GENES SUPRESSORES DE TUMOR, CANDIDATOS E CANÔNICOS, DIFERENCIALMENTE EXPRESSOS NOS SUBTIPOS LUMINAIS A E B	74
FIGURA 15 - ONCOGENES E GENES SUPRESSORES DE TUMOR, CANDIDATOS E CANÔNICOS, DIFERENCIALMENTE EXPRESSOS NOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO	75
FIGURA 16 - VIAS E PROCESSOS BIOLÓGICOS DIFERENCIALMENTE ENRIQUECIDOS E COMPARTILHADOS PELOS SUBTIPOS	78
FIGURA 17 - VIAS E PROCESSOS BIOLÓGICOS SIGNIFICATIVAMENTE ENRIQUECIDOS EXCLUSIVAMENTE NOS SUBTIPOS LUMINAIS	81
FIGURA 18 - VIAS E PROCESSOS BIOLÓGICOS SIGNIFICATIVAMENTE ENRIQUECIDOS EXCLUSIVAMENTE NOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO	82
FIGURA 19 - ANÁLISE DE ENRIQUECIMENTO BASEADA EM ONTOLOGIA GÊNICA PARA AS DEPS DO TOP 30 DA CLASSIFICAÇÃO VIA SUPPORTED VECTOR MACHINES DOS SUBTIPOS LUMINAIS	86
FIGURA 20 - ANÁLISE DE ENRIQUECIMENTO BASEADA EM ONTOLOGIA GÊNICA PARA AS DEPS DO TOP 30 DA CLASSIFICAÇÃO VIA SUPPORTED VECTOR MACHINES DOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO	87
FIGURA 21 - VALIDAÇÃO CRUZADA DOS MODELOS DE CLASSIFICAÇÃO DE SUBTIPOS DE CÂNCER DE MAMA E DEPS QUE COMPÕEM CADA PAINEL	89
FIGURA 22 - <i>BOXPLOTS</i> REPRESENTANDO A EXPRESSÃO DAS DEPS QUE COMPÕEM OS PAINÉIS DE CADA SUBTIPO	91

FIGURA 23 - *OPERATING CHARACTERISTIC CURVES* E VALORES DE AUC OBTIDOS PELOS PAINÉIS PROTEICOS QUANDO APLICADOS A UMA NOVA COORTE DE AMOSTRAS 93

CAPÍTULO 2

FIGURE 1 - DIFFERENTIALLY EXPRESSED RIBOSOMAL PROTEINS (RPs) AND PROTEOMIC EXPRESSION LEVELS OF SPECIFIC RPs..... 175

FIGURE 2 - DIFFERENTIALLY EXPRESSED SMALL RIBOSOMAL PROTEINS AT mRNA LEVELS..... 176

FIGURE 3 - DIFFERENTIALLY EXPRESSED LARGE RIBOSOMAL PROTEINS AT mRNA LEVELS..... 177

FIGURE 4 - ASSOCIATIONS BETWEEN RPs mRNA EXPRESSION LEVELS AND CLINICOPATHOLOGICAL PARAMETERS OF BREAST CANCER..... 179

FIGURE 5 - PROGNOSTIC SIGNIFICANCE OF EACH RP REGARDING OS, DMFS AND DFS IN BREAST CANCER PATIENTS..... 182

FIGURE 6 - CBIOPORTAL ANALYSIS OF MUTATIONAL LANDSCAPE OF THE RPs IN METABRIC AND TCGA BREAST TUMOR SAMPLES..... 184

FIGURE 7 - FREQUENCY, CLINICAL RELEVANCE OF MUTATIONS AND GENOMIC NEIGHBORS OF THE RP GENES..... 185

FIGURE 8 - CO-EXPRESSION AND FUNCTIONAL ENRICHMENT ANALYSIS OF THE RPs IN BREAST CANCER..... 187

FIGURE 9 - GENE ONTOLOGY FUNCTIONAL ENRICHMENT ANALYSIS..... 188

FIGURE 10 - GENE REGULATORY NETWORK OF RIBOSOMAL PROTEINS EXPRESSION BY miRNAs AND TRANSCRIPTION FACTORS..... 192

LISTA DE TABELAS

TABELA 1 - PRINCIPAIS TIPOS HISTOLÓGICOS E SUAS CARACTERÍSTICAS	27
TABELA 2 - CLASSIFICAÇÃO MOLECULAR E IMUNOISTOQUÍMICA DOS SUBTIPOS DE CÂNCER DE MAMA E SUAS PRINCIPAIS CARACTERÍSTICAS	30
TABELA 3 - CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DAS PACIENTES DOADORAS DAS AMOSTRAS BIOLÓGICAS UTILIZADAS DURANTE ESTE ESTUDO	50
TABELA 4 - CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DOS TUMORES UTILIZADOS POR KRUG E COLABORADORES PARA OBTENÇÃO DOS DADOS DE EXPRESSÃO PROTEICA	59
TABELA 5 - VALORES DOS PARÂMETROS C E SIGMA E MARGEM DE ERRO PREVISTA PARA CADA UM DOS MODELOS PROPOSTOS	84
TABELA SUPLEMENTAR 1 – PARÂMETROS UTILIZADOS PARA EXECUÇÃO DAS CORRIDAS DE LC-MS/MS.....	105
TABELA SUPLEMENTAR 2 - DEPS RESULTANTES DA ANÁLISE DE ANOVA NAS COMPARAÇÕES ENTRE SUBTIPOS, E IDENTIFICAÇÃO DOS PARES ONDE A DIFERENÇA DE EXPRESSÃO É SIGNIFICATIVA ATRAVÉS DO PÓS-TESTE DE TUKEY	107
TABELA SUPLEMENTAR 3 - DEPS IDENTIFICADAS, PARA CADA SUBTIPO, ATRAVÉS DA ABORDAGEM DE UM-CONTRA-TODOS.....	113
TABELA SUPLEMENTAR 4 – ONCOGENES E GENES SUPRESSORES DE TUMOR (CANÔNICOS E CANDIDATOS) IDENTIFICADOS DENTRE AS DEPS DE CADA SUBTIPO.....	132
TABELA SUPLEMENTAR 5 - VIAS E PROCESSOS IDENTIFICADOS COMO DIFERENCIALMENTE ATIVADOS PARA CADA SUBTIPO.....	137

TABELA SUPLEMENTAR 6 - RESULTADO DO RANQUEAMENTO ORIUNDO DA ANÁLISE DE <i>MACHINE LEARNING</i> DE CADA SUBTIPO	152
--	-----

CAPÍTULO 2

TABLE 1 – ASSOCIATION BETWEEN RPs EXPRESSION AND NOTTINGHAM PROGNOSTIC INDEX AND SCARFF-BLOOM-RICHARDSON GRADES.....	181
--	-----

TABLE 2 – REACTOME ENRICHMENT ANALYSIS PERFORMED TO THE SET OF GENES THAT SHARE GENOMIC NEIGHBORHOOD WITH THE RPs GENES.....	190
--	-----

SUPPLEMENTARY TABLE 1 – PARAMETERS USED TO PERFORM THE LC-MS/MS RUNS.....	206
---	-----

SUPPLEMENTARY TABLE 2 - RIBOSOMAL PROTEINS DIFFERENTIALLY EXPRESSED.....	208
--	-----

SUPPLEMENTARY TABLE 3 – DETAILED RESULTS OF THE KAPLAN-MEIER ANALYSIS OF OVERALL-SURVIVAL, DISTANT METASTASIS-FREE SURVIVAL AND DISEASE-FREE SURVIVAL.....	211
--	-----

SUPPLEMENTARY TABLE 4 – PREDICTED REGULATORY INTERACTIONS BETWEEN miRNAs AND TRANSCRIPTION FACTORS (TFs).....	221
---	-----

SUMÁRIO

1 INTRODUÇÃO	16
2 REVISÃO DE LITERATURA.....	18
2.1 O CÂNCER DE MAMA.....	18
2.1.1 Biologia da glândula mamária humana.....	18
2.1.2 Epistemologia e etiologia do câncer de mama.....	20
2.1.3 Biologia da tumorigênese mamária.....	23
2.1.4 Classificação histológica do câncer de mama	26
2.1.5 Classificação molecular e imunoistoquímica do câncer de mama	29
2.2 ABORDAGENS PARA O ESTUDO DO CÂNCER DE MAMA.....	33
2.2.1 A proteômica, interatômica e suas abordagens.....	36
2.2.2 Análise e interpretação de dados proteômicos	38
2.2.3 As contribuições da proteômica e interatômica para o estudo do câncer de mama.....	43
3 DESCRIÇÃO DOS CAPÍTULOS	47
4 CAPÍTULO 1.....	48
4.1 RESUMO CAPÍTULO 1.....	48
5 OBJETIVOS	49
5.1 OBJETIVO GERAL.....	49
5.2 OBJETIVOS ESPECÍFICOS.....	49
6 MATERIAL E MÉTODOS.....	49
6.1 OBTENÇÃO E PROCESSAMENTO DO PROTEOMA	49
6.1.1 Caracterização das amostras biológicas	49
6.1.2 Obtenção dos extratos proteicos.....	52
6.1.3 Quantificação dos extratos pelo método do triptofano.	53
6.1.4 Separação eletroforética dos extratos proteicos.....	53
6.1.5 Digestão enzimática em gel das proteínas e preparo para LC-ESI-MS/MS. 54	
6.2 QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA TOTAL	56
6.2.1 Análise dos peptídeos por LC-ESI-MS/MS.....	56
6.2.2 Quantificação e análise dos espectros de massa obtidos	57
6.3 ANÁLISE E INTERPRETAÇÃO DOS DADOS.....	57
6.3.1 Processamento dos dados brutos	57
6.3.2 Análises estatísticas dos dados proteômicos.....	60

6.3.3	Identificação de oncogenes e genes supressores de tumor diferencialmente expressos em cada subtipo	61
6.3.4	Análise de enriquecimento e anotação funcional	61
6.3.5	Análise de predição computacional de interatoma	62
6.3.6	Análises de <i>Machine learning</i> e <i>clusterização</i> hierarquizada	63
7	RESULTADOS	65
7.1	QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA TOTAL DOS TUMORES ANALISADOS.....	65
7.2	ALTERAÇÕES DE EXPRESSÃO PROTEICA IDENTIFICADAS NAS COMPARAÇÕES ENTRE OS PROTEOMAS DE CADA SUBTIPO	69
7.3	ONCOGENES E GENES SUPRESSORES DE TUMOR COM DIFERENÇAS DE EXPRESSÃO ENTRE OS SUBTIPOS.....	74
7.4	ANÁLISES DIFERENCIAIS DE ENRIQUECIMENTO FUNCIONAL E INTERATOMA PARA CADA SUBTIPO	76
7.5	CONSTRUÇÃO DE PAINÉIS PROTEICOS A PARTIR DE MODELOS DE CLASSIFICAÇÃO BASEADOS EM <i>MACHINE LEARNING</i>	83
7.6	VALIDAÇÃO DOS PAINÉIS EM UMA NOVA COORTE DE AMOSTRAS	90
8	DISCUSSÃO	94
9	CONCLUSÃO.....	104
10	ANEXOS.....	105
11	REFERÊNCIAS.....	155
12	CAPÍTULO 2	167
	ABSTRACT.....	167
12.1	INTRODUCTION	168
12.2	MATERIALS AND METHODS.....	169
12.2.1	Patients and sample preparation.....	169
12.2.2	Protein identification and quantification by LC-MS/MS	170
12.2.3	Proteomic data analysis and statistical approach	171
12.2.4	Breast Cancer Gene-Expression Miner analysis	171
12.2.5	Genomic alteration analysis by cBioPortal database	172
12.2.6	Gene enrichment and functional analysis of GSEA/MSigDB.....	173
12.2.7	Prediction of regulation of RPs expression by miRNA and transcription factor.....	173

12.3 RESULTS	174
3.1 Transcriptomic and proteomic expression levels of large and small ribosome proteins in cancer	174
3.2 Association between RPs expression and clinicopathological parameters of breast cancer	178
3.3 Prognostic value of ribosomal proteins in breast cancer	182
3.4 Genomic alterations in RPs coding genes associated with breast cancer samples	183
3.5 Correlation between mRNA and protein levels of the RPs and enrichment analysis.....	186
3.6 Prediction of main regulators of RP expression	191
12.4 DISCUSSION.....	192
12.4 CONCLUSIONS.....	197
12.5 REFERENCES	198
12.6 SUPPLEMENTARY MATERIAL.....	206
13. REFERÊNCIAS (CAPÍTULO 1 & CAPÍTULO 2).....	224

1 INTRODUÇÃO

O câncer é um dos principais problemas de saúde a nível mundial. Em um cenário de crescimento e envelhecimento populacional, e de aumento na prevalência dos principais fatores de risco, com ênfase naqueles relacionados com o desenvolvimento econômico, o câncer se tornou uma das maiores causas de morte prematura (abaixo dos 70 anos) no mundo (INCA, 2019).

O câncer de mama é uma neoplasia multifatorial, sendo decorrente de fatores ambientais e genéticos - hereditários ou não. Dentre os fatores de risco da doença, os principais são a idade - acima dos 40 anos ou próxima à menopausa - menarca precoce e gravidez tardia, além de aspectos relacionados com o estilo de vida, como hábitos alimentares, prática de atividades físicas e uso prolongado de contraceptivos, terapias de reposição hormonal e outros (KAMIŃSKA et al., 2015). Dentre os fatores hereditários, as mutações nos genes *BRCA1* e *BRCA2* são as mais relevantes, apresentando grande correlação com o risco de desenvolvimento de câncer de mama (KAMIŃSKA et al., 2015).

Outros fatores interferem no prognóstico e seleção do tratamento dos tumores mamários, tais como *status* dos receptores de estrogênio (ER) e progesterona (PR), a expressão do receptor do fator de crescimento epidérmico (HER2) e do antígeno de proliferação celular Ki-67 (PARISE & CAGGIANO, 2014). Tumores com ausência de PR constituem a maior parte dos casos diagnosticados em mulheres jovens (< 30 anos) e com mutações em *BRCA1*, enquanto a alta expressão de ER está relacionada com tumores em mulheres mais velhas e menarca precoce (MOMENIMOVAHED & SALEHINIYA, 2019). O receptor do fator de crescimento epidérmico (HER2) possui expressão aumentada em 20-30% dos tumores mamários, sendo este aumento de expressão associado com maior agressividade, maiores taxas de recorrência e sobrevida mais breve, enquanto Ki-67 é um marcador do *status* proliferativo e de desregulação de ciclo celular (MANNELL, 2016). Além de sua utilização como marcadores de prognóstico, estes quatro fatores (ER, PR, HER2 e Ki-67) são também utilizados para a classificação dos tumores mamários em subtipos.

O câncer de mama é uma doença heterogênea, afetando sítios anatômicos diferentes na mama e apresentando tumores com características

distinguíveis entre si. Esses aspectos conferem aos tumores diferentes prognósticos e respostas a tratamentos, de maneira que a classificação dos tumores mamários em subtipos se faz necessária para que a tomada de decisões em relação a cada paciente seja a melhor possível, considerando o tipo de tumor para o qual o paciente que será tratado (FRAGOMENI et al., 2019). Atualmente, a tomada de decisão é feita individualmente levando em consideração características clínico-patológicas como morfologia, tamanho e grau do tumor, presença de metástase em linfonodos e expressão de ER, PR, HER2 e Ki-67 (GOLDHIRSCH et al., 2013).

Os sistemas de classificação tradicionais, baseados em características clínico-patológicas, possuem limitações, uma vez que estes parâmetros não são suficientes para identificar as alterações genéticas envolvidas no desenvolvimento e progressão do tumor e, além disso, mesmo tumores com parâmetros clínico-patológicos semelhantes apresentam comportamentos completamente diferentes. Assim, estudos que facilitem a identificação dos subtipos e sua caracterização molecular, a partir da identificação de marcadores moleculares, são necessários para aprimorar a estratificação dos pacientes visando facilitar o direcionamento do paciente para o tratamento do qual ele terá os melhores benefícios e menor toxicidade, além de aperfeiçoar os tratamentos disponíveis atualmente e facilitar a implementação de tratamentos customizados no futuro (YERSAL & BARUTCA, 2014).

Dentre as novas ciências “ômicas”, a proteômica e a interatômica possuem como foco de estudo a caracterização do perfil de expressão proteica (proteoma) em diferentes estados fisiológicos (GRAVES & HAYSTEAD, 2002) e a análise das interações proteicas (interatoma) e suas consequências biológicas (KANHAIYA et al., 2017). O proteoma de uma célula representa grande parte do produto de seus genes funcionais, sendo fonte de informações que permitem identificar os processos biológicos alterados nos tumores e responsáveis por sua patologia (KOOMEN et al., 2008; GAM, 2012; LOBO et al., 2017). Diversos estudos têm alcançado resultados relevantes na diferenciação proteômica e interatômica entre tecidos tumorais e não tumorais (DA COSTA et al., 2015; GOMIG et al., 2019), e também na caracterização e discriminação dos subtipos através da indicação de novos biomarcadores proteicos, sendo a espectrometria

de massas uma das técnicas mais acuradas e utilizadas nestes estudos (PANIS et al., 2014; ZAMANIAN-AZODI et al., 2015; CORRÊA et al., 2017).

Contudo, a identificação de marcadores proteicos que sejam utilizáveis para a maioria dos pacientes é um grande desafio, visto que diversos fatores, como idade, ação de vírus, mutações em genes reguladores e supressores de tumor, dieta e atividade física podem alterar a expressão proteica. Além disso, diferentes países possuem composições étnicas e culturais que tornam o padrão de expressão proteica diferente entre as populações. Esta heterogeneidade entre os perfis proteômicos faz com que as alternativas clínicas baseadas em marcadores proteicos sejam pouco adotadas atualmente, sendo necessárias mais pesquisas visando enriquecer o entendimento da expressão proteica no contexto tumoral e sua relevância clínica (GAM, 2012).

Como apontado por Tyanova e colaboradores (2016), discriminações entre os perfis proteômicos populacionais podem ser uma das grandes contribuições para a apuração da classificação dos tumores mamários em subtipos, para a compreensão das diferenças biológicas entre eles, e também para o entendimento das diferenças de respostas a tratamentos no câncer de mama. Assim sendo, este trabalho visa contribuir para o entendimento e caracterização dos subtipos do câncer de mama através da avaliação do perfil proteômico diferencial destes subtipos, da investigação de diferenças biológicas e de interatoma entre eles e da indicação de marcadores proteicos que facilitem o diagnóstico e discriminação entre subtipos.

2 REVISÃO DE LITERATURA

2.1 O CÂNCER DE MAMA

2.1.1 Biologia da glândula mamária humana

Ao nascimento, as mamas feminina e masculina não apresentam diferenças perceptíveis estrutural ou funcionalmente, permanecendo na infância como um conjunto de ductos lactíferos arranjados de maneira convergente e circunferencial em relação aos mamilos (ELLIS & MAHADEVAN, 2013). Do período da infância até a puberdade, a mama feminina irá permanecer sem alterações, até que aproximadamente dois anos antes da menarca, ela comece

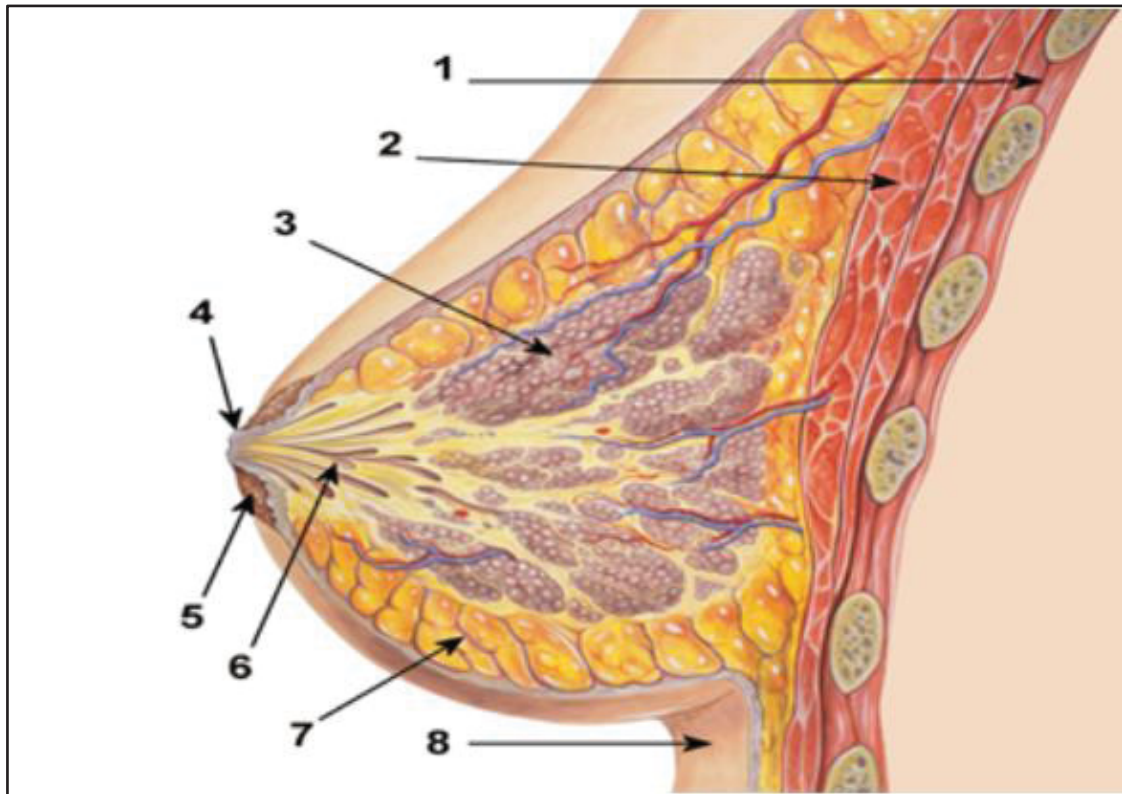
a se modificar. A partir desta fase, ocorrerão diversas alterações morfológicas e funcionais no epitélio e estroma mamários, resultantes de respostas a várias influências hormonais no microambiente da mama, sendo o estrogênio e a progesterona os hormônios mais influentes (STINGL, 2010). A presença do estrogênio é um dos sinais do início da puberdade feminina, sendo produzido pelos folículos dos ovócitos quando começam a amadurecer. A partir deste momento, o estrogênio será o desencadeador do desenvolvimento de diversas características sexuais primárias e secundárias da mulher, iniciando o desenvolvimento da mama feminina na puberdade.

Do mesmo modo, a progesterona possui função vital para o desenvolvimento do tecido mamário, estimulando a proliferação celular e a formação de alvéolos e ductos (STINGL, 2010). Outros fatores também são responsáveis pelo desenvolvimento deste tecido, como os fatores de crescimento *EGF* (*Epidermal growth factor*) e *VEGF* (*Vascular endothelial growth factor*), além das vias de sinalização WNT e Notch (ERCAN et al., 2011). A mama, uma vez madura, será constituída basicamente de estroma (tecido adiposo, conjuntivo, vasos e nervos) e tecido parenquimático, que é o conjunto de glândulas circundadas pelo estroma. O parênquima (ou tecido glandular) é formado por ácinos, os quais representam a menor parte do tecido e são responsáveis pela secreção de leite: Os conjuntos formados por ácinos são chamados de lóbulos, que em conjunto, serão conhecidos como lobos mamários (FIGURA 1). A secreção que se forma nos lobos mamários é transportada pelos ductos até a papila através do sistema lobular-ductal, e da papila estará pronta para ser exteriorizada pelos mamilos (CASTRO, 2011), cumprindo a principal função das mamas, que é a lactação. Na menopausa, a diminuição do aporte de hormônios como a progesterona e o estrogênio irá reduzir a estimulação do tecido, gerando um atrofiamento e declínio na presença de tecido glandular, o qual será substituído gradualmente por tecido adiposo (ELLIS & MAHADEVAN, 2013).

O tecido mamário, deste modo, apresenta diversos ciclos de estimulação hormonal, proliferação e involução, sendo que esta constante transformação favorece o aparecimento de alterações que podem resultar em uma proliferação

incorreta de células epiteliais dos lóbulos ou ductos, gerando, respectivamente, carcinomas lobulares ou ductais (ERCAN et al., 2011).

FIGURA 1 – ESQUEMA DA ESTRUTURA ANATÔMICA DA MAMA FEMININA.



Fonte: El Sharkawy, 2014.

(Legenda: (1) Parede da caixa torácica. (2) Músculos peitorais. (3) Lóbulos. (4) Mamilo. (5) Aréola. (6) Ductos. (7) Tecido adiposo. (8) Pele).

De fato, diversas vias vitais para o controle fisiológico e morfológico do desenvolvimento da mama encontram-se desreguladas no câncer, como as vias de sinalização mediadas por WNT/beta-catenina e pelos receptores *NRG3* de *EGFR*, além da via de sinalização Notch e várias outras. Além disso, a desregulação das vias hormonais de progesterona e estrogênios também pode contribuir para alterar o desenvolvimento normal da mama e, conseqüentemente, para um maior risco de câncer (ERCAN et al., 2011).

2.1.2 Epistemologia e etiologia do câncer de mama

O câncer é um conjunto de doenças que possui altos índices de incidência, mortalidade e impactos sociais e econômicos, sendo uma das principais barreiras para o aumento da expectativa de vida na maioria dos países

do globo. Segundo a Organização Mundial da Saúde (OMS, 2019), o câncer figura como a segunda ou principal causa de mortes prematuras em 112 dos 183 países membros da organização, estando na contramão em relação à incidência e mortalidade de outras doenças: Enquanto ataques cardíacos e doenças cardiológicas em geral têm diminuído sua proporção de mortalidade, o câncer possui suas taxas de incidência e mortalidade em plena ascensão, principalmente devido ao envelhecimento e crescimento populacional, além do aumento da prevalência e distribuição de vários dos fatores de risco da doença, principalmente aqueles relacionados com o desenvolvimento econômico (SUNG et al., 2021).

Segundo a última estimativa mundial para incidência e mortalidade do câncer, realizada pela agência GLOBOCAN de pesquisa em câncer, 19,3 milhões de casos de câncer, com aproximadamente 10 milhões de mortes devido à doença, ocorreram no ano de 2020 (SUNG et al., 2021). Ainda de acordo com a agência, 2020 ficou marcado como o ano em que o câncer de mama se tornou o câncer mais diagnosticado no mundo (sem considerar o câncer de pele não melanoma) ultrapassando o câncer de pulmão e alcançando uma ocorrência de 2,3 milhões de casos. Assim, o câncer de mama passa a representar aproximadamente 11,7% de todos os casos de câncer diagnosticados no globo, seguido por câncer de pulmão (11,4%), colorretal (10%), próstata (7,3%) e de estômago (5,6%), e a ocupar a 4ª colocação em termos de letalidade (6,9%) quando considerados ambos os sexos. Porém, quando consideramos apenas o sexo feminino, o câncer de mama é, além do mais diagnosticado (24,5%), aquele com a maior taxa de mortalidade (15,5%) no mundo.

No Brasil, estima-se a ocorrência de 625 mil novos casos de câncer no triênio 2020-2022, sendo o câncer de mama o mais frequente entre as mulheres, com 66.280 casos para cada ano do triênio, apenas atrás do câncer de pele não melanoma, e com um risco aproximado de 61,61 novos casos a cada 100 mil mulheres (INCA, 2019), constituindo-se um risco estimado maior do que a média mundial para este tipo de câncer, que é de 55,2 casos a cada 100 mil mulheres.

Epistemologicamente, o câncer de mama é uma doença complexa, multifatorial e heterogênea. Ao contrário do que se pode imaginar, o câncer de

mama possui apenas uma pequena parcela de 5-10% de casos cuja origem é hereditária, provenientes de mutações germinativas em genes de alta penetrância. Os genes *BRCA1* e *BRCA2* são os genes cujas mutações são mais reconhecidamente associadas com câncer de mama hereditário, porém genes como *PTEN*, *CHEK2*, *PALB2*, *TP53*, *ATM*, *STK11* e *CDH1* têm se destacado neste contexto devido a sua estreita relação com o desenvolvimento da neoplasia mamária (HADDAD, 2020).

Deste modo, os fatores de risco associados com a maior parcela de incidência do câncer de mama são aqueles intrínsecos ao indivíduo e ao ambiente no qual ele se insere. O risco de ocorrência do câncer de mama é consistentemente associado com a idade e com fatores hormonais e reprodutivos, como menarca precoce, primeira gravidez em idade avançada, poucos eventos de gravidez, curtos períodos de amamentação e menopausa tardia, além da utilização de terapias hormonais e uso de contraceptivos orais. Fatores relacionados com o estilo de vida, como obesidade, falta de exercícios físicos, consumo de álcool e tabaco, densidade da mama e exposição à radiação ionizante também são fatores que influenciam a chance de desenvolver câncer de mama, além da predisposição genética para a doença (SIBIO et al., 2016).

Em uma meta-análise realizada em 2012 pelo Grupo Colaborativo sobre Fatores Hormonais no Câncer de Mama (*CGHFBC*), que incluía mais de 118 mil mulheres com câncer de mama, foi indicado que o risco de incidência de câncer de mama aumenta em 5% para cada ano mais jovem da mulher em sua menarca e em 2.9% para cada ano de atraso na menopausa. Neste sentido, o *CGHFBC* conclui que há um risco maior (43%) de incidência de câncer de mama em mulheres em pré-menopausa em relação a mulheres em pós-menopausa da mesma idade. Outro exemplo de fator de risco para o câncer de mama que possui grande suporte científico é a ocorrência de obesidade e altos índices de massa corporal: Suzuki e colaboradores (2009) conduziram uma meta-análise baseada em mais de 30 estudos e determinaram que cada 5-kg/m² de aumento no índice de massa corporal, em relação ao peso ideal, está relacionado com uma chance 33% maior de desenvolver câncer de mama em mulheres pós-menopausa, principalmente na incidência de tumores ER/PR positivos.

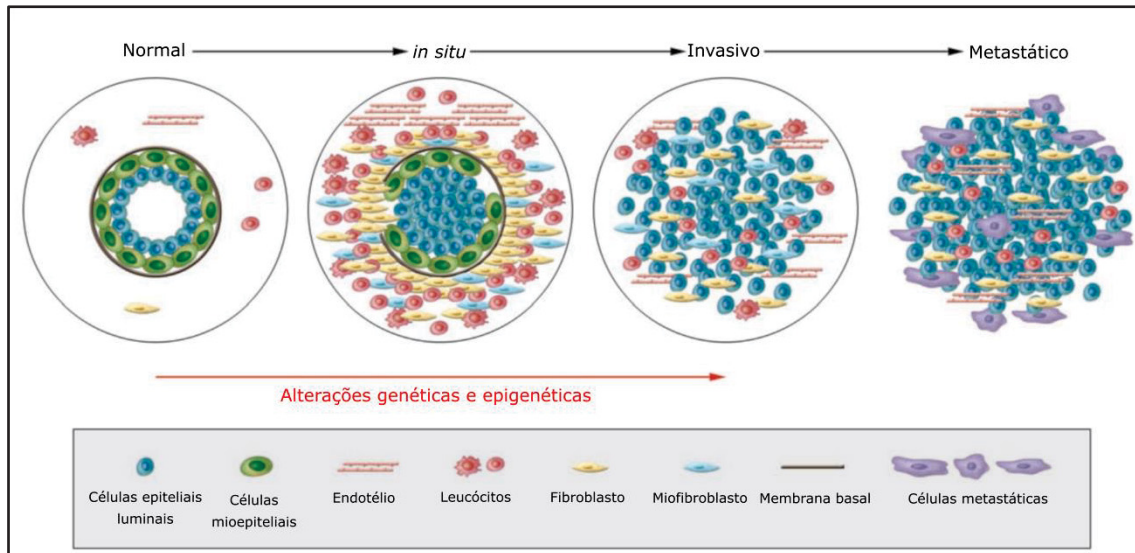
2.1.3 Biologia da tumorigênese mamária

Apesar de o entendimento acerca dos mecanismos exatos que iniciam o processo de transformação maligna das células ainda ser limitado, assim como a atual capacidade de se detectar e interferir em seu estabelecimento, sabe-se que a iniciação de uma neoplasia se dá através de transformações (genéticas e epigenéticas) que ocorrem inicialmente em uma única célula, e que a progressão do tumor se dará conforme novas alterações genéticas se acumularem e forem selecionadas durante a expansão clonal das células malignas originais (POLYAK et al., 2007). Assim, a biologia do câncer de mama envolve o início deste estado neoplásico e sua progressão através de estágios clínicos e patológicos bem definidos, iniciando-se com uma proliferação sustentada de células da mama, e seguindo através de uma evolução em um carcinoma *in situ*, que se torna invasivo e acarreta, por fim, o início de processos metastáticos (ZHANG, et al., 2017) (FIGURA 2).

O progresso da doença através dos passos que constituem a tumorigênese é causado por diversas alterações moleculares que variam desde a expressão diferencial de genes, proteínas e RNAs reguladores, alterações epigenéticas e até mesmo participação de variados tipos celulares, como fibroblastos, neutrófilos e macrófagos, que contribuem com alterações de microambiente que podem favorecer o desenvolvimento do tumor (POLYAK, et al., 2007). Atualmente, sabemos que cada estágio do desenvolvimento do tumor mamário possui assinaturas moleculares distintas, as quais conferem para cada etapa da progressão tumoral alterações em vias de sinalização e processos biológicos específicos para cada etapa (SIMPSON, et al., 2005). A conversão metastática das células do tumor mamário, que define a última etapa da evolução do câncer de mama possui, por exemplo, assinaturas moleculares bem definidas. Esta etapa é caracterizada pela aquisição de características mesenquimais por células epiteliais através de um processo denominado transição epitelial-mesenquimal (EMT). A EMT, por sua vez, possui alterações genéticas razoavelmente específicas: O processo se inicia com a super-expressão de proteínas mesenquimais, como a fibronectina e metaloproteinases, e perda de moléculas de adesão celular, como a E-caderina e B-caderina. A diminuição da expressão de E-caderina, por exemplo, irá resultar

na perda de adesão das células epiteliais aos desmossomos e, conseqüentemente, no aumento da motilidade e capacidade de migração destas células. (OTTEWELL, et al., 2015).

FIGURA 2 – ESQUEMA DAS ETAPAS DA PROGRESSÃO DO CARCINOMA MAMÁRIO E CÉLULAS ENVOLVIDAS.



Fonte: Adaptado de POLYAK et al., 2007.

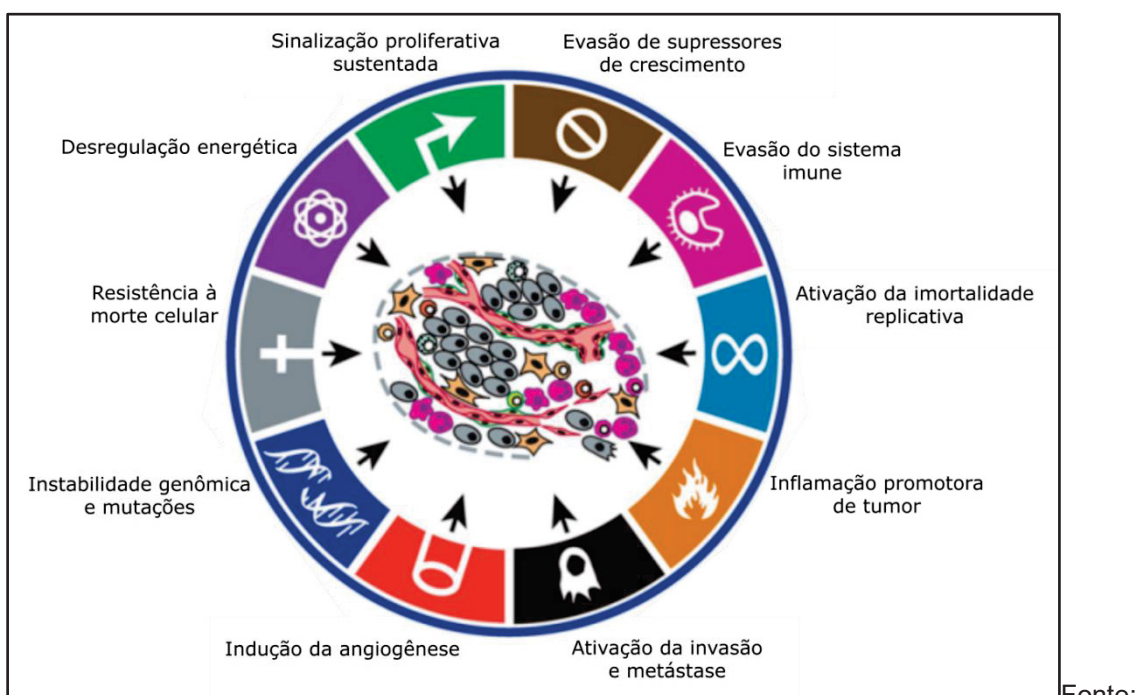
De maneira geral, as alterações moleculares intra ou extracelulares que acompanham a reprogramação celular que resulta no instaurar e progressão de uma neoplasia (como o câncer de mama), estão reunidas dentro dos *hallmarks* do câncer (HANAHAN & WEINBERG, 2011). Os *hallmarks* do câncer propõem a organização das principais características das células malignas em categorias, de forma a sistematizar a diversidade e complexidade dos processos que induzem a progressão tumoral e facilitar seu entendimento (FIGURA 3). Cada *hallmark* representa, deste modo, uma característica adquirida pela célula tumoral durante a tumorigênese que permite esta célula crescer, proliferar e se disseminar através da metástase.

Os *hallmarks* incluem categorias como as de sinalização proliferativa sustentada, desregulação energética e instabilidade genômica e mutações. A sustentação e prolongamento do estado proliferativo é a característica mais fundamental de uma célula tumoral. Enquanto uma célula normal consegue cuidadosamente regular a produção e liberação de fatores de crescimento, e conseqüentemente sua entrada e progressão no ciclo celular, células tumorais

perdem essa capacidade de regulação e se tornam ‘independentes’ em termos de crescimento e proliferação. Porém, a capacidade proliferativa não depende apenas da perda do controle sobre o ciclo celular, estando também envolvida a reprogramação do metabolismo celular, o qual garantirá à célula a energia necessária para permitir seu crescimento e divisão exacerbada. O *hallmark* de instabilidade genômica e mutações, por fim, está no cerne da aquisição de todas as outras características. Através da perda da capacidade de reparo e dos sistemas de manutenção da integridade genômica, mutações poderão ser mais facilmente retidas e selecionadas nos clones gerados pela célula maligna original, mutações as quais orquestrarão a aquisição das demais capacidades ao longo da tumorigênese (HANAHAN & WEINBERG, 2011).

De modo geral, cada um dos *hallmarks* possui alterações moleculares peculiares, que irão alterar vias de sinalização, de reparo, metabólicas e outras, cada uma delas relacionadas com algum aspecto do câncer. Dentre as alterações moleculares que resultam na aquisição destes *hallmarks* estão as alterações proteômicas e interatômicas (KIM et al., 2012; CARVALHO et al., 2020; CHAKAFANA & SHONHAI, 2021).

FIGURA 3 – HALLMARKS DO CÂNCER. CARACTERÍSTICAS ADQUIRIDAS PELAS CÉLULAS TUMORAIS AO LONGO DA TUMORIGÊNESE.



Adaptado de HANAHAN & WEINBERG, 2011.

2.1.4 Classificação histológica do câncer de mama

O câncer de mama apresenta grande heterogeneidade histológica, sendo possível, através de estudos morfológicos, determinar diferentes classificações histológicas para estes tumores. Por meio desta classificação é possível estimar se o tumor terá um comportamento mais agressivo; a chance do tumor metastatizar; a expectativa de sobrevida da paciente e, desta forma, facilitar a determinação da melhor abordagem de tratamento (MAKKI, 2015).

O estudo morfológico do câncer de mama irá investigar se o tumor está contido no componente epitelial do tecido, ou se as células já invadiram o estroma, e também se o tumor se originou nos ductos mamários ou nos lóbulos. A localização do tumor, nos ductos ou lóbulos, determina a classificação histológica mais ampla: Aproximadamente 50-80% dos tumores mamários diagnosticados são carcinomas ductais invasivos (CDIs), sendo o restante (aprox. 20-50%) classificado como carcinomas lobulares invasivos (CLIs). Porém, além da localização do tumor no tecido mamário, outros elementos, tais como características morfológicas das células, número de células e mitoses, tipo e localização de secreções, perfil imunohistoquímico e arquitetura do tecido afetado, são investigados para aperfeiçoar a classificação dos tumores (NASCIMENTO & OTONI, 2020). Caso os CDIs/CLIs não possuam características morfológico-histológicas suficientes para se determinar um tipo histológico especial, eles serão classificados como “sem tipo específico” e “clássico” respectivamente, porém, caso eles apresentem características distintivas suficientes, eles serão classificados em “tipos especiais”.

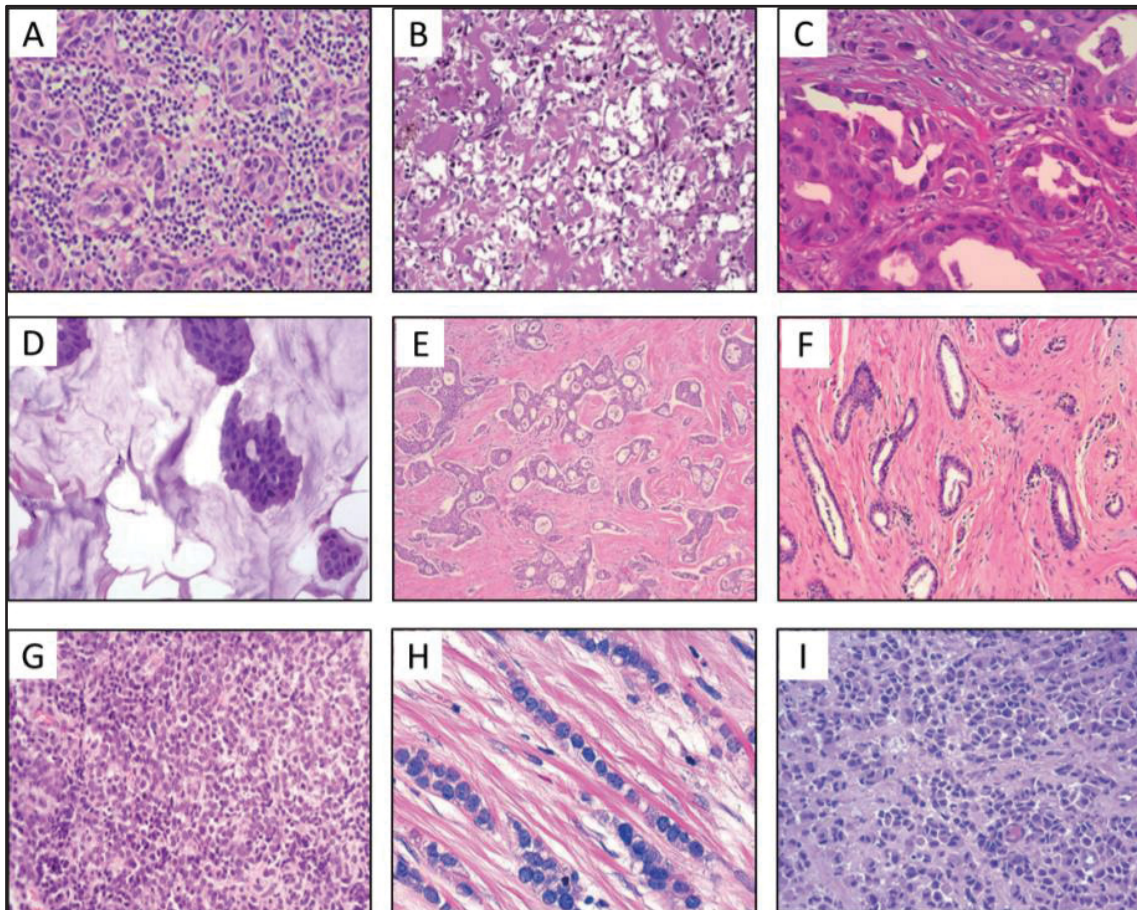
Os tipos especiais de CDI mais comuns são os carcinomas medulares, metaplásicos, apócrinos, mucinosos, cribriformes, tubulares e neuroendócrinos, enquanto os CLIs podem ser clássicos ou pleomórficos (TABELA 1; FIGURA 4).

TABELA 1 – PRINCIPAIS TIPOS HISTOLÓGICOS E SUAS CARACTERÍSTICAS
(NASCIMENTO & OTONI, 2020).

Classificação histológica	Principais características
Carcinoma ductal invasivo “sem tipo específico”	O subtipo CDI “sem tipo específico” representa aproximadamente 40-75% dos casos de carcinomas mamários diagnosticados, apresentando grande variação morfológica e comportamento clínico. Os tumores são pleomórficos e possuem muitos nucléolos e mitoses: Necroses e calcificações podem ser notadas na maioria das pacientes.
Carcinoma medular	É um tipo especial de carcinoma invasivo, representando aproximadamente 5% dos casos diagnosticados. Este tipo especial está associado com resultados clínicos favoráveis e baixas taxas de comprometimento de linfonodos axilares. Normalmente afeta pacientes com 30-40 anos de idade, frequentemente sendo associado à mutações do gene <i>BRCA1</i> . Normalmente apresenta células tumorais gigantes e metaplásicas.
Carcinoma metaplásico	Este tipo especial representa apenas 1% dos casos diagnosticados, afetando principalmente mulheres em pós-menopausa. Sua principal característica é a predominância de componentes em diferenciação metaplásicos, acompanhado de altas taxas de comprometimento de linfonodos axilares e comportamento agressivo. Em termos morfológicos, este tipo é altamente heterogêneo, contendo células ductais misturadas com outros tipos celulares, como células mesenquimais e células escamosas.
Carcinoma apócrino	Apresentando uma proeminente diferenciação apócrina em mais de 90% das células do tumor, este tipo especial representa de 1-4% dos casos diagnosticados.
Carcinoma mucinoso	Também conhecido como carcinoma mucoso, mucoide, coloide ou gelatinoso, este tipo especial representa aproximadamente 2% dos diagnósticos de câncer de mama. Ele é frequentemente associado com pacientes com mais de 60 anos de idade, e possui em geral

	<p>um prognóstico favorável. Morfologicamente, estes tumores são caracterizados pela presença de uma abundante mucina extracelular, a qual rodeia pequenos amontoados de células tumorais.</p>
Carcinomas cribriformes	<p>Tipo especial associado com bom prognóstico, baixas taxas de metástase e com pacientes com mais de 50 anos de idade, representando aproximadamente 1-3,5% dos casos diagnosticados. Este tipo apresenta amontoados isolados de células tumorais com baixo grau de atipia, mas com aspecto cribriforme em mais de 90% das células.</p>
Carcinomas tubulares	<p>O subtipo especial tubular é bem diferenciado, ocorrendo em mulheres com 50-60 anos de idade e representando 2% dos casos diagnosticados. A maioria dos tumores está associada com lesões pré-malignas na mama, e com uma proliferação proeminente dos túbulos, os quais apresentam lúmen aberto, sem presença de necrose ou mitoses.</p>
Carcinomas neuroendócrinos	<p>Representam de 0,5-5% dos tumores mamários diagnosticados, atingindo principalmente pacientes com idade bem avançada. Estes tumores possuem características parecidas com tumores neuroendócrinos do trato respiratório e gastrointestinal, apresentando expressão dos marcadores cromogranina A e sinaptofisina em mais de 50% das células do tumor.</p> <p>Morfologicamente, estes tumores possuem um padrão de crescimento infiltrativo com agregados sólidos com formato alveolar, trabecular ou em roseta,</p>
Carcinoma lobular (clássico e pleomórficos)	<p>O carcinoma lobular representa de 5-15% das neoplasias mamárias diagnosticadas, afetando principalmente pacientes com idade avançada. A forma clássica é caracterizada por pequenas células tumorais, com pouca atipia e distribuídas uniformemente e de forma concêntrica pelo estroma. Já o tipo pleomórfico é caracterizado pela presença de células hiper-cromáticas, apócrinas e com núcleo excêntrico e mitoses proeminentes. Percebe-se uma forte associação deste tipo especial com mutações no gene <i>TP53</i>.</p>

FIGURA 4 – IMUNOISTOQUÍMICA DOS PRINCIPAIS TIPOS HISTOLÓGICOS DO CÂNCER DE MAMA.



Fonte: NASCIMENTO & OTONI, 2020.

(Legenda: (A) Carcinoma medular. (B) Carcinoma metaplásico. (C) Carcinoma apócrino. (D) Carcinoma mucinoso. (E) Carcinoma cribriforme. (F) Carcinoma tubular. (G) Carcinoma neuroendócrino. (H) Carcinoma lobular clássico. (I) Carcinoma lobular pleomórfico)

2.1.5 Classificação molecular e imunohistoquímica do câncer de mama

O câncer de mama é um conjunto de doenças que possui um amplo espectro fenotípico, com diferentes comportamentos clínicos e de resposta a tratamentos. Apenas a classificação histológica/morfológica, mesmo amparada pela análise de fatores clínico-patológicos como índice mitótico, tamanho do tumor, comprometimento de linfonodos sentinelas e presença de metástase, não é suficiente para pautar a escolha do melhor tratamento ou abordagem clínica, ou mesmo para prever o real comportamento patofisiológico do tumor. Neste sentido, diversos pesquisadores têm explorado os aspectos moleculares dos

tumores mamários visando facilitar a gestão clínica dos pacientes através do agrupamento destes pacientes em classes mais homogêneas (NASCIMENTO & OTONI, 2020).

Em um trabalho considerado pioneiro, liderado por Perou e Sorlie (2000), os tumores mamários foram classificados em subtipos com base no padrão de similaridades e diferenças dos perfis de expressão gênica, os quais refletem mais fielmente o fenótipo, prognóstico e padrão de resposta a tratamento destes tumores. Esta classificação molecular resultou no surgimento de quatro subtipos: luminal A, luminal B, HER2+ enriquecido e *basal-like*. Os principais grupos de genes que pautaram esta divisão são compostos por genes relacionados com a expressão do receptor de estrogênio (ER), receptor de progesterona (PR), HER2 (*Human epidermal growth factor receptor 2*) e o regulador/marcador de proliferação celular Ki-67. Posteriormente, um painel imunistoquímico baseado em quatro marcadores (ER/PR/HER2/Ki-67) foi aceito como suficiente para uma satisfatória estratificação dos tumores em quatro subtipos (luminal A, luminal B, HER2+ enriquecido e triplo negativo) (TABELA 2).

TABELA 2 – CLASSIFICAÇÃO MOLECULAR E IMUNOISTOQUÍMICA DOS SUBTIPOS DE CÂNCER DE MAMA E SUAS PRINCIPAIS CARACTERÍSTICAS.

Subtipo	Luminal A	Luminal B		HER2+ enriquecido	<i>Basal-like</i> /Triplo negativo
		HER2-	HER2+		
Biomarcadores	ER+ PR+ HER2- Ki-67 < 14%	ER+ PR- HER2- Ki-67 > 14%	ER+ PR-/+ HER2+ Qualquer Ki-67	ER- PR- HER2+ Ki-67 > 14%	ER- PR- HER2- Ki-67 > 14%
Frequência	40-50%	20-30%		15-20%	10-20%
Grau histológico	Bem diferenciado (Grau 1)	Moderadamente diferenciado (Grau 2)		Pouco diferenciado (Grau 3)	Pouco diferenciado (Grau 3)
Prognóstico	Favorável	Intermediário		Desfavorável	Desfavorável
Resposta a terapias	Endócrina	Endócrina Quimioterapia	Endócrina Quimioterapia Terapia alvo	Quimioterapia Terapia alvo	Quimioterapia Inibidores PARP

Fonte: Adaptado de GOLDHIRSCH et al., 2013; NASCIMENTO & OTONI, 2020.

Esta simplificação, baseada na imunoistoquímica destes quatro marcadores, além de significativamente representar a classificação molecular (apesar de não haver completa correspondência entre as classificações), é mais acessível e possui custo relativamente mais baixo, sendo a classificação mais comumente utilizada na prática clínica (GOLDHIRSCH et al., 2013; BALIC et al., 2019).

Os quatro subtipos clássicos possuem características e comportamentos bastante próprios (GOLDHIRSCH et al., 2013; DAI et al., 2015; FRAGOMENI et al., 2019; NASCIMENTO & OTONI, 2020):

LUMINAL A: Este subtipo representa aproximadamente metade (50%) dos casos de câncer de mama diagnosticados. O subtipo luminal A é definido através do seguinte perfil imunoistoquímico: ER+ ($\geq 1\%$), alta expressão de PR ($\geq 20\%$), HER2- ($\leq 10\%$) e Ki-67 ($< 14\%$). Molecularmente este subtipo é caracterizado pela alta expressão dos genes *ESR1* e *XBP1* e por mutações nos genes *PIK3CA*, *MAP3KI*, *GATA3* e *FOXA1*, além de ganhos em 1q e 8q e perdas em 8p e 16q e, devido às características das células luminiais-epiteliais, das quais os tumores deste subtipo se originam, este subtipo apresenta alta expressão das citoqueratinas 7/8/18/19.

O subtipo luminal A é associado frequentemente a tumores de baixo grau histológico, tais como CDIs “sem tipo específico” e carcinomas tubulares, cribriformes, mucinosos e CLIs “clássicos”, e com prognóstico favorável, menor agressividade, evolução mais lenta (em relação aos demais subtipos) e com menores taxas de comprometimento de linfonodos e metástase. Devido à presença do receptor de estrogênio, estes tumores respondem favoravelmente à terapia endócrina.

LUMINAL B: O subtipo luminal B representa de 20-30% dos tumores mamários diagnosticados mundialmente, podendo ser subdividido em luminal B HER2- (ER+ [$\geq 1\%$], PR- ou $< 20\%$, HER2- [$\leq 10\%$] e alto nível de Ki-67 [$\geq 14\%$]) e luminal B HER2+ (ER+ [$\geq 1\%$], HER2+ [$>10\%$], e qualquer expressão de ER ou Ki-67). Em termos moleculares, o subtipo luminal B apresenta ampliações dos genes *CCND1* e *MDM2*, mutações nos genes *TP53* e *PIK3CA*, e perda da expressão de *ATM*, além de ampliações focais em 8p12 e 11q13, por exemplo,

e expressão de citoqueratinas de baixo peso molecular devido à sua origem derivada de células luminiais-epiteliais. A principal diferença molecular entre os subtipos luminiais é a maior expressão, em luminal B, de genes relacionados com proliferação celular, como *CCNE1* e *NSEP1*, e maior ativação relativa das vias alternativas dos fatores de crescimento codificados pelos genes *PI3K* e *SRC*.

Este subtipo apresenta normalmente baixos graus histológicos, na maioria CDIs, e é associado a um prognóstico intermediário e maiores chances de metástase em relação ao subtipo luminal A. Dos subtipos ER positivos, o luminal B é o mais agressivo, e nem sempre as pacientes apresentam boa resposta a terapia endócrina, sendo necessária a complementação com quimioterapia (HER2+/-) ou terapia alvo (HER2+).

HER2+ ENRIQUECIDO: Este subtipo representa de 15-20% dos casos de câncer de mama diagnosticados, sendo caracterizado imunoistoquimicamente pela alta expressão de HER2 (>10%), expressão negativa de ER (< 1%) e PR (<20%) e alta expressão do marcador Ki-67 (>14%). Para este subtipo, a técnica de hibridização *in situ* fluorescente (FISH) é comumente utilizada para confirmar a alta expressão de HER2, verificando se o gene *ERBB2* (que codifica HER2) está amplificado, caso a imunoistoquímica apresente uma coloração apenas 'fraca' de HER2. Molecularmente, o subtipo HER2+ enriquecido apresenta amplificação dos genes *ERBB2* e *CCND1*, mutações em *TP53*, *PIK3CA* e *APOBEC*, além de alta expressão de *FGFR4* e *EGFR*, baixa expressão de genes relacionados com ER e elevada instabilidade genômica.

A alta expressão de HER2 faz este subtipo ser frequentemente relacionado a tumores de alto grau histológico, alta taxa proliferativa e de metástases, acometimento de linfonodos e, conseqüentemente, com pior prognóstico e sobrevida. Apesar de tudo, pacientes com tumores mamários deste subtipo respondem bem a drogas que bloqueiam a atividade de HER2, como os anticorpos monoclonais (Trastuzumab) e o bloqueador do receptor tirosina-quinase (Lapatinib).

TRIPLO NEGATIVO: O subtipo triplo negativo é normalmente considerado como um correspondente ao subtipo molecular *basal-like* (sobreposição em cerca de 80% dos casos), e representa cerca de 10-20% dos casos de câncer de mama

diagnosticados atualmente. Imunoistoquimicamente, este subtipo é caracterizado pela ausência de expressão dos receptores ER (< 1%), PR (<20%) e da oncoproteína HER2 (\leq 10%), sendo extremamente proliferativo (Ki-67 > 30%).

Este subtipo é altamente heterogêneo, tanto molecularmente quanto clinico-morfologicamente, porém a alta expressão de Ki-67, *VIM* e de proteínas da família das lamininas, além de ganhos nas regiões 9p24-p21, 10p15-p13, 12p13, por exemplo, podem ser consideradas típicos deste subtipo. O triplo negativo se manifesta principalmente em CDIs “sem tipo específico”, mas também em carcinomas medulares, apócrinos e metaplásicos.

Além disso, tumores categorizados neste subtipo são mais prevalentes em pacientes jovens e portadoras de mutações no gene *BRCA1*, possuindo alto grau histológico e de agressividade, acometimento contralateral e de linfonodos axilares, altos índices metastáticos e de recaída, e pior prognóstico e sobrevida, conseqüentemente. Devido a sua heterogeneidade e falta de alvos para terapia, este subtipo representa um grande desafio para a medicina, sendo tratado principalmente por meio de cirurgia, quimioterapia à base de platina e de inibidores de *PARP* (*Poly ADP-ribose polymerase*).

Apesar de a classificação imunoistoquímica ser eficiente, a necessidade de meios mais exatos de estratificação de pacientes e de determinação de prognóstico, além da necessidade de uma compreensão mais completa da biologia destes tumores, faz com que novos estudos sejam necessários, e justifica a criação de novos ensaios multigênicos (DAI et al., 2015), como o *Oncotype DX* (McVEIGH et al., 2017), *Prosigna PAM50* (WALLDEN, 2015) e *MammaPrint* (SOLIMAN et al., 2020). Estes três ensaios são atualmente utilizados para prever, por exemplo, quais pacientes mais se beneficiariam com terapias neoadjuvantes; o risco de recorrência e metástase; estágio do tumor e, mais especificamente o *PAM50*, a classificação molecular dos tumores.

2.2 ABORDAGENS PARA O ESTUDO DO CÂNCER DE MAMA

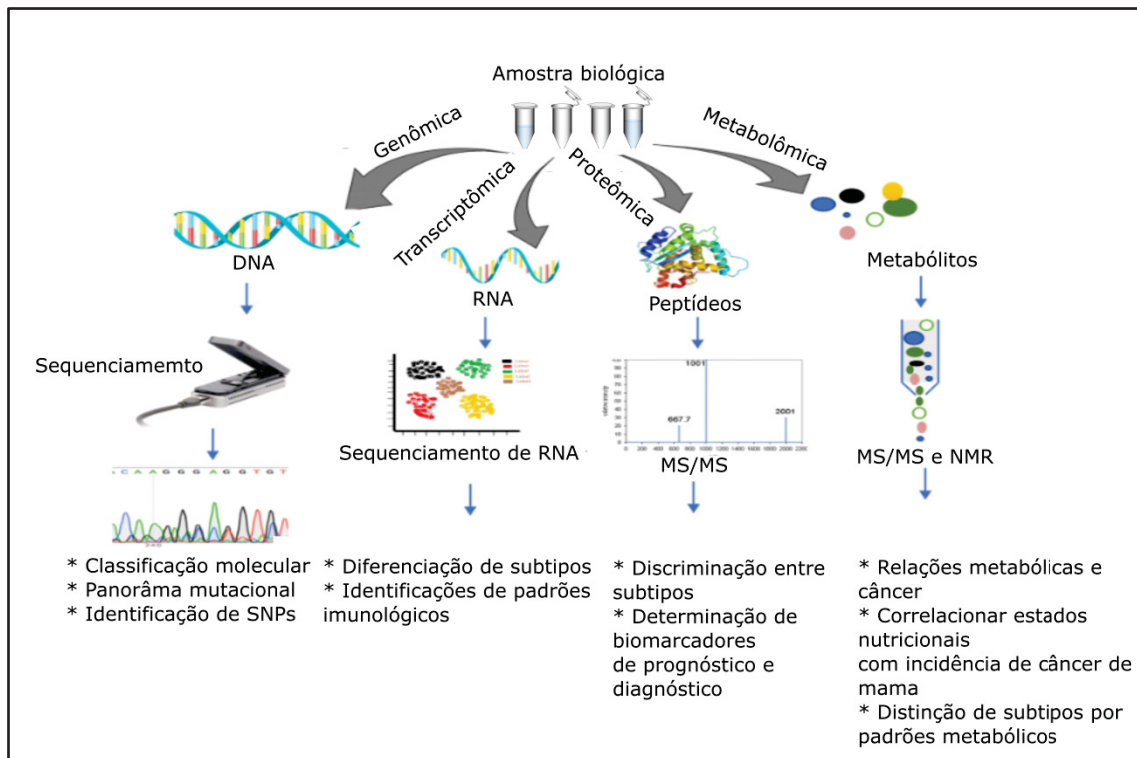
O grande risco à saúde que o câncer de mama representa para a maioria dos países do globo pressiona a comunidade científica e médica para o caminho

do diagnóstico precoce, da melhor gestão de pacientes, de melhorias nos tratamentos e da redução da mortalidade. Também devido a esta pressão, abordagens antiquadas de pesquisa têm sido deixadas de lado, e a biologia de sistemas, baseadas nas análises de redes e abordagens “ômicas”, vem ganhando força e popularidade no estudo de doenças complexas, como o câncer de mama (BARH et al., 2014).

As ciências ômicas são tecnologias utilizadas para quantificar e identificar componentes celulares em larga escala (PARSONS & FRANCAVILLA, 2020). Com o advento das tecnológicas ômicas, novos marcadores de prognóstico e diagnóstico têm sido indicados, bem como novos alvos de terapia e tratamentos personalizados. Contudo, devido a grande complexidade do câncer – incluso o câncer de mama - e a diversos fatores associados com a doença, as taxas de incidência e mortalidade do câncer de mama têm demonstrado uma ascensão significativa.

Assim, as ômicas ainda têm muito a contribuir para a compreensão a um nível mais profundo da biologia do câncer de mama, e para que possamos obter marcadores de diagnóstico e preditivos de prognóstico, e novos alvos terapêuticos dentro de um padrão-ouro ideal (BARH et al., 2014) (FIGURA 5).

FIGURA 5 – PRINCIPAIS LINHAS DE PESQUISA, DENTRE AS NOVAS CIÊNCIAS ÔMICAS, UTILIZADAS PARA A INVESTIGAÇÃO DO CÂNCER DE MAMA.



Fonte: Adaptado de DALAL et al., 2020.

(Legenda: MS = Espectrometria de massas. NMR = Ressonância nuclear magnética)

Dentre as abordagens ômicas mais utilizadas para o estudo do câncer de mama estão a genômica, fundamentalmente baseada na tecnologia de sequenciamento de nova geração, que permitiu que os tumores mamários fossem classificados a nível molecular (PEROU & SORLIE, 2000), contribuiu para a investigação do panorama mutacional do câncer de mama (LANG et al., 2020) e da determinação de polimorfismos de nucleotídeos únicos (SNPs) relacionados com susceptibilidade ao câncer de mama e respostas terapêuticas (FUNG et al., 2018); A transcriptômica, que se utiliza de micro arranjos/RNAseq para determinar a sequência de RNAs e a quantificar estes transcritos, utilizada, por exemplo, para classificar linhagens celulares e amostras de pacientes em subtipos (NEVE et al., 2006; JIANG et al., 2019) e para explorar as características imunológicas do tecido mamário (AZIZI et al., 2020); A metabolômica, que identifica e quantifica metabólitos endógenos de fluidos corporais, e que contribuiu grandemente para relacionar alterações metabólicas com a incidência de câncer de mama (JERBY et al., 2012), para diferenciar

aspectos metabólicos entre os subtipos (HAUKAAS et al., 2016) e correlacionar estados nutricionais com o risco de desenvolvimento da neoplasia mamária (PLAYDON et al., 2017) e a proteômica e interatômica, que serão abordadas no restante deste trabalho.

2.2.1 A proteômica, interatômica e suas abordagens

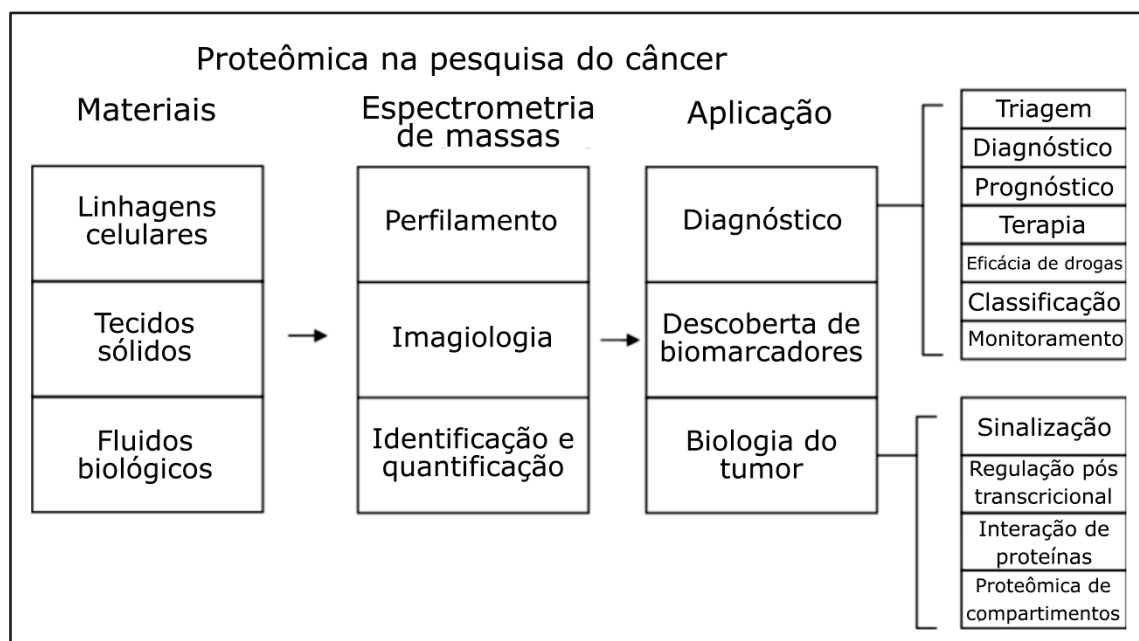
A proteômica compreende o conjunto de métodos e tecnologias utilizadas para a sistemática análise de peptídeos e proteínas em larga escala. Assim sendo, o proteoma – conjunto de todas as proteínas expressas por determinado tecido, em dada condição fisiológica – é o objeto de estudo desta ciência ômica (JAIN & FRANCS, 2002).

Os ácidos nucleicos presentes em determinado tecido são altamente relevantes como alvos de estudo no contexto do câncer de mama, porém estudar apenas a sequência que codifica uma proteína não é suficiente para compreender a relevância biológica e funcional da proteína para a biologia do tumor. Estudos baseados em mRNAs, por exemplo, podem quantificar a quantidade relativa de transcrito e prever a expressão da proteína, porém sabe-se que os dados de expressão de mRNA nem sempre se correlacionam fielmente com a realidade proteômica da célula (PARSONS & FRANCAVILLA, 2020). Deste modo, alterações transcriptômicas nem sempre refletem as diferenças biológicas entre tumores e tecido não tumoral, ou mesmo as diferenças biológicas entre os subtipos do câncer de mama.

Por outro lado, as proteínas são as moléculas que participam mais ativamente dos processos biológicos, além de serem as principais bioefetoras do fenótipo maligno, atuando diretamente em toda a evolução neoplásica. Desta forma, estudar o perfil proteico do câncer de mama – expressão, função, estrutura, interações e regulação - não apenas contribui para uma compreensão mais aprofundada da biologia, mecanismos intrínsecos, microambiente e demais particularidades do processo tumorigênico, mas também é uma necessidade clínica: As proteínas são plausíveis e acessíveis como alvos terapêuticos, possuindo também alto valor para determinação da existência de processos e condições anormais no tecido, sendo utilizadas para detecção precoce da doença, como biomarcadores de diagnóstico e prognóstico e para seleção da

melhor estratégia de tratamento para cada paciente (MACKLIN et al., 2020). Dentre as estratégias mais utilizadas para a obtenção, identificação e quantificação do proteoma, está a espectrometria de massas (MS). A MS é uma técnica de alta resolução utilizada para a avaliação do proteoma total contido em misturas complexas provenientes de tecido sólido, linhagens celulares ou fluidos corporais, sendo capaz de analisar profundamente as alterações proteicas em diferentes estados clínico-fisiológicos (PIETRO et al., 2014). Neste trabalho, a espectrometria de massas foi utilizada como metodologia base para identificação e quantificação do proteoma total de nossas amostras biológicas.

FIGURA 6 – POSSIBILIDADES DE ABORDAGENS E APLICAÇÕES DA PROTEÔMICA BASEADA EM MS NA PESQUISA DO CÂNCER.



Fonte: Adaptado de MERWE e colaboradores (2007).

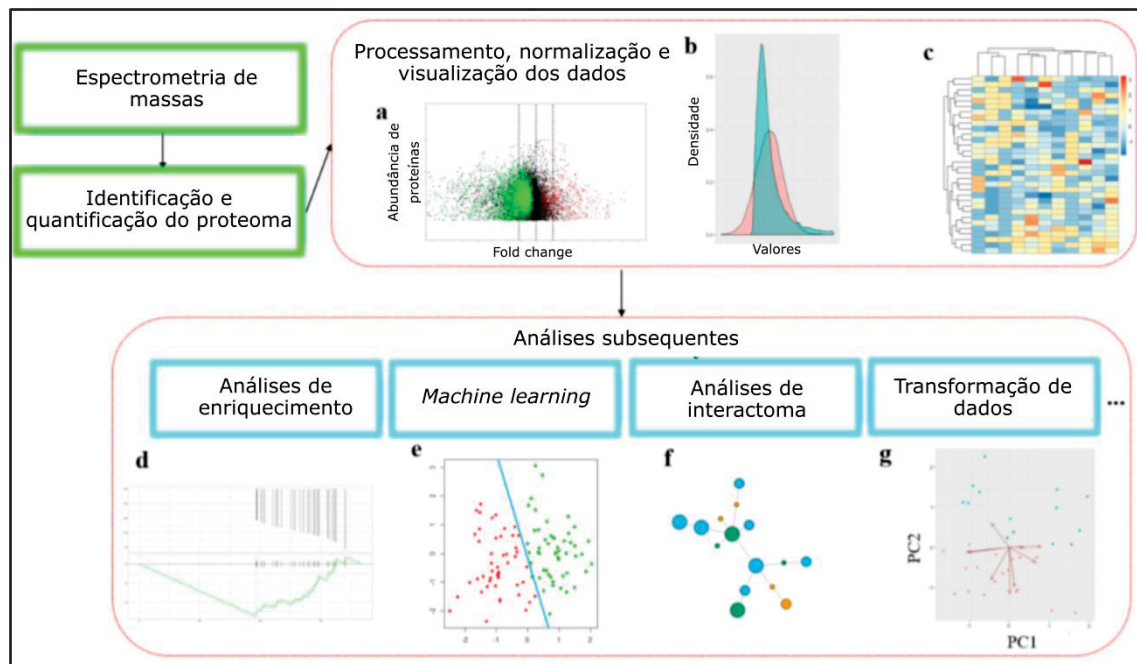
Outra ômica comumente associada à proteômica é a interatômica. A interatômica possui como objeto de estudo as interações genéticas e físicas proteína-proteína e também de proteínas com outras moléculas, como ácidos nucleicos e pequenas moléculas (TWYMAN, 2012). As informações contidas dentro do padrão de interações do proteoma podem contribuir grandemente para a compreensão do funcionamento individual de uma proteína, além de seu funcionamento dentro de vias, redes e complexos multi-proteicos (LUCK et al., 2017).

2.2.2 Análise e interpretação de dados proteômicos

As proteínas são moléculas essenciais para o funcionamento da maquinaria biológica da célula, podendo atuar como fosforilases, fosfatases, polimerases, nucleases, proteases, fatores de transcrição entre outros, desempenhando funções variadas, tais como mensageiras e sinalizadoras, reguladoras de pH, transportadoras, catalizadoras de reações intra e extracelulares, reguladoras da expressão de variadas moléculas e demais funções celulares (ANDRADE et al., 2016). De fato, alterações na regulação, interação e expressão proteica estão entre os fatores mais proeminentes para o desenvolvimento de uma doença como o câncer de mama. Assim, a compreensão destas alterações de proteoma não apenas colabora para a compreensão da biologia do câncer de mama, mas também é um objetivo clínico dos mais importantes (LUALDI & FASANO, 2018).

A proteômica baseada em espectrometria de massas contribui para a pesquisa do câncer de mama ao identificar e quantificar a expressão proteica de amostras biológicas, gerando uma imensa quantidade de informações sobre a biologia celular. Esta enorme quantidade de dados disponibilizados pela proteômica inspirou a criação e cooptação de diversas abordagens e métodos de análise bioinformática para facilitar na compreensão e descoberta das relações entre mecanismos proteicos - regulatórios e funcionais - com comportamentos biológicos alterados, por exemplo, durante o desenvolvimento e progressão do câncer (CHEN et al., 2020). Estes dados, quando corretamente analisados, nos permitem compreender o complexo e dinâmico sistema proteico do câncer como uma unidade integrada. A figura 7 ilustra esquematicamente uma *pipeline* de análise de dados de proteômica baseada em espectrometria de massas, juntamente com as abordagens bioinformáticas mais utilizadas.

FIGURA 7 – PIPELINE CLÁSSICA PARA ANÁLISES BIOINFORMÁTICAS DE DADOS PROVENIENTES DE ESTUDOS DE PROTEÔMICA BASEADA EM ESPECTROMETRIA DE MASSAS



Fonte: Adaptado de CHEN e colaboradores (2020).

(Legenda: (a) Análise de expressão diferencial. (b) Normalização dos dados. (c) Visualização dos padrões de expressão proteica via *heatmaps*. (d) Análises de enriquecimento de vias e processos biológicos para diferentes conjuntos de proteínas diferencialmente expressas. (e) Discriminação de amostras e/ou condições clínicas baseada em *machine learning*. (f) Análises de interações proteína-proteína. (g) Transformação e seleção de proteínas com diferença de expressão mais acentuada entre as amostras e/ou condições comparadas.)

2.2.2.1 Identificação e quantificação de proteomas através de espectrometria de massas.

A identificação e quantificação do proteoma total das amostras biológicas acontece através da passagem dos extratos proteicos via cromatógrafo e espectrômetro de massas. Após a fragmentação e ionização dos peptídeos através do espectrômetro de massas, os espectros proteicos são obtidos e a sequência das proteínas é identificada. A abordagem mais utilizada para identificação destas proteínas se baseia na comparação das sequências obtidas com sequência de peptídeos já identificados e disponíveis em banco de dados. Para cada comparação calcula-se uma pontuação, que será mais alta quanto maior a correspondência entre as sequências. Assim, a sequência da base de

dados que receber a maior pontuação de correspondência será utilizada para identificação da proteína em questão (CHEN et al., 2020).

O próximo passo é a quantificação da abundância de cada proteína nas amostras biológicas, que pode ocorrer através de duas abordagens: A quantificação por marcação (isobárica ou por isótopos) e a quantificação livre de marcação. Na abordagem por marcação, as amostras diferencialmente marcadas e então agrupadas e quantificadas em uma mesma 'corrida' de MS (ANAND et al., 2016). Já na abordagem livre de marcação, a quantificação de cada amostra é realizada separadamente em múltiplas passagens pelo espectrômetro, exigindo métodos de normalização para ajustar a variação decorrente (ANAND et al., 2016).

Após a quantificação e identificação do proteoma das amostras biológicas de interesse, os dados deverão passar por diversas etapas de processamento, incluindo etapas de remoção de peptídeos identificados como contaminantes, de identificações duvidosas e amostras *outliers*; normalização e logaritmização dos valores de expressão e demais tratamentos que se julgarem necessários (LUALDI & FASANO, 2018). Uma vez processados, os dados estarão prontos para as subseqüentes análises estatísticas, a partir das quais o significado biológico e clínico deste proteoma poderá ser acessado.

2.2.2.2 Análises de enriquecimento e anotação funcional

Um dos maiores desafios da abordagem proteômica baseada em MS é a interpretação biológica dos dados gerados. Uma das opções de metodologias disponíveis para a interpretação em larga escala de dados proteômicos é a análise de enriquecimento funcional (CHANG et al., 2017). A abordagem de enriquecimento funcional se baseia na utilização de aproximações estatísticas para investigar anotações funcionais significativamente enriquecidas dentro de um conjunto de proteínas, buscando, por exemplo, associações entre listas de proteínas com os termos de ontologia gênica (*Gene ontology terms*), vias metabólicas, de sinalização e de regulação transcricional, ou ainda vias/processos relacionados com os *hallmarks* do câncer (WU et al., 2014).

Através destas análises é possível inferir não apenas a significância biológica de um conjunto de proteínas, mas também determinar quais as diferenças funcionais entre diferentes estados clínico-fisiológicos. Por exemplo, pode-se determinar as diferenças funcionais entre tumores e tecidos não tumorais através do enriquecimento de um conjunto de proteínas com alta ou baixa expressão no tecido tumorais, ou ainda as diferenças biológicas entre diferentes tipos ou subtipos de câncer através de uma análise funcional de conjuntos de proteínas com expressão diferencial entre estes tipos/subtipos tumorais.

2.2.2.3 Análises de predição de interatoma.

Cada vez mais se percebe que a análise das interações proteína-proteína do ambiente celular neoplásico é relevante para a compreensão do funcionamento proteico desregulado no câncer, já que nenhuma proteína consegue exercer suas funções isoladamente. Devido às conexões entre proteínas, mudanças de expressão em apenas uma proteína central podem afetar o funcionamento de toda uma rede, a qual, uma vez perturbada, afetará processos biológicos que podem levar à aquisição de algum dos *hallmarks* por células tumorais (RICHARDS et al., 2020). Da mesma forma, esta premissa promete a possibilidade de, através de análises de predição de interatoma, se aperfeiçoar a seleção de proteínas para terapias alvo, permitindo a seleção de proteínas-chave dentro do contexto interatômico que, uma vez atingidas, influenciarão todo o contexto de funcionamento da rede (KOVACS et al., 2019).

Uma abordagem que nos permite explorar estas relações proteína-proteína é a de predição computacional de redes de interação. Estas redes podem ser preditas baseando-se em algoritmos que levam em conta medidas de co-expressão e co-localização entre proteínas ou na análise de motivos/sítios de interação, por exemplo. Através da determinação destas redes, pode-se explorar com maior precisão a relação entre enzimas e seus substratos, quinases e seus alvos, receptores e seus ligantes, funcionamento de complexos multi-proteicos e toda uma miríade de interações proteína-proteína que interferem no contexto do funcionamento celular e do tumor (CHEN et al., 2020).

2.2.2.4 Discriminação entre condições biológicas via abordagens de aprendizado de máquina (*machine learning*)

Outro grande desafio de interpretação dos dados derivados da proteômica baseada em MS é a seleção de proteínas com potencial para serem biomarcadores e/ou moléculas discriminativas de diferentes condições biológicas, podendo-se citar a seleção de marcadores que discriminem pacientes com baixa ou alta expectativa de sobrevida, de diagnóstico de presença ou ausência de neoplasia, ou até mesmo para a discriminação de subtipos do câncer de mama. Como determinar, dentro de um conjunto numeroso de proteínas, qual o menor número necessário, e quais as proteínas com maior potencial discriminatório? As abordagens de aprendizado de máquina (*machine learning*) podem ser uma opção.

O *machine learning* é um conjunto de metodologias baseadas na determinação de padrões e relações entre grupos de dados – por exemplo, de expressão proteica – e estado fisiológico, como doente/não doente, ou estimativas clínicas, como baixo tempo de sobrevida/alto tempo de sobrevida. Mais especificamente, o *machine learning* atua através da criação de um modelo de predição a partir da coleta de amostras de pacientes com estados clínicos conhecidos: No caso de modelos proteicos, o resultado final apresentará um conjunto de proteínas que possuiu um padrão de expressão relacionado diretamente com os estados clínico-fisiológico de interesse. Idealmente, este conjunto de proteínas poderá ser usado futuramente para a discriminação de pacientes com estados patológicos ainda indeterminados ou para reforçar sistemas de classificação atuais (KOUROU et al., 2015).

Assim, o principal objetivo do *machine learning* é de classificar condições clínico-patológicas com base em dados moleculares. Cada modelo preditivo possui parâmetros que determinam o quão eficaz o modelo é para classificar as condições de interesse, como valores de sensibilidade, especificidade, acurácia e área sob a curva (AUC), sendo a utilidade clínica do modelo avaliada conforme o desempenho obtido pela medição destes parâmetros (IQBAL et al., 2021).

2.2.3 As contribuições da proteômica e interatômica para o estudo do câncer de mama.

Apesar do significativo avanço nas pesquisas sobre o câncer de mama, não só a complexidade de sua biologia ainda é apenas parcialmente conhecida, como a rotina clínica ainda necessita de biomarcadores de diagnóstico, prognóstico e estratificação de pacientes, além de novos alvos terapêuticos. Neste contexto, a proteômica e interatômica têm contribuído significativamente.

Diferentes abordagens proteômicas têm sido aplicadas visando compreender as diferenças biológicas entre os tecidos tumorais mamários em relação aos tecidos não tumorais e entre tumores de diferentes subtipos. Gomig e colaboradores (2019) utilizaram a abordagem MS livre de marcação para comparar tecidos sólidos tumorais de câncer de mama com tecidos não tumorais adjacentes e contralaterais, explorando o impacto das alterações de expressão proteica do tecido tumoral em vias e processos biológicos relacionados com a tumorigênese mamária. As principais alterações de vias canônicas influenciadas pelas proteínas diferencialmente expressas do tecido tumoral foram a inibição da via de *LXR/RXR*, relacionada com a homeostase de colesterol e ácidos graxos e com o controle do metabolismo de glucose; a via de eNOS, relacionada com a produção e/ou metabolização de espécies reativas de oxigênio, tomando parte em processos de inflamação e estimulação do crescimento tumoral, e as vias de sinalização mediadas por eIF2 e sirtuínas, que participam de alterações no microambiente celular relacionadas com hipóxia, angiogênese e inflamação. Outros trabalhos baseados em MS também investigaram o panorama proteico em comparações entre tecido tumoral e não tumoral mamário, porém considerando relações adicionais, como do estágio do tumor (AL-WAJEEH et al., 2012) e metástase em linfonodos (MILIOLI et al., 2015), ou ainda utilizando outros tipos de amostras biológicas, como o líquido de aspiração de mamilo (BRUNORO et al., 2019), soro (LEE et al., 2020) ou exossomos (RISHA et al., 2020). Conjuntamente, estes trabalhos contribuem para a compreensão da relação entre alterações de expressão de proteínas específicas, mecanismos celulares desregulados por estas proteínas, e os diferentes fenótipos malignos observados.

Também utilizando a abordagem de espectrometria de massas livre de marcação, Panis e colaboradores (2014) investigaram o proteoma diferencial entre tumores de diferentes subtipos de câncer de mama, caracterizando diferenças biológicas entre os subtipos através de análises de enriquecimento funcional. Através de suas análises, os autores foram capazes de determinar desregulações biológicas comuns a todos os subtipos, como remodelagem de citoesqueleto e perda de adesão celular, indicando que estas alterações podem ser mediadas por variações de expressão proteica comuns a todos os subtipos, como a baixa expressão da proteína codificada pelo gene *FAK1*, mas também que estes processos podem ser mediados por diferentes proteínas dependendo do subtipo do tumor. Por exemplo, a desregulação da adesão celular nos subtipos luminal A e HER2+ enriquecido foi relacionada principalmente a proteínas da família das integrinas, enquanto as quimiocinas foram as mais relevantes para o conjunto dos subtipos luminal A, HER2+ enriquecido e triplo negativo, com o subtipo luminal B relacionado principalmente à adesão celular desregulada por caderinas. Porém, outros processos, como apoptose, regulação do ciclo celular e coagulação sanguínea, apresentaram alterações que possuíam padrões completamente diferentes dependendo do subtipo em que eram observadas. O subtipo luminal B apresentou uma relação estreita com desregulação do ciclo celular, com alta expressão de *ATM* e baixa expressão de *ATR* em relação aos demais subtipos, enquanto o subtipo HER2+ enriquecido apresentou uma relação proeminente com apoptose, mais especificamente com apoptose induzida por danos do DNA, e desregulação de proteínas específicas, como as codificadas pelos genes *ATM*, *ATR*, *FANCD2* e *BLM*. O subtipo triplo negativo apresentou uma relação muito mais estreita com coagulação sanguínea do que a encontrada em outros subtipos, com alta expressão dos fibrinogênios alfa, beta e gama, além de outras proteínas como fibrinas, plasminas e plasminogênios. Além de processos e vias com padrões de desregulação específicos de cada subtipo, alterações de interatoma também foram descritas, evidenciando que alterações de padrões de interação também se relacionam com os diferentes comportamentos clínicos e biológicos apresentados pelos subtipos de câncer.

Outras contribuições da proteômica para o estudo do câncer de mama consistem na determinação de proteínas que possuem padrões característicos de expressão no tecido mamário e que possuem a capacidade de atuar como biomarcadores de presença e ausência de neoplasia (e outros aspectos clínicos), e como biomarcadores com potencial para discriminação e estratificação de tumores em diferentes subtipos. Apesar da utilização de características patológicas como morfologia, tamanho e grau do tumor, presença de metástase em linfonodos e expressão de ER, PR, HER2 e Ki-67 ser de grande valor para tomada de decisão e estratificação de pacientes, novos biomarcadores são necessários para aperfeiçoar estes processos e possibilitar diagnósticos mais precisos e precoces. Neste sentido, diversos autores têm contribuído com a indicação de biomarcadores proteicos, tanto em tecido tumoral sólido, quanto em biopsia líquida.

Chung e colaboradores (2013), por exemplo, utilizaram a proteômica baseada em MS para comparar o proteoma de tecidos mamários tumorais e tecidos adjacentes, determinando um conjunto de duas proteínas (ubiquitina e *S100*) que, juntas, se associavam com alto grau histológico e tamanho tumoral, apresentando potencial para serem utilizadas como marcadoras para diagnóstico precoce. Ösz e colaboradores (2020) também se basearam em dados de MS e imunohistoquímica para determinar proteínas que pudessem ser marcadoras de diagnóstico para câncer de mama e preditoras de sobrevida, concluindo, por exemplo, que a alta expressão das proteínas E-caderina e Bcl-2 se relaciona com melhor expectativa de sobrevida, e indicando um conjunto de outras proteínas que poderiam ser utilizadas para diagnóstico da presença de neoplasia mamária. Em relação à biopsia líquida, a revisão de Li e colaboradores (2020) explora a ocorrência de alguns estudos nesta área, com a existência de painéis de proteínas de soro, urina e outros fluidos que podem ser potencialmente utilizados para o diagnóstico do câncer de mama.

No campo da discriminação dos subtipos de câncer de mama, alguns trabalhos (TYANOVA et al., 2016; YANOVICH et al., 2018; BOUCHAL et al., 2019; TERKELSEN et al., 2020) têm indicado que padrões de expressão e interação proteica podem ser associados com diferentes subtipos, sendo que alguns destes conjuntos proteicos apresentam potencial para serem utilizados

na clínica, no sentido de aperfeiçoar a classificação dos tumores em cada um dos subtipos. Tyanova e colaboradores (2016), por exemplo, utilizaram a estratégia de machine learning para determinar conjuntos de proteínas que têm padrões de expressão capazes de discriminar entre os subtipos luminais (genes *MAPK3* e *EEF1G*), HER2+ enriquecido (HER2 e Grb7) e triplo negativo (*MCM5*, *STMN1*, *GLS*, *RCL1*, *C9ORF114*, *ENO1* e outras).

Em geral, as pesquisas proteômicas recentes têm demonstrado que é possível traçar correlações entre padrões de expressão proteica com diferenças biológicas entre tumores mamários e de classificação em subtipos, indicando que de fato as proteínas são ótimas opções como biomarcadores de diagnóstico e prognóstico. Contudo, devido à grande variação da expressão proteica encontrada entre populações, e mesmo dentro das populações, ainda são necessários mais estudos que busquem a determinação de painéis proteicos, não apenas visando a mais profunda compreensão da biologia do câncer de mama e seus subtipos, mas também o diagnóstico mais precoce e uma discriminação mais robusta entre seus subtipos.

3 DESCRIÇÃO DOS CAPÍTULOS

Capítulo 1 – Este capítulo compreende a etapa principal deste trabalho de mestrado. Nele são contemplados os itens 4 a 11 (objetivos às referências), tratando da obtenção do proteoma total de tumores mamários de diferentes subtipos, e da análise dos proteomas visando compreender como os proteomas diferenciais entre os subtipos de câncer de mama estão relacionados com as diferenças biológicas entre os mesmos. As análises proteômicas permitiram apontar, além disso, painéis proteicos com potencial para discriminação entre os tumores de cada subtipo.

Capítulo 2 – Este capítulo é derivado dos resultados descritos no capítulo 1, e é apresentado no formato para publicação em revista científica a ser definida. Durante a elaboração das análises do primeiro capítulo, uma família específica de proteínas – a família das subunidades ribossomais (RPs) – se destacou devido ao padrão de expressão e associações com funções moleculares altamente relevantes no câncer: Tendo isto em vista, o capítulo 2 aprofunda a análise das *large ribosomal subunits* (RPLs) e *small ribosomal subunits* (RPSs) adicionando aos dados proteômicos informações sobre expressão a nível transcriptômico (mRNA) obtidas da base de dados TCGA. Mais especificamente, o capítulo 2 consiste de uma análise abrangente acerca da expressão das RPs, incluindo; seu potencial de utilização como marcadoras de diagnóstico e prognóstico; o panorama mutacional e funcional das RPs, além dos principais mecanismos regulatórios da expressão das RPs. Neste capítulo são incluídos dados proteômicos obtidos de tecidos não tumorais, obtidos conforme a metodologia descrita no capítulo 1.

4 CAPÍTULO 1

4.1 RESUMO CAPÍTULO 1

O câncer de mama é uma doença com altos índices de incidência, mortalidade e impactos sociais e econômicos. Os tumores mamários são classificados em subtipos, visando direcionar o tratamento e prever o prognóstico dos pacientes, porém, a classificação atual, baseada em quatro marcadores imunoistoquímicos, não é suficiente para uma classificação satisfatória destes tumores. Além disso, as características biológicas que geram as diferenças entre tumores não são completamente conhecidas. Neste sentido, as tecnologias proteômicas e interatômicas têm grande relevância, tanto na investigação das diferenças biológicas entre subtipos através da identificação de proteínas diferencialmente expressas, as quais influenciam vias e processos relacionados com as diferenças clínicas e biológicas dos subtipos, quanto na identificação de marcadores proteicos que possam complementar e otimizar a classificação e discriminação entre subtipos. Neste estudo, métodos proteômicos baseados em espectrometria de massas LC-MS/MS livre de marcação foram utilizados para obtenção e comparação do proteoma total de tumores mamários de diferentes subtipos e, através da aplicação de análises computacionais de enriquecimento funcional e de predição de interatoma, a relevância biológica de proteínas diferencialmente expressas foi evidenciada. Além disso, a tecnologia de *machine learning* baseada em *supported vector machines* foi aplicada para a obtenção de painéis proteicos com potencial para servirem como painéis de discriminação e de classificação de tumores mamários em subtipos. Proteínas diferencialmente expressas (DEPs) puderam ser identificadas entre todos os subtipos, identificando oncogenes e supressores de tumor que possuem padrões de expressão que variam segundo a subdivisão entre subtipos. Através das análises de enriquecimento, os contextos biológicos nos quais se incluem estas DEPs puderam ser explorados, sendo identificados processos e vias biológicas que possuem níveis de atuação e ativação/inativação que diferem segundo o subtipo do tumor, incluindo alterações, por exemplo, nas vias de endocitose mediada por clatrina, metabolismo de glicose e carboidratos, citoproteção por *HMOX1* e regulação da estabilidade e atividade de *PTEN*. Os painéis proteicos aqui propostos obtiveram bom desempenho, alcançando taxas de falso-positivos sempre menores que 10%, e sensibilidade e especificidade sempre maiores que 75%. No conjunto, este trabalho aponta padrões de expressão proteica que podem ser associados com subtipos específicos, e acessa a significância destas DEPs em vias e processos biológicos relacionados com as diferenças biológicas e clínicas apresentadas por tumores mamários de diferentes subtipos. Além disso, painéis proteicos de discriminação entre subtipos foram propostos, os quais, apesar de necessitarem de mais estudos comprobatórios, apresentam bom desempenho e robustez para a classificação de tumores em diferentes subtipos.

Palavras-chave: Câncer de mama; subtipos; proteômica; painéis proteicos.

5 OBJETIVOS

5.1 OBJETIVO GERAL

Obter e analisar o proteoma de tumores dos quatro subtipos do câncer de mama (Luminal A, Luminal B, HER2+ enriquecido e triplo negativo), investigando as diferenças de expressão seus proteomas no contexto das diferenças biológicas que os mesmos apresentam, e determinando painéis proteicos com potencial para discriminar os diferentes subtipos da doença.

5.2 OBJETIVOS ESPECÍFICOS

- Quantificar e identificar o proteoma de tecidos sólidos tumorais mamários de quatro subtipos do câncer de mama (Luminal A, Luminal B, HER2+ enriquecido e triplo negativo).
- Determinar proteínas que possuam diferenças significativas de expressão na comparação entre as amostras de cada subtipo.
- Identificar, através de análises de enriquecimento funcional, alterações em vias e processo biológicos que possuam padrões de ocorrência que sejam específicos e/ou variem conforme o subtipo analisado.
- Prever e descrever, computacionalmente, redes de interação proteína-proteína.
- Indicar, através de análises pautadas em *machine learning*, painéis de expressão proteica que possuam potencial para discriminar os diferentes subtipos do câncer de mama.

6 MATERIAL E MÉTODOS

6.1 OBTENÇÃO E PROCESSAMENTO DO PROTEOMA

6.1.1 Caracterização das amostras biológicas

Este estudo foi realizado como parte dos projetos de pesquisa desenvolvidos pelo Laboratório de Citogenética Humana e Oncogenética (LabCHO) do Departamento de Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná (UFPR), em parceria com o Centro de Doenças da Mama do Hospital Nossa Senhora das Graças (HNSG) e da Plataforma de

Espectrometria de Massas (RPT02H) do Instituto Carlos Chagas, Fiocruz, Paraná. Este estudo foi aprovado pelo Comitê de Ética do Setor de Ciências da Saúde da UFPR (CAAE 19870319.3.0000.0102).

As amostras biológicas utilizadas para o desenvolvimento deste estudo são amostras de tecido tumoral sólido de pacientes com câncer de mama. Tais amostras foram coletadas durante procedimentos cirúrgicos realizados no HNSG mediante assinatura do termo de consentimento livre e esclarecido. Após processamento para retirada de resquícios de vasos sanguíneos e tecido adiposo, as amostras foram fragmentadas e armazenadas em solução de *RNAlater* a -80°C até o momento de obtenção dos extratos proteicos. As informações clínicas e patológicas das pacientes e amostras biológicas são descritas na tabela 3. A classificação imunoistoquímica dos subtipos do câncer de mama foi realizada segundo os critérios de Goldhirsch e colaboradores (2013). Neste estudo, um total de 23 amostras foram utilizadas, sendo os subtipos luminais A e B e o triplo negativo sendo representados por seis amostras cada, e o subtipo HER2+ enriquecido sendo representado por cinco amostras.

TABELA 3 – CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DAS PACIENTES DOADORAS DAS AMOSTRAS BIOLÓGICAS UTILIZADAS DURANTE ESTE ESTUDO.

Código	Idade (Anos)	Classificação histológica e grau do tumor	Marcadores imunoistoquímicos				Classificação imunoistoquímica do tumor
			ER	PR	HER2	Ki-67	
T932	66	CDI – I	POS	POS	NEG	10%	Luminal A
T957	46	CDI – II	POS	POS	NEG	10%	Luminal A
T986	62	CDI – I	POS	POS	NEG	5%	Luminal A
T993	35	CDI – II	POS	POS	NEG	5%	Luminal A

T976	63	CDI – II	POS	POS	NEG	13 %	Luminal A
T972	61	CDI – II	POS	POS	NEG	10 %	Luminal A
T917	52	CDI – III	POS	POS	NEG	70 %	Luminal B
T924	70	CDI – III	POS	POS	NEG	60 %	Luminal B
T930	32	CDI – II	POS	POS	NEG	60 %	Luminal B
T949	55	CDI – II	POS	POS	NEG	25 %	Luminal B
T965	60	CDI – II	POS	POS	NEG	30 %	Luminal B
T981	60	CDI – II	POS	POS	NEG	20 %	Luminal B
T868	76	CIM – III	NE G	NE G	POS	25 %	HER2 + Enriquecido
T935	88	CD – III	NE G	NE G	POS	60 %	HER2 + Enriquecido
T967	54	CDI – III	NE G	NE G	POS	75 %	HER2 + Enriquecido
T995	59	CDI – II	NE G	NE G	POS	20 %	HER2 + Enriquecido
T997	70	CDI – II	NE G	NE G	POS	70 %	HER2 + Enriquecido
T752	40	CDI – III	NE G	NE G	NEG	38 %	Triplo negativo
T779	66	CDI – III	NE G	NE G	NEG	80 %	Triplo negativo

T952	65	CDI – III	NE G	NE G	NEG	90 %	Triplo negativo
T955	53	CDI – III	NE G	NE G	NEG	80 %	Triplo negativo
T956	67	CDI – III	NE G	NE G	NEG	80 %	Triplo negativo
T960	31	CDI – III	NE G	NE G	NEG	80 %	Triplo negativo

(Legenda – CDI = Carcinoma ductal invasivo. CIM = Carcinoma invasivo misto. CD = Carcinoma Ductal. ER = Receptor de estrogênio. PR = Receptor de progesterona. HER2 = Receptor de crescimento epidérmico 2. Ki-67 = Antígeno marcador de proliferação celular)

6.1.2 Obtenção dos extratos proteicos

A obtenção dos extratos proteicos se deu através de lise química (baseada no detergente dodecil sulfato de sódio, SDS) e de lise mecânica (homogeneização do tecido via *TissueLyser*), segundo protocolo compatível com análises de LC-MS/MS, adaptado de Gomig e colaboradores (2019).

Cada tumor analisado foi pesado, armazenado em tubo de 2 mL (com espessura reforçada, visando as próximas etapas da extração) e, para cada 10 mg de tecido utilizado, 100 µl de tampão de lise (SDS 4%, Tris-HCl 0,1M; pH 7,5, DTT 0,1M) foram acrescentados. Logo após, cada tubo recebeu uma *bead* magnética de 3,5 mm de raio, e os tecidos foram homogeneizados através do disruptor de tecidos *TissueLyser* (Qiagen), em três ciclos de 3 minutos (frequência 30 Hz) intercalados por intervalos de 1 minuto de aquecimento à 95°C em banho seco, e 5 minutos de esfriamento em gelo, para completa homogeneização dos tecidos e solubilização das proteínas. Após os três ciclos, as amostras foram mantidas em gelo por 30 minutos para atuação do tampão de lise. Depois deste intervalo, as amostras foram submetidas ao processo de sonicação, visando otimizar a lise celular, quebrar ácidos nucleicos e diminuir a viscosidade do extrato. A sonicação se deu em três ciclos de 30 segundos (frequência 25 Hz), intercalados por pausas de 1 minuto, nas quais as amostras eram mantidas em gelo. Por fim, duas etapas de centrifugação (primeira etapa

[5 minutos, 12.000 rpm a 4°C], segunda etapa [10 minutos, 12.000 rpm a 4°C]) foram realizadas com o objetivo de precipitar restos celulares e moléculas indesejadas, de modo a manter no sobrenadante apenas o conteúdo proteico, o qual foi coletado e estocado a -80°C até o momento da quantificação.

6.1.3 Quantificação dos extratos pelo método do triptofano.

O método do triptofano (WISNIEWSKI & GOUGAZ, 2015) foi utilizado para quantificação da concentração total de proteínas presentes em cada extrato proteico. Esta técnica, diferentemente de outras técnicas de quantificação, como a de Bradford (1976), é completamente compatível com detergentes como o SDS, descartando a necessidade de métodos de filtragem para a retirada do mesmo.

Vários estudos demonstraram que o conteúdo de triptofano presente em extratos obtidos a partir de SDS pode ser mensurado através da medição da fluorescência emitida pelo triptofano de proteínas completamente desnaturadas (WISNIEWSKI & GOUGAZ, 2015). Segundo os autores, o proteoma humano possui 0,0117g de triptofano a cada 1 g de proteína: Baseando-se nesta proporção, o método do triptofano utiliza a montagem de uma curva padrão a partir de medidas de fluorescência derivadas de concentrações conhecidas de triptofano para uma posterior comparação das medidas de fluorescência de cada amostra com esta curva padrão, utilizando a proporção massa triptofano/massa proteína para determinar a quantidade total de proteína por amostra. A fluorescência emitida foi analisada em um aparelho TECAN, com parâmetros de emissão 350 nm, excitação 295 nm, *wavelength 1* e ganho 100.

6.1.4 Separação eletroforética dos extratos proteicos

Visando não apenas uma breve separação das proteínas de cada extrato, mas também a remoção de contaminantes (como moléculas de DNA, detergente, sais e outros compostos não proteicos) que possam estar presentes no extrato, géis unidimensionais de poliacrilamida (1D-SDS-PAGE) foram utilizados, contendo um gel inferior para separação das proteínas, e um gel superior para empilhamento. O gel de separação foi elaborado a 14% (acrilamida 0,9%, SDS 10%, Tris-HCl 400mM, persulfato de amônio [APS] 1%, tetrametiletilenodiamina [TEMED] 0.08%) enquanto o gel de empilhamento foi

elaborado a 5% (acrilamida 0,5%, Tris-HCl 120mM, SDS 10%, APS 1% e TEMED 0,08%).

Para cada amostra, um volume de extrato contendo 25 µg de proteína foi misturado ao tampão de amostra (Tris-HCl 40mM, SDS 1%, beta-mercaptoetanol 2,5%, Glicerol 6%, Azul de bromofenol 0,005%) e, após aquecimento por 5 minutos em banho seco a 95°C, a mistura extrato + tampão de amostra foi aplicada em gel, sendo a corrida eletroforética realizada a 100 mV por tempo suficiente para que toda a amostra saísse do gel de empilhamento e adentrasse aproximadamente 3 centímetros dentro do gel de separação. Além disso, outra corrida eletroforética independente foi realizada nos mesmos padrões, porém permitindo que a amostras corresse até o fim do gel, no intuito de permitir a posterior visualização de todo o padrão de bandas exibido em cada amostra.

Posteriormente a ambas as corridas, os géis foram corados em solução de Azul de Coomassie (R-250) 0,1% por 20 minutos, sendo em seguida deixados por cerca de 12-14 horas (*overnight*) em solução de descoloração (metanol 30% e ácido acético 10%). No dia seguinte, os géis foram lavados com água ultrapura. O gel de corrida longa permitiu a visualização de todo o padrão de bandas das amostras, enquanto o gel de corrida curta foi processado para as próximas etapas.

Após as lavagens, os géis de corrida curta foram dispostos em placas de vidro limpas e as bandas contendo as proteínas foram excisadas com o auxílio de um bisturi. Cada banda foi cortada em pedaços de 1 mm x 1 mm, sendo acondicionados em água ultra pura até os próximos passos de preparo para MS.

6.1.5 Digestão enzimática em gel das proteínas e preparo para LC-ESI-MS/MS.

As etapas aqui descritas compreendem a descoloração dos géis após o corte, bem como as etapas de redução, alquilação e digestão enzimática em gel das proteínas, que precedem a extração dos peptídeos do gel.

Uma vez cortados, os pedaços de gel de cada amostra foram armazenados em tubos de 1,5 mL e tratados com solução de descoloração

(bicarbonato de amônio $50\text{mmol}^{-1}\text{L}$, diluído em etanol 50%), permanecendo em agitação por no mínimo 10 minutos em temperatura ambiente, ou até que todo o coomassie fosse retirado, ficando *overnight* caso a coloração fosse muito forte. Em seguida, os géis foram desidratados com etanol absoluto em duas etapas seguidas de agitação por 10 minutos cada, a 25°C , até que os pedaços de gel ficassem endurecidos e brancos. Após, os tubos foram levados à secagem no equipamento *SpeedVac* (Thermo Fisher Scientific) por 7 minutos (Pressão de vácuo 0,1).

Em seguida, o tampão de redução (solução de DTT $10\text{mmol}^{-1}\text{L}$ em ABC $50\text{mmol}^{-1}\text{L}$) foi adicionado e seguiu-se para uma incubação em rotação por 60 minutos a 56°C . Logo após, a alquilação das amostras foi realizada com a adição do tampão de alquilação (solução de Iodoacetamida $55\text{mmol}^{-1}\text{L}$ em ABC $50\text{mmol}^{-1}\text{L}$) seguida por incubação sob agitação por 45 minutos a 25°C no escuro. A etapa de digestão ocorreu logo em seguida, com a adição da solução de digestão (solução aquosa de ABC $50\text{mmol}^{-1}\text{L}$) e incubação sob agitação por 20 minutos a 25°C , seguida de outra etapa de desidratação em etanol absoluto, nos mesmos termos citados anteriormente, e então de outra incubação com o tampão de digestão seguida de desidratação, culminando com a secagem dos pedaços de gel da amostra no *SeedVac*. Uma vez secos, os pedaços de gel receberam uma solução de tripsina fresca ($12,5\text{ng}^{-1}\ \mu\text{L}$ diluída com solução ABC $50\text{mmol}^{-1}\text{L}$) e ficaram por 20 minutos a 4°C para que o gel pudesse reidratar nesta solução. Passado este tempo, o excesso de solução de tripsina foi retirado, e as amostras foram incubadas *overnight* em estufa a 37°C , imersas em tampão ABC $50\text{mmol}^{-1}\text{L}$.

No dia seguinte, as amostras foram retiradas da estufa e todo o líquido foi coletado em tubos novos de 2 mL. Em seguida, os pedaços de gel foram tratados com ACN 100% e agitados por 10 minutos a 25°C e, após o período, o sobrenadante foi coletado e adicionado ao mesmo tubo de 2 mL mencionado anteriormente. Este procedimento foi repetido duas vezes. O volume total contido no tubo de 2 mL foi diminuído até 20-30% do volume inicial através do *SpeedVac* para retirada da acetonitrila.

As próximas etapas consistem na purificação e armazenamento dos peptídeos em *StageTips* C18 (ponteiros de 200 μ L com membranas C18 acopladas no seu interior). Estas ponteiros, previamente a seu uso, devem ser ativadas com 100 μ L de metanol, passando por uma centrifugação a 1000G por 2 minutos e equilíbrio das membranas com ácido fórmico 0,1%, finalizando com uma centrifugação (1000G por 6 minutos) antes de sua utilização. Os peptídeos (secos pelo *SpeedVac*) foram suspensos em 100 μ L da solução de ácido fórmico e então transferidos para as *StageTips* C18 ativadas, passando por uma centrifugação à 1000G até os peptídeos passarem através da membrana C18. Em seguida, as *StageTips* foram lavadas duas vezes com 200 μ L da solução de ácido fórmico e centrifugadas a 100G por 6 minutos por lavagem, sendo por fim armazenadas a 4°C até o momento de sua eluição.

A eluição foi realizada previamente à submissão dos peptídeos à LC-MS/MS. As *StageTips* foram novamente lavadas com 200 μ L da solução de ácido fórmico e centrifugadas por 6 minutos à 1000G, descartando-se o líquido em seguida. A eluição de cada *StageTips* ocorreu duas vezes com 20 μ L de solução ácido fórmico 0,1% e acetonitrila 40% em novos tubos. Estes tubos foram em seguida submetidos ao *SpeedVac* para evaporação da acetonitrila, sendo que no final os peptídeos foram ressuspensos em 20 μ L de ácido fórmico 0,1%, quantificados em espectrofotômetro por absorvância em 280nm, ajustada a concentração para 0,1 μ g/ μ L⁻¹ e então transferidos para uma placa de 96 poços, a qual foi levada ao cromatógrafo.

6.2 QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA TOTAL

6.2.1 Análise dos peptídeos por LC-ESI-MS/MS

Os peptídeos referentes aos extratos proteicos de cada amostra foram analisados em nanocromatografia líquida acoplada à espectrometria de massas (LC-MS/MS) com todo o processo ocorrendo na plataforma de espectrometria de massas do Instituto Carlos Chagas (Fiocruz, PR), sob protocolo próprio da plataforma. Os instrumentos utilizados foram o Cromatógrafo líquido Ultimate 3000 RSLCnano (*Thermo Fisher Scientific*), diretamente conectado ao espectrômetro de massas Orbitrap Fusion Lumos (*Thermo Fisher Scientific*). A carbamidometilação de cisteína foi definida como modificação fixa, e a oxidação

da metionina e acetilação N-terminal proteica foram definidas como modificações variáveis. Um FDR de 1% foi aplicado independentemente para identificação de peptídeos e proteínas, com pelo menos sete aminoácidos necessários para identificação dos peptídeos. Os demais parâmetros da cromatografia e espectrometria são apresentados na tabela suplementar 1 (anexo).

6.2.2 Quantificação e análise dos espectros de massa obtidos

Os espectros de massa obtidos ao final do processo de LC-MS/MS foram analisados ainda dentro da plataforma de espectrometria de massas por nossos parceiros, seguindo protocolos próprios. Os espectros de massa foram analisados através do programa MaxQuant (v. 1.6.17), sendo a identificação das proteínas realizada através do recurso de busca no banco de dados Uniprot (7577 sequências, obtido em 11/01/2021), seguindo a opção disponível para dados oriundos de quantificação livre de marcação. Cada amostra submetida à plataforma foi aplicada em duplicata em todo o processo, sendo a opção de sobreposição entre corridas utilizada para agrupar as intensidades de quantificação livre de marcação (*Label-free quantification*; LQF) e determinar a expressão de cada proteína da amostra.

6.3 ANÁLISE E INTERPRETAÇÃO DOS DADOS

6.3.1 Processamento dos dados brutos

Os dados brutos obtidos através da plataforma MaxQuant foram trabalhados através do programa Perseus (v. 1.16.15.0), considerando-se os valores de *Label-free quantification* (LFQ) para a determinação da expressão de cada proteína nas diferentes amostras.

O processamento dos dados se iniciou com a remoção de proteínas identificadas como potenciais contaminantes, peptídeos reversos e proteínas apenas identificadas por um sítio modificado. Em seguida, os valores de LFQ foram transformados em logarítmicos de base 2 ($\log_2 X$), inserindo para cada proteína a simbologia NaN (*not a number*) nas amostras nas quais elas obtiveram quantificação nula. Com a presença de NaNs na matriz de dados, foi possível filtrar e remover as proteínas que estavam ausentes (quantificação nula) em mais de 30% das amostras. Esta filtragem diminuiu razoavelmente a

quantidade de proteínas que seguiram para futuras análises, porém ao mesmo tempo, garantiu maior robustez aos resultados. Por fim, os valores de LFQ foram normalizados através da aplicação do *Z-score* e então os valores NaN remanescentes foram distribuídos por valores seguindo uma distribuição normal (*Width 0,3* e *down shift 1.8*). Uma vez processadas, as amostras foram agrupadas segundo seus subtipos e, através da análise de componente principal (PCA) e mensuração da correlação de Pearson entre as amostras, *outliers* foram removidas antes de se dar prosseguimento para as análises posteriores.

6.3.1.1 Processamento dos dados de expressão proteica de Krug e colaboradores (2020)

Além dos dados de expressão proteica cuja obtenção está descrita nos itens anteriores, outro conjunto de dados, também baseado em espectrometria de massas, foi utilizado para a verificação da reprodutibilidade dos painéis proteicos obtidos a partir da abordagem de *machine learning* aplicada nos dados *in house* (item 6.3.6).

Este outro conjunto de dados foi obtido a partir da plataforma *LinkedOmics* (VASAIKAR et al., 2017) que hospeda dados experimentais depositados por pesquisadores de várias partes do mundo. Os dados resgatados e utilizados nesta pesquisa foram originados do trabalho de Krug e colaboradores (2020). Estes dados de expressão foram obtidos de tecidos sólidos de carcinoma mamário através de espectrometria de massas baseada em marcação *tandem mass tag* (TMT), seguindo o protocolo descrito no artigo original dos autores (KRUG et al. 2020). Antes de sua análise, os mesmos procedimentos de normalização e processamento descrito no item 4.3.1 foram aplicados para estes dados através do programa Perseus, selecionando-se apenas amostras com *status* de ER, PR, HER2 disponíveis. A ausência de informações acerca do marcador Ki-67 limitou a diferenciação entre amostras do subtipo luminal A e luminal B, de modo que apenas amostras de luminal B HER2+ enriquecido puderam ser selecionadas, uma vez que elas não dependem, necessariamente, de informações de Ki-67 para serem classificadas. As informações clínicas e patológicas destas amostras biológicas são descritas na tabela 4, porém informações sobre idade das pacientes, grau e tipo histológico dos tumores não

estão disponíveis. A classificação imunoistoquímica do subtipo de cada amostra seguiu os mesmos critérios utilizados para as amostras *in house*.

TABELA 4 - CARACTERÍSTICAS CLÍNICO-PATOLÓGICAS DOS TUMORES UTILIZADOS POR KRUG E COLABORADORES (2020) PARA OBTENÇÃO DOS DADOS DE EXPRESSÃO PROTEICA.

Código da amostra na plataforma - <i>LinkedOmics</i>	Marcadores imunoistoquímicos			Classificação imunoistoquímica do tumor
	ER	PR	HER2	
X03BR005	POS	POS	NEG	Luminal A
X06BR005	POS	POS	NEG	Luminal A
X11BR020	POS	POS	NEG	Luminal A
X11BR025	POS	POS	NEG	Luminal A
X11BR027	POS	POS	NEG	Luminal A
X11BR054	POS	POS	NEG	Luminal A
X18BR004	POS	POS	NEG	Luminal A
X18BR019	POS	POS	NEG	Luminal A
X20BR001	POS	POS	NEG	Luminal A
X21BR002	POS	POS	NEG	Luminal A
X22BR006	POS	POS	NEG	Luminal A
X05BR001	POS	POS	POS	Luminal B
X09BR005	POS	POS	POS	Luminal B
X11BR004	POS	POS	POS	Luminal B
X05BR026	POS	POS	POS	Luminal B
X11BR010	POS	POS	POS	Luminal B
X22BR005	POS	POS	POS	Luminal B
X11BR019	POS	POS	POS	Luminal B
X01BR017	NEG	NEG	POS	HER2+ Enriquecido
X06BR003	NEG	POS	POS	HER2+ Enriquecido
X11BR028	NEG	NEG	POS	HER2+ Enriquecido
X11BR030	NEG	NEG	POS	HER2+ Enriquecido
X21BR001	NEG	NEG	POS	HER2+ Enriquecido
X01BR027	NEG	NEG	NEG	Triplo Negativo

X01BR040	NEG	NEG	NEG	Triplo Negativo
X05BR029	NEG	NEG	NEG	Triplo Negativo
X05BR042	NEG	NEG	NEG	Triplo Negativo
X05BR043	NEG	NEG	NEG	Triplo Negativo
X05BR045	NEG	NEG	NEG	Triplo Negativo
X11BR003	NEG	NEG	NEG	Triplo Negativo
X11BR023	NEG	NEG	NEG	Triplo Negativo
X11BR024	NEG	NEG	NEG	Triplo Negativo
X11BR042	NEG	NEG	NEG	Triplo Negativo
X20BR005	NEG	NEG	NEG	Triplo Negativo
X20BR007	NEG	NEG	NEG	Triplo Negativo

Fonte: Krug e colaboradores (2020) em *LinkedOmics*.

6.3.2 Análises estatísticas dos dados proteômicos

Todas as análises estatísticas abaixo descritas foram realizadas através do programa Perseus (v. 1.16.15.0).

Os dados, uma vez processados e normalizados, foram submetidos à análise de variância (*One-Way ANOVA*) com aplicação da correção de *False discovery rate* (FDR), baseada em permutação (250 randomizações). Cada grupo (subtipo) foi comparado um contra o outro, e apenas proteínas com valor de P-FDR <0,05 foram consideradas diferencialmente expressas. Visando determinar exatamente em quais comparações as proteínas apresentavam diferenças de expressão significativas, o pós-teste de Tukey (FDR 0,05) foi aplicado sobre o resultado do teste de ANOVA. No final, obteve-se uma lista de proteínas diferencialmente expressas (DEPs), utilizada para parte das análises posteriores.

Outra abordagem estatística utilizada sobre os dados processados foi a análise de teste-*t* através da abordagem um-contra-todos. Para tanto, os dados foram separados em quatro grupos para comparação em pares (amostras luminal A vs não luminal A; Luminal B vs não luminal B; HER2+ enriquecido vs não HER2+ enriquecido; Triplo negativo vs não triplo negativo). Para cada comparação, o teste-*t* foi aplicado com o valor de P <0,05 sendo utilizado para

determinar as proteínas com diferença de expressão. No final, quatro listas de DEPs foram obtidas, uma lista para cada subtipo.

Para as listas de DEPs geradas na comparação um-contra-todos, um filtro adicional foi aplicado para determinar quais proteínas seriam mantidas para a continuidade das análises: o valor de $\log_2\text{-fold-change}$ (Log^2FC). Apenas proteínas consideradas super-expressas ($\text{Log}^2\text{FC} > 0,58$) ou sub-expressas ($\text{Log}^2\text{FC} < - 0,58$) foram mantidas.

6.3.3 Identificação de oncogenes e genes supressores de tumor diferencialmente expressos em cada subtipo

Após a obtenção de proteínas diferencialmente expressas através da abordagem de um-contra-todos, os genes que codificam as DEPs foram submetidos à plataforma *Network of cancer genes & Healthy drivers (NCG v. 7.0)*, uma plataforma que possui coleções manualmente curadas de oncogenes, genes supressores de tumor, candidatos a oncogenes e candidatos a genes supressores de tumor. A classificação dos genes é realizada a partir da frequência de duplicações e mutações, e prevalência de alterações de ganho ou perda de função em tumores da base de dados TCGA. A classificação de oncogenes ou genes supressores de tumor em “candidatos” ou “canônicos” se dá através do número de publicações, métodos de análise e cânceres afetados que suportam a classificação em genes com potencial para serem oncogenes / supressores de tumor (candidatos), ou a classificação em oncogenes / supressores de tumor bem estabelecidos como tal (canônicos).

6.3.4 Análise de enriquecimento e anotação funcional

Inicialmente, cada lista de DEPs (originadas do teste- t da comparação um-contra-todos) foi submetida à análise na base de dados de assinaturas moleculares (MsigDB, versão 7.4) (LIBERZON et al., 2011), buscando-se sobreposições entre os genes que codificam cada DEP com as assinaturas moleculares disponíveis na coleção *Reactome*. O compêndio de perfis de expressão *Global Cancer Map (Broad Institute)* foi utilizado como parâmetro de obtenção de expressão gênica, e apenas sobreposições significativas (valor de $P\text{-FDR} < 0,05$) foram mantidas. Como resultado desta análise, obtivemos vias e

processos biológicos significativamente enriquecidos para cada uma das listas de DEPs, ou seja, para cada um dos subtipos. Estes resultados obtidos através do MsigDB foram explorados com o auxílio da ferramenta online *interactivenn* (HEBERLE et al., 2015) com o intuito de determinar quais vias e/ou processos são compartilhados entre os subtipos, e quais eram enriquecidos exclusivamente em um ou outro subtipo. Além disso, para cada via/processo enriquecido foi calculado um *Z-score* de enriquecimento: esta pontuação considera a proporção de genes super expressos e sub expressos para determinar se, para aquele conjunto de DEPs, a via/processo está ativada ou desativada.

Posteriormente, outra análise de enriquecimento foi realizada a partir do resultado da análise de *machine learning*. As DEPs ranqueadas no *top 30* da classificação de cada subtipo foram avaliadas quanto a funções moleculares, processos biológicos e compartimentos celulares segundo os termos de *Gene Ontology*. A análise foi conduzida na *Database for Annotation, Visualization and Integrated Discovery* (DAVID, v. 6.8) (HUANG et al., 2009), sem restrições de significância devido à baixa quantidade de DEPs submetidas à análise. Os resultados da análise no DAVID foram importados e analisados na plataforma R (*RStudio* v. 1.1.463) segundo a *pipeline* de análise do pacote *GOplot* (WALTER et al., 2015), que inclui a plotagem dos resultados e cálculo do *Z-score* de enriquecimento para cada termo de GO e seu conjunto de DEPs.

6.3.5 Análise de predição computacional de interatoma

Para as análises computacionais de predição de interatoma, foram utilizados os conjuntos de DEPs de cada subtipo oriundos da comparação um-contratodos. Contudo, apenas as DEPs relacionadas com vias enriquecidas (vindas das análises de enriquecimento do MsigDB) em subtipos específicos foram analisadas. Desta forma, um interatoma para cada subtipo foi predito, utilizando-se apenas DEPs envolvidas em vias/processos específicos de cada subtipo.

Todas as predições de interatoma foram conduzidas na plataforma *Search Tool for the Retrieval of Interacting Genes/Proteins* (STRING, v. 11.5). Este programa possui um algoritmo próprio que calcula uma pontuação para cada par de interação proteína-proteína – diretas ou indiretas – com base em

fatores como: Interações validadas experimentalmente, co-expressão, co-ocorrência, co-localização e fusão gênica. Essa pontuação exprime o quão confiáveis são estas interações, ou seja, qual a chance de estas interações de fato ocorrerem no ambiente celular, de modo que apenas interações com pontuação maior do que 0,4 foram mantidas para análise, visando aumentar a robustez do interatoma.

Após a predição das interações, os dados foram exportados e analisados pelo programa *cytoscape*. Através deste programa, os dados de interação foram plotados de forma a permitir uma visualização clara de cada par de interação, e cada conjunto de DEPs foi identificado com as vias/processos em que estavam envolvidas, segundo os dados vindos da análise do MsigDB.

6.3.6 Análises de *Machine learning* e clusterização hierarquizada

Nesta pesquisa, metodologias de *machine learning* foram utilizadas para ranquear, selecionar e determinar o mínimo número possível de DEPs necessárias para compor painéis proteicos que podem ser utilizados para discriminar entre os subtipos de câncer de mama. Em outras palavras, modelos de classificação foram montados com o intuito de construir painéis proteicos que possam classificar e discriminar tumores de câncer de mama em subtipos, com base na expressão das proteínas que compõem o painel. As amostras foram separadas de forma a permitir que o modelo de predição fosse criado para quatro comparações diferentes: Tumores luminais A x não luminais A; Tumores luminais B x não luminais B; Tumores HER2+ enriquecido x não HER2+ enriquecido e tumores triplo negativos x não triplo negativos.

O método de *machine learning* utilizado está implementado na plataforma Perseus na forma de três etapas sequenciais: (1) Seleção e otimização do método e parâmetros de classificação, (2) seleção das moléculas e (3) validação cruzada. A etapa inicial do método consiste na determinação do método de classificação e dos parâmetros de restringência utilizados. O algoritmo de classificação escolhido foi o *Support Vector Machines* (SVMs). O SVM é um dos modelos de predição mais robustos disponíveis atualmente, que utiliza estatísticas de regressão não-probabilística binária para classificar as moléculas com base em seu poder de determinar em qual dos dois grupos cada amostra

pertence. O SVM atua separando o conjunto total de amostras em um conjunto de treino e outro de validação. O modelo é construído a partir da repetição da classificação no modelo de treino, para uma posterior validação do modelo no conjunto de validação. A classificação foi realizada em conjuntos de treino que representavam 15% das amostras, em 250 partições *randômicas* de classificação. Como nosso conjunto de dados não foi composto por moléculas com comportamento de expressão linear, o *kernel* utilizado foi a função radial básica, que se utiliza dos parâmetros C e Sigma para determinar a restringência da classificação. Neste estudo, foram utilizados os valores C e Sigma que garantissem menos de 15% de erro na classificação, que pode ser considerada uma margem restritiva e robusta para um modelo de predição.

A próxima etapa do processo consiste na classificação das moléculas segundo o método e os parâmetros previamente selecionados. Utilizando o SVM com parâmetros C e Sigma, juntamente com o teste de ANOVA para selecionar as proteínas com expressão mais destoante entre os grupos, as DEPs foram ranqueadas conforme seu poder discriminativo, e o menor número possível de DEPs necessárias para diferenciar os grupos também foi determinado. Ao fim, uma lista de DEPs ranqueadas foi obtida para cada subtipo, bem como um número de DEPs necessárias para discriminar cada subtipo dos demais.

A última etapa foi a validação cruzada do modelo de predição, que utiliza as DEPs mais bem ranqueadas de cada modelo segundo o número mínimo necessário determinado na etapa anterior. Ou seja, se o modelo de predição determina que um painel com apenas cinco DEPs é suficiente para discriminar, por exemplo, quais tumores são luminais B e quais não são, as cinco DEPs mais bem ranqueadas na comparação tumores luminais B x não luminais B irão compor este painel de cinco proteínas. A validação cruzada do modelo irá avaliar o quão eficazmente o modelo consegue classificar o subtipo dos tumores, baseando-se na criação de uma matriz de confusão e medição dos valores de especificidade, sensibilidade e taxa de falsos positivos da classificação das amostras em cada subtipo. Após a validação cruzada, o método de *clusterização* hierarquizada, baseado em distâncias Euclidianas, foi utilizado para visualização do padrão de expressão das proteínas de cada painel.

Visando confirmar o poder dos modelos de predição e dos painéis proteicos para a classificação de tumores mamários em subtipos, os dados de expressão proteica obtidos do trabalho de Krug e colaboradores (2020) foram submetidos à análise de regressão logística e construção de curvas ROC (*Receiver Operating Characteristic Curves*) no programa SPSS, utilizando-se do valor de área sob a curva (*AUC*) para determinar o quão bons nossos painéis são para classificar amostras desta coorte independente. O painel proteico proposto para cada subtipo foi testado neste novo conjunto de amostras, utilizando-se 0,5 como valor de corte para a classificação via regressão logística, juntamente com um intervalo de confiança de 95% para a curva ROC.

7 RESULTADOS

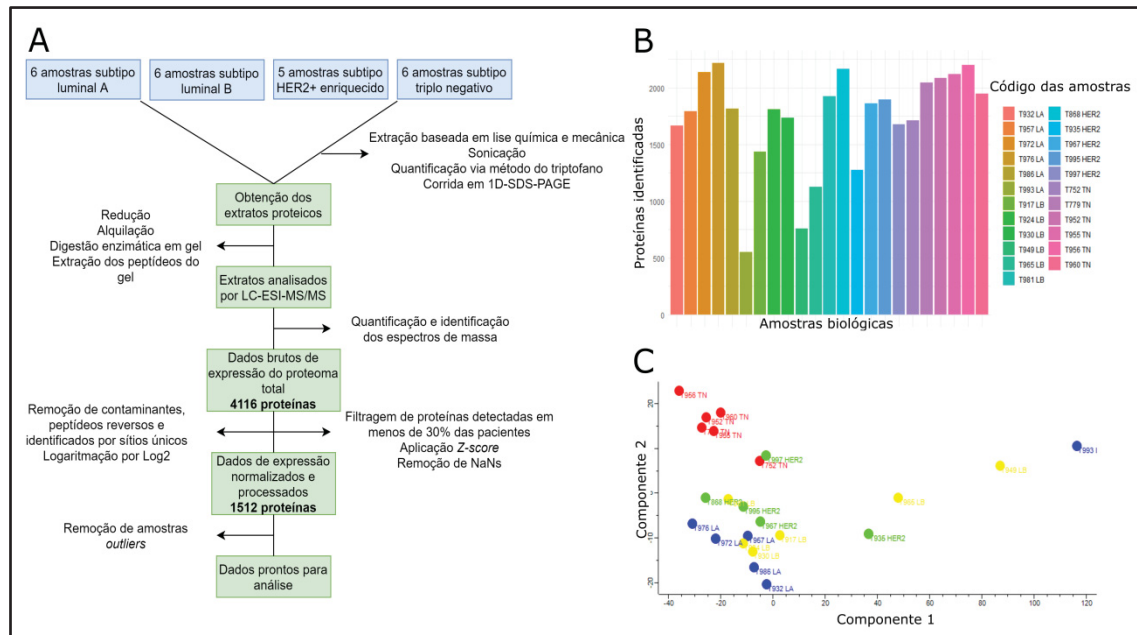
7.1 QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA TOTAL DOS TUMORES ANALISADOS

Neste estudo, metodologias proteômicas baseadas em espectrometria de massas livre de marcação e de bioinformática foram utilizadas para obtenção do proteoma total de tumores de câncer de mama, e para posterior comparação do proteoma de tumores de diferentes subtipos. O fluxograma da imagem 8A ilustra o processo de obtenção, processamento dos dados brutos e as quantidades de proteínas identificadas e filtradas no decorrer da análise.

Em média, cada amostra apresentou de 1500 a 2000 proteínas compondo seu proteoma (FIGURA 8B), com 4116 proteínas diferentes identificadas considerando o conjunto de proteínas contidas em todas as amostras. Contudo, já de início foi possível perceber que as amostras T935, T993 e T956 apresentaram uma quantidade muito menor de proteínas identificadas em relação às demais amostras. Após o processamento dos dados brutos, o proteoma total foi filtrado de maneira a manter apenas proteínas que foram expressas em pelo menos 70% das amostras, o que diminuiu a quantidade total de proteínas para 1512. A análise de componentes principais (PCA), aplicada nas amostras após a filtragem supracitada, nos mostrou que quatro amostras apresentavam um padrão de expressão completamente diferente das demais (FIGURA 8C), incluindo as amostras T935, T993 e T956, juntamente com a

amostra T949. Estas quatro amostras foram consideradas *outliers* e removidas das análises posteriores.

FIGURA 8 – ESQUEMA E RESULTADOS DA QUANTIFICAÇÃO E IDENTIFICAÇÃO DO PROTEOMA DE CADA AMOSTRA BIOLÓGICA.

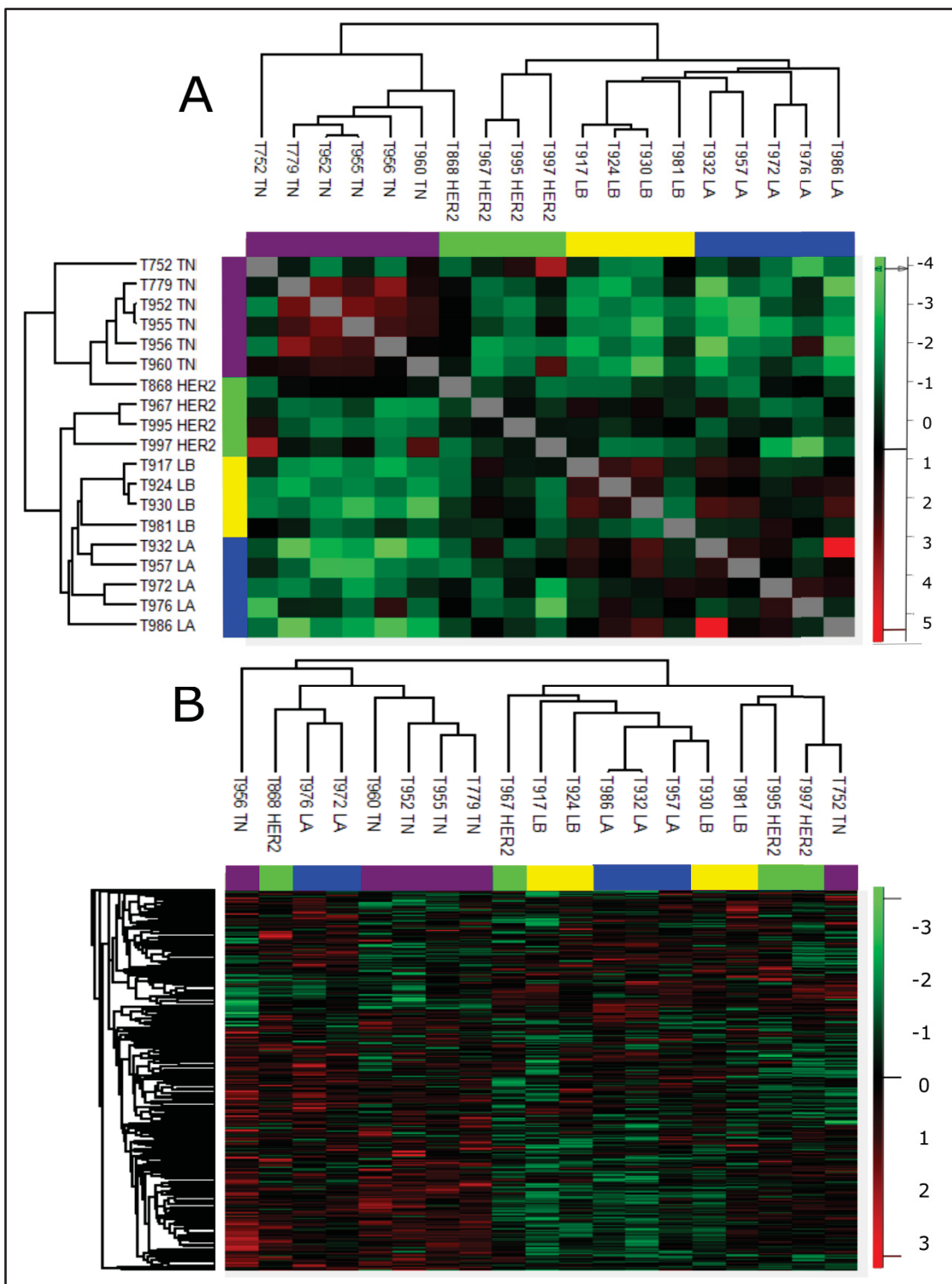


(Legenda: (A) Fluxograma do processo de preparo das amostras para LC-ESI-MS/MS e etapas de processamento que culminam com os proteomas prontos para análise. (B) Gráfico de barras representando a quantidade de proteínas identificadas por amostra. (C) Análise de componente principal de todas as amostras considerando o proteoma total. Quatro amostras foram consideradas *outliers* e retiradas das análises posteriores. Azul = Amostras LA. Amarelo = Amostras LB. Verde = Amostras HER2. Vermelho = Amostras TN).

Em seguida, as amostras mantidas foram analisadas através de uma *clusterizaçã* baseada em correlaçã de Pearson, visando determinar se, considerando o proteoma total, as amostras de mesmo subtipo possuem maior similaridade entre si do que em relaçã a amostras de outros subtipos. O *heatmap* obtido (FIGURA 9A) demonstra que, apesar da tendênciã geral de maior correlaçã entre amostras de mesmo subtipo, o proteoma total não reflete de maneira fiel a classificaçã imunistoquímica baseada em ER, PR, HER2 e Ki-67. Também considerando o proteoma total, uma *clusterizaçã* hierárquica baseada em distâncias Euclidianas foi realizada para verificar se o proteoma total é capaz de agrupar as amostras em um padrã que reflita a classificaçã imunistoquímica entre os subtipos (FIGURA 9B). O *heatmap* resultante ilustra

que o proteoma total não reflete, necessariamente, a classificação em subtipos imunoistoquímicos, agrupando as amostras apenas parcialmente segundo seus subtipos. Ambos os resultados sugerem que as diferenças proteômicas entre tumores são influenciadas por fatores que não se refletem na classificação em subtipos imunoistoquímicos, como por exemplo, diferenças individuais entre as pacientes no que diz respeito ao microambiente tumoral, e à heterogeneidade intrínseca dos tumores. Ademais, estes resultados indicam a necessidade de filtrar o proteoma total e identificar proteínas que de fato possuam diferenças de expressão entre os subtipos e que, conseqüentemente, apresentem padrões de expressão relacionados a esta classificação.

FIGURA 9 – CLUSTERIZAÇÕES PARA DETERMINAÇÃO DA SIMILARIDADE ENTRE AMOSTRAS, CONSIDERANDO O PROTEOMA TOTAL.



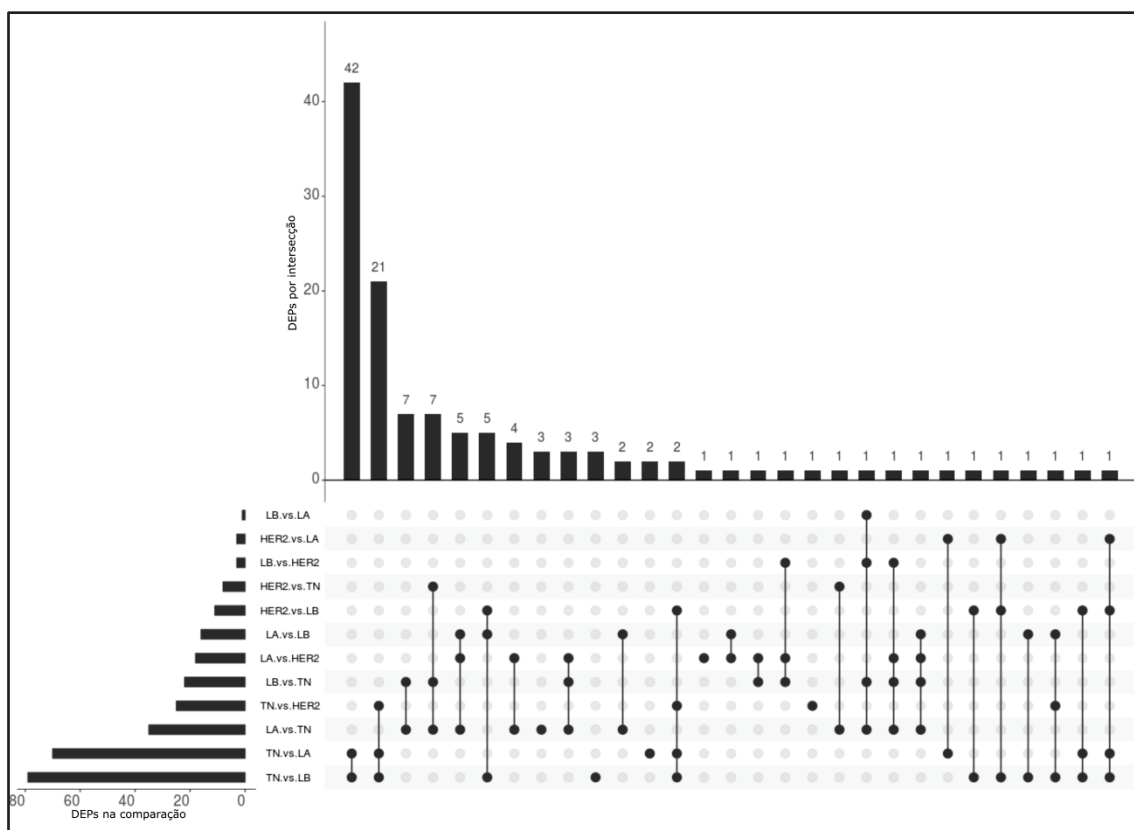
(Legenda: Azul = Amostras LA. Amarelo = Amostras LB. Verde = Amostras HER2. Roxo = Amostras TN. (A) *Clusterização* baseada em correlação de Pearson entre as amostras, considerando o proteoma total. (B) *Clusterização* hierarquizada das amostras considerando o proteoma total).

7.2 ALTERAÇÕES DE EXPRESSÃO PROTEICA IDENTIFICADAS NAS COMPARAÇÕES ENTRE OS PROTEOMAS DE CADA SUBTIPO

A partir da observação de que o proteoma total não necessariamente reflete a classificação imunistoquímica, abordagens estatísticas foram aplicadas visando encontrar proteínas diferencialmente expressas (DEPs), ou seja, proteínas cujo padrão de expressão reflita a classificação dos subtipos luminais A e B, HER2+ enriquecido e triplo negativo.

Primeiramente, o teste de análise de variância (*one-way* ANOVA) foi aplicado, seguido do pós-teste de Tukey, para determinar, a um nível de significância de valor de P-FDR < 0,05, proteínas que possuíam diferenças de expressão entre os subtipos. Resultaram desta análise 122 DEPs que apresentavam diferença de expressão em pelo menos uma comparação entre subtipos (FIGURA 10).

FIGURA 10 – *UPSET PLOT* INDICANDO A QUANTIDADE DE DEPs DE CADA COMPARAÇÃO.



(Legenda: Gráfico de barras lateral = Quantidade de DEPs por comparação. Gráfico de barras superior e "pontos" = Quantidade de DEPs compartilhadas em cada intersecção entre comparações).

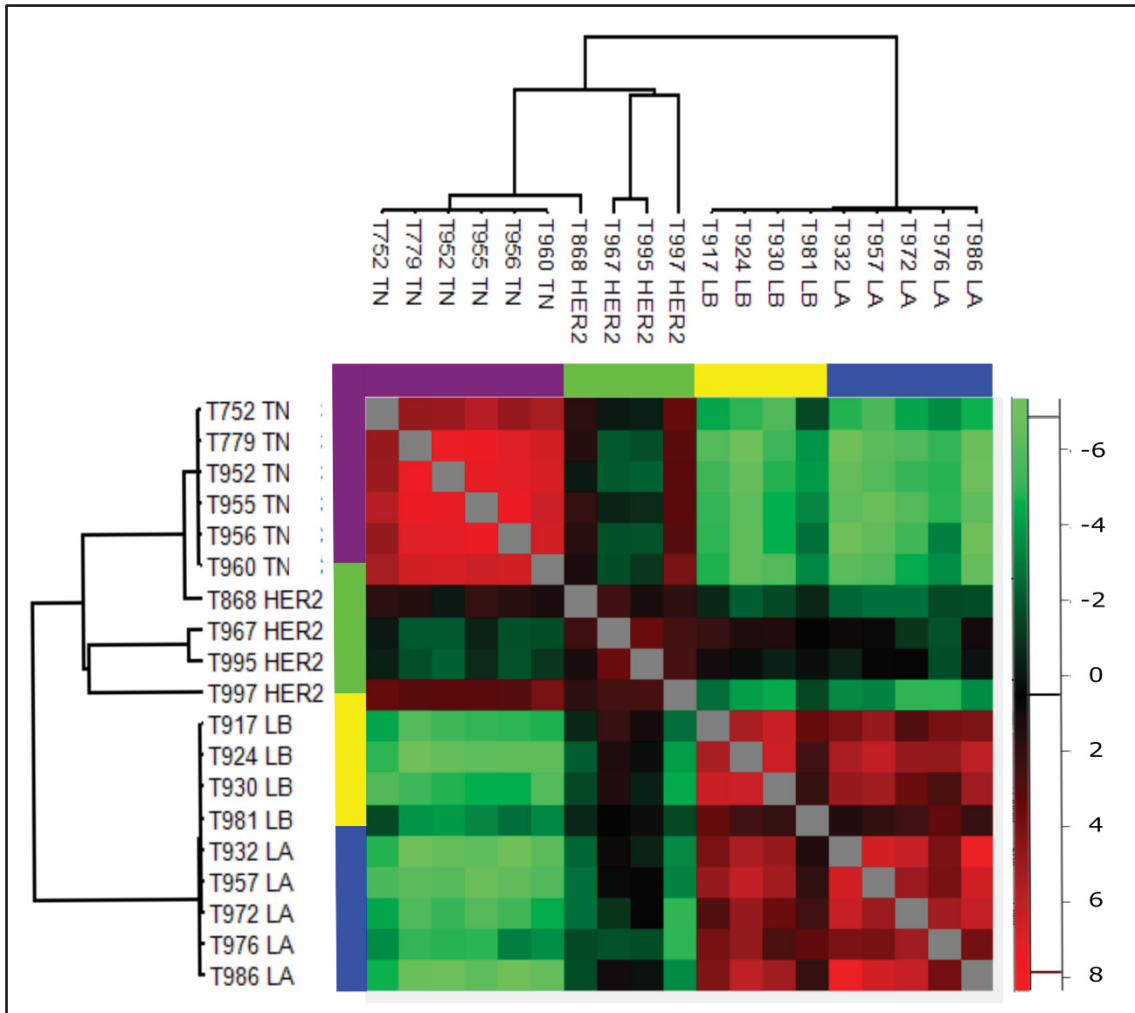
Semelhantemente ao que foi indicado pela *clusterização* baseada em correlação de Pearson e em distância Euclidiana, mesmo considerando apenas as DEPs, os subtipos LA e LB são bastante parecidos em termos de proteoma, havendo poucas proteínas com diferença de expressão entre estes subtipos (< 30). O subtipo TN, por sua vez, apresentou uma quantidade maior de DEPs em relação aos outros subtipos (> 80), demonstrando que, em termos de proteoma, este subtipo é bastante diferenciado em relação aos demais, principalmente em relação aos subtipos LA e LB. A maior parte das DEPs do subtipo TN esteve com expressão maior em relação aos outros subtipos, com poucas DEPs sub expressas. O subtipo HER2, por fim, obteve um padrão de expressão de DEPs mais variado, possuindo DEPs em relação a todos os outros três subtipos, tanto DEPs super expressas, quanto sub expressas.

Também foi possível perceber que a maioria das DEPs do subtipo TN estavam simultaneamente super expressas em relação aos subtipos LA e LB, incluindo 14 proteínas ribossomais, como as codificadas pelos genes *RPL13A*, *RPL23A*, *RPL27*, *RPS12* e *RPS15A*, e outra parte considerável de suas DEPs em relação aos subtipos LA, LB e HER2, incluindo as proteínas codificadas pelos genes *FSCN1*, *FABP5*, *EEF1A1*, *PDIA6* e *PPIB*, demonstrando uma boa capacidade inicial do proteoma de diferenciar o subtipo TN dos subtipos luminais e do subtipo HER2+ enriquecido. No subtipo LB, cinco DEPs (*EPB41L2*, *LGALS3*, *PDIA5*, *S100A10*, *S100A11*) foram sub expressas simultaneamente nas comparações com os outros três subtipos, enquanto o subtipo LA apresentou cinco DEPs (*DDX1*, *HP1BP3*, *IQGAP2*, *RTCB*, *TPP1*) super expressas simultaneamente em relação a todos os outros três subtipos, demonstrando que, assim como o TN, os subtipos luminais possuem DEPs que podem diferenciá-los, individualmente, dos outros subtipos. O subtipo HER2 não obteve nenhuma DEP que apresentasse diferenças de expressão simultâneas a todos os outros subtipos, ou seja, foram encontradas apenas DEPs com diferenças de expressão em relação a um ou outro subtipo. A tabela suplementar 2 (anexo) contém as 122 DEPs identificadas, valores de P e comparações nas quais as proteínas possuíam diferenças de expressão.

Após a obtenção e seleção destas 122 DEPs, uma nova *clusterização* baseada em correlação de Pearson foi realizada (FIGURA 11). O resultado

evidência que as DEPs destes subtipos formam um sub proteoma que reflete mais acuradamente a classificação imunohistoquímica destes tumores.

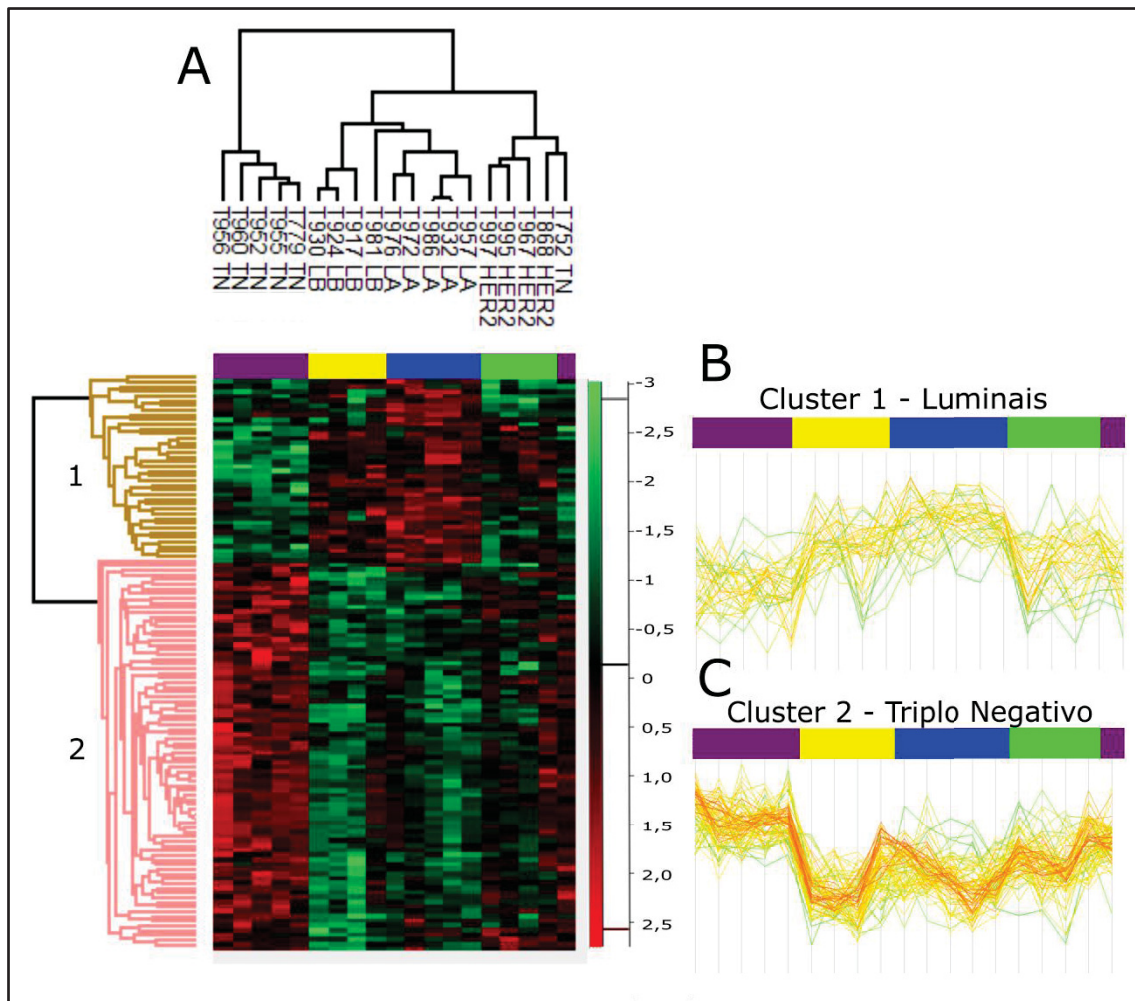
FIGURA 11 – CLUSTERIZAÇÃO BASEADA EM CORRELAÇÃO DE PEARSON, CONSIDERANDO APENAS AS 122 DEPS PROVENIENTES DA ANÁLISE DE ANOVA.



(Legenda: Azul = Amostras LA. Amarelo = Amostras LB. Verde = Amostras HER2. Roxo = Amostras TN).

Através da análise de *clusterização* hierarquizada (FIGURA 12A), foi possível visualizar mais precisamente as 122 DEPs e os padrões citados anteriormente, verificando que as amostras passaram a ser melhor agrupadas de acordo com seus subtipos. Um *cluster* muito claro de tumores TN foi formado (FIGURA 12C), diferenciando estes tumores dos demais, e da mesma forma, um *cluster* bastante evidente foi formado separando os tumores luminais dos outros, porém sem diferenciar o subtipo LA do LB (FIGURA 12B). Nenhum *cluster* foi formado de modo a separar os tumores HER2 dos outros subtipos.

FIGURA 12 – CLUSTERIZAÇÃO HIERARQUIZADA COM BASE NAS 122 DEPS
PROVENIENTES DA ANÁLISE DE ANOVA.

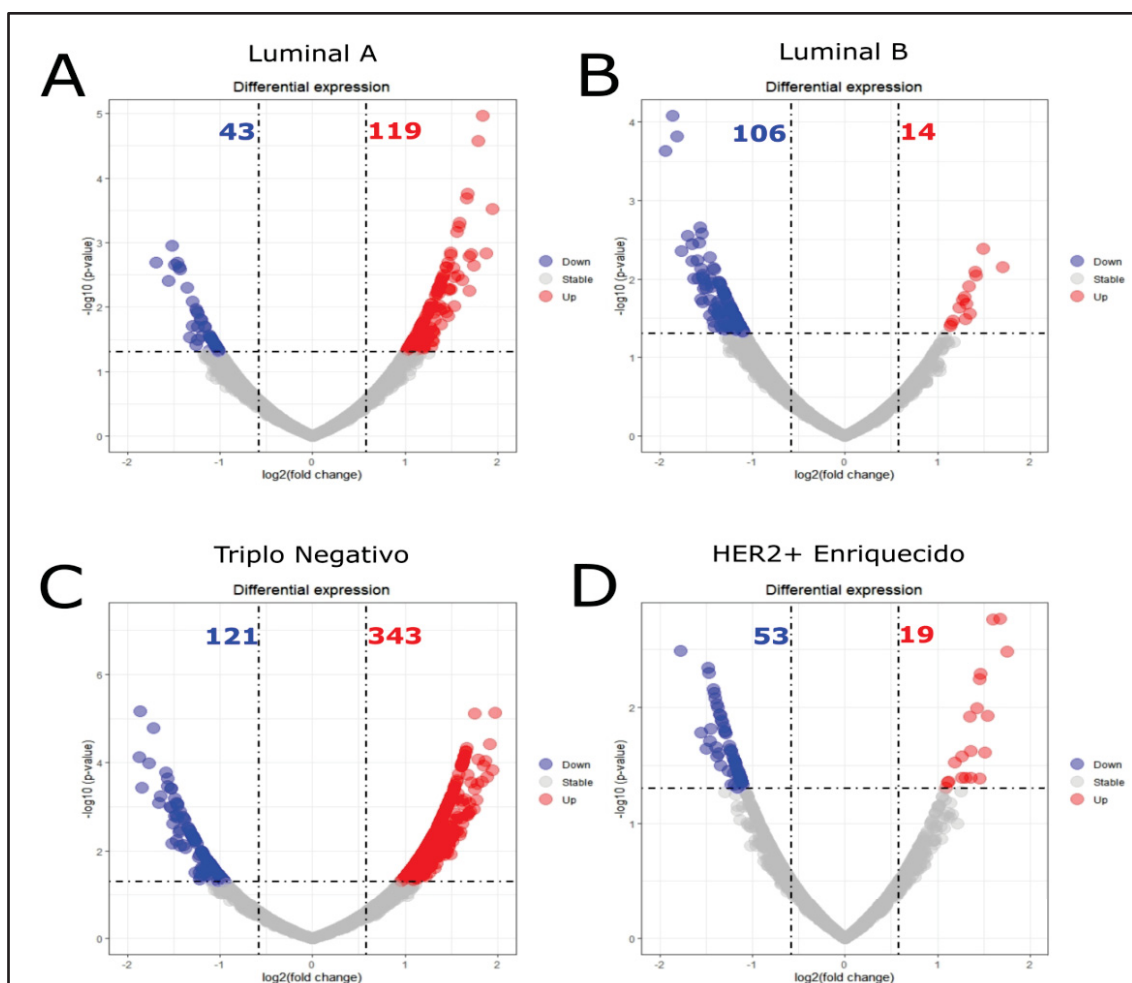


(Legenda: Azul = Amostras LA. Amarelo = Amostras LB. Verde = Amostras HER2. Roxo = Amostras TN. (A) *Clusterização* hierarquizada das amostras considerando apenas as DEPs provenientes da análise de ANOVA. (B) *Profileplot* indicando o comportamento do *cluster* formado pelas amostras luminais. (C) *Profileplot* indicando o comportamento do *cluster* formado pelas amostras triplo negativas).

Em seguida, outra abordagem de comparação de proteomas foi aplicada: A partir da utilização do teste-*t* (valor de $P < 0,05$, $\log^2FC < -0,58$; $> 0,58$), cada subtipo foi comparado contra os tumores dos outros subtipos, em uma abordagem de um-contra-todos, ou seja, tumores LA foram comparados contra um grupo de “tumores não LA”; tumores TN foram comparados contra “tumores não TN” e assim por diante. Esta abordagem teve como objetivo detectar DEPs não em comparações entre subtipos específicos, mas sim determinar DEPs que possuam padrões de expressão que diferenciem um subtipo em relação a todos

os outros. O subtipo luminal A apresentou, proporcionalmente, mais DEPs super expressas do que DEPs sub expressas (FIGURA 13A), assim como o subtipo triplo negativo (FIGURA 13C). O subtipo LB obteve mais DEPs sub expressas do que super expressas em relação aos outros subtipos (FIGURA 13B), enquanto o subtipo HER2 obteve relativamente poucas DEPs com padrão de expressão típico do subtipo (FIGURA 13D). Vale ressaltar que as “super expressões” ou “sub expressões” das DEPs são relativas aos outros subtipos, não necessariamente significando que, em relação a tecidos não tumorais, estas DEPs também teriam o mesmo padrão de super ou sub expressão. A tabela suplementar 3 (anexo) contém as DEPs identificadas, valores de P e Log^2FC de cada proteína.

FIGURA 13 – VULCANO PLOTS REPRESENTANDO O RESULTADO DAS ANÁLISES DA ABORDAGEM UM-CONTRA-TODOS.

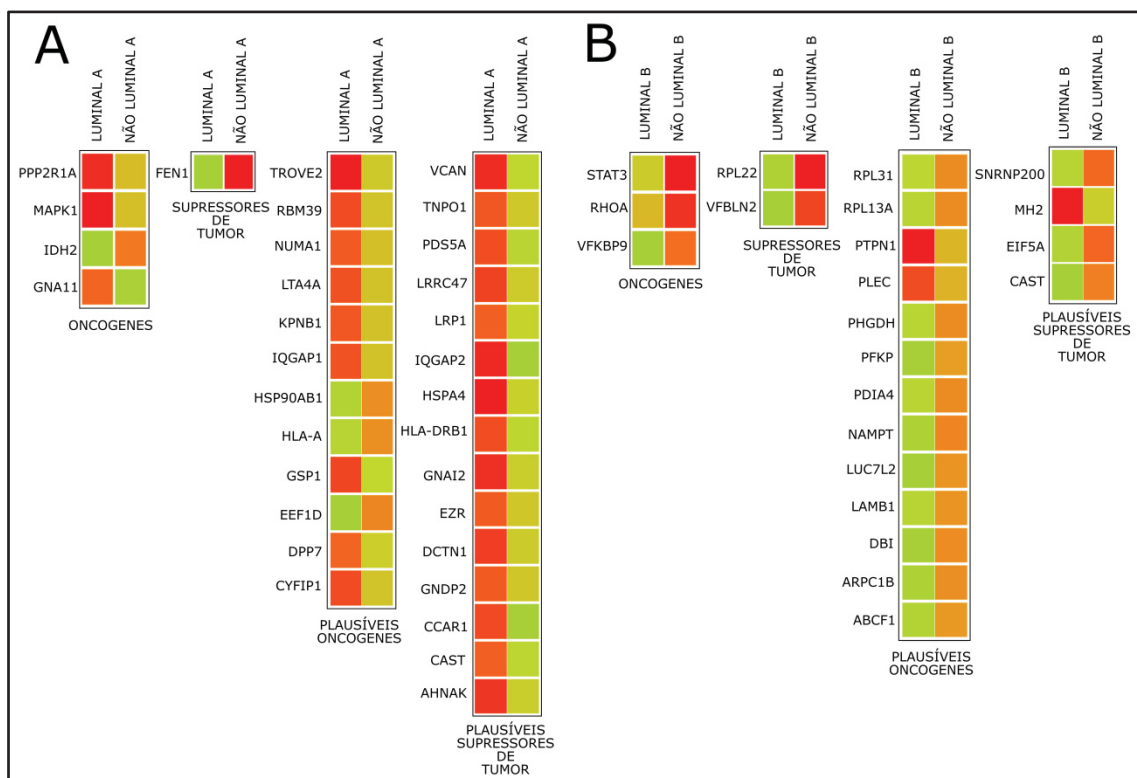


(Legenda: Azul – DEPs sub expressas (valor de P < 0,05, Log^2FC < -0,58). Vermelho – DEPs super expressas (valor de P < 0,05, Log^2FC > 0,58)

7.3 ONCOGENES E GENES SUPRESSORES DE TUMOR COM DIFERENÇAS DE EXPRESSÃO ENTRE OS SUBTIPOS.

O primeiro aspecto avaliado após a obtenção das DEPs da comparação um-contra-todos foi a expressão diferencial de oncogenes e genes supressores de tumor entre os subtipos. Estes genes são os grandes responsáveis pelo início e progressão do fenótipo maligno de células mamárias, e a expressão diferencial destes genes entre os subtipos, observada através da expressão de DEPs, pode contribuir para as diferenças de comportamento biológico e clínico observadas entre subtipos.

FIGURA 14 – ONCOGENES E GENES SUPRESSORES DE TUMOR, POTENCIAIS E CANÔNICOS, DIFERENCIALMENTE EXPRESSOS NOS SUBTIPOS LUMINAIS A E B.

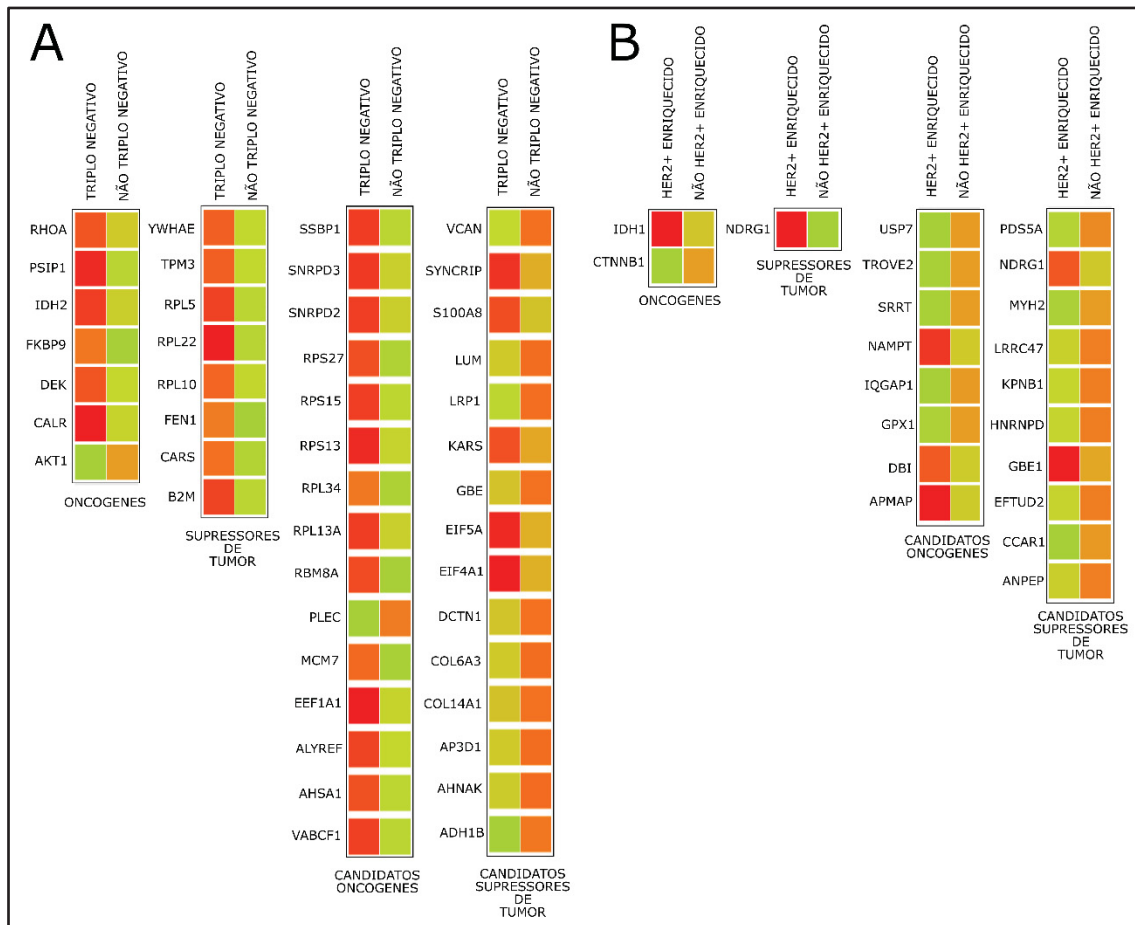


(Legenda: (A) Oncogenes e genes supressores de tumor diferencialmente expressos no subtipo luminal A em relação aos demais subtipos. (B) Oncogenes e genes supressores de tumor diferencialmente expressos no subtipo luminal B em relação aos demais subtipos. (Verde) Menor expressão relativa da DEP. (Vermelho) Maior expressão relativa da DEP).

O subtipo Luminal A obteve quatro oncogenes canônicos e 12 oncogenes candidatos diferencialmente expressos em relação aos demais subtipos, bem como 15 genes supressores de tumor candidatos, sendo um único canônico.

Grande parte dos oncogenes (12/16) apresentaram maior expressão neste subtipo em relação aos demais, e dentre os supressores de tumor, apenas o gene *FEN1* (1/16) esteve sub expresso neste subtipo em relação aos outros (FIGURA 14A).

FIGURA 15 - ONCOGENES E GENES SUPRESSORES DE TUMOR, POTENCIAIS E CANÔNICOS, DIFERENCIALMENTE EXPRESSOS NOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO



(Legenda: (A) Oncogenes e genes supressores de tumor diferencialmente expressos no subtipo triplo negativo em relação aos demais subtipos. (B) Oncogenes e genes supressores de tumor diferencialmente expressos no subtipo HER2+ enriquecido em relação aos demais subtipos. Um máximo de quinze oncogenes e/ou genes supressores de tumor foram ilustrados por categoria. (Verde) Menor expressão relativa da DEP. (Vermelho) Maior expressão relativa da DEP).

O subtipo luminal B, por sua vez, obteve três oncogenes canônicos e 13 candidatos entre suas DEPs, com dois genes supressores de tumor canônicos e quatro candidatos (FIGURA 14B). Dentre os oncogenes, apenas dois

estiveram super expressos nestes tumores (2/16), porém a maior parte dos supressores de tumor se apresentaram sub expressos em tumores luminais B em relação aos demais subtipos (4/6).

O subtipo triplo negativo foi o que obteve a maior quantidade de oncogenes (seis canônicos e cinquenta e dois candidatos) e também de genes supressores de tumor (oito canônicos e 19 candidatos), sendo 47 oncogenes super expressos neste subtipo em relação aos demais (47/59) e 12 supressores (12/27) com menor expressão neste subtipo do que nos outros (FIGURA 15A). Os tumores do subtipo HER2+ enriquecido, por sua vez, apresentaram 10 oncogenes (dois canônicos e oito candidatos) e 10 candidatos genes supressores de tumor, com pouco menos da metade (4/10) dos oncogenes sendo super expressos nestes tumores relativamente aos de outros subtipos, porém com a maior parte dos supressores (8/10) com uma acentuada diminuição de expressão neste subtipo (FIGURA 15B). O total de oncogenes e supressores de tumor (canônicos e candidatos) identificados por subtipo pode ser acessado na tabela suplementar 4 (anexo).

7.4 ANÁLISES DIFERENCIAIS DE ENRIQUECIMENTO FUNCIONAL E INTERATOMA PARA CADA SUBTIPO

Uma vez obtidas as proteínas que possuem padrões de expressão que refletem a classificação dos tumores em subtipos imunoistoquímicos, a significância biológica destas DEPs foi analisada a partir de análises funcionais e de interatoma. A plataforma MsigDB foi utilizada para investigar cada um dos conjuntos de DEPs oriundos da abordagem de comparação um-contra-todos, resultando em vias e processos biológicos significativamente enriquecidas (valor de $P\text{-FDR} < 0,05$) para cada subtipo. Vias e processos não relacionados com o câncer foram desconsiderados. Estas vias e processos não representam o total de vias desreguladas em cada um destes subtipos, mas sim aquelas que possuem diferenças de enriquecimento ou ativação (*Z-score*) em um subtipo em relação aos demais, uma vez que as DEPs utilizadas nestas análises são provenientes de comparações entre tumores de cada subtipo, e não de uma comparação de tecidos tumorais vs. tecidos não tumorais.

O subtipo LA, que possui o segundo maior número de DEPs, apresentou 92 vias/processos, retendo grande parte delas não compartilhadas com outros subtipos (54), enquanto o subtipo LB apresentou 88 vias/processos enriquecidos, com 16 delas enriquecidas apenas neste subtipo. O subtipo TN, que possui o maior número de DEPs, foi o que apresentou o maior número de vias/processos enriquecidos (93), mas ficando apenas com 15 vias não compartilhadas com outros subtipos. E por fim, o subtipo HER2, que possui o menor número de DEPs, teve apenas 20 vias/processos significativamente enriquecidos, com sete delas únicas deste subtipo. Para cada via e/ou processo compartilhado entre subtipo, uma pontuação de *Z-score* foi calculada, indicando como os mecanismos se apresentam em cada subtipo: *Z-scores* positivos indicam que o mecanismo está mais fortemente ativado naquele subtipo em relação aos demais, e *Z-scores* negativos indicam mecanismos com menor participação em determinado subtipo em relação aos demais (FIGURA 16).

Os subtipos luminais obtiveram duas vias em comum e que não aparecem nos subtipos TN ou HER2 enriquecido, sendo elas a via de endocitose mediada por clatrininas e tráfego através da membrana. O subtipo LA esteve envolvido na via das clatrininas (*Z-score* = 2,23) através da alta expressão das DEPs codificadas pelos genes *RAB5B*, *AP2B1*, *GPS1*, *COPS4* e *COPS2*, e na ativação (*Z-score* = 2,88) do tráfego através da membrana envolvendo, por exemplo, as DEPs codificadas pelos genes *RAB5B* e *RAB6A*, membros da família *RAS* de oncogenes, além dos genes *DCTN1* e *ACTR1A*. O subtipo luminal B, por sua vez, obteve uma desativação relativa da via de endocitose mediada por clatrininas (*Z-score* = -2,44) através da sub expressão das DEPs codificadas pelos genes *ARPC5*, *ACTR2*, *ARPC3*, *ACTR3* e *ARPC2*, todas estas relacionadas com complexos de actina, e também uma desativação relativa do tráfego membranar (*Z-score* = -3,5) envolvendo a sub expressão das mesmas cinco DEPs citadas anteriormente relacionadas com actina, além da sub expressão das DEPs *TUBA4A*, *TMED10* e *CHMP4B*.

FIGURA 16 – VIAS E PROCESSOS BIOLÓGICOS DIFERENCIALMENTE ENRIQUECIDOS E COMPARTILHADOS PELOS SUBTIPOS.



Os subtipos LB e TN compartilharam vias como, por exemplo, a via de sinalização mediada por *CLEC7A/DECTIN1*, degradação de *DVL*, regulação negativa de *NOTCH4*, sinalização mediada por *NOTCH4* e *WNT*, citoproteção por *HMOX1* e regulação da estabilidade e atividade de *PTEN*. Todas estas vias estiveram desativadas em LB em relação aos demais subtipos, com uma ativação mais conspícua no subtipo TN em relação aos subtipos LA, LB e HER2. Em relação à regulação da estabilidade e atividade de *PTEN* (*Z-score* = -2,23), o subtipo LB apresentou a sub expressão de subunidades de proteassomos (*PSMA5*, *PSMA7*, *PSMB2*, *PSMD4* e *PSMD7*), os quais também se relacionam à relativa inativação da sinalização via *NOTCH4* (*Z-score* = -2,23) e *WNT* (*Z-score* = -2,64) e degradação de *DVLs* (*Z-score* = -2,23). Em contraste, o subtipo TN apresentou super expressão de diversas subunidades de proteassomos (*PSMA1*, *PSMA5*, *PSMA7*, *PSMB1*, *PSMB2*, *PSMB9*, *PSMD4*, *PSMD7* e *PSMD9*) e outras DEPs como *H2AFV*, *SNX3* e *VPS29*, e baixa expressão relativa de *AKT1* e *CAV1*, envolvidas na relativa ativação da degradação de *DVL* (*Z-score* = 2,88), regulação da estabilidade e atividade de *PTEN* (*Z-score* = 2,49), sinalização mediada por *NOTCH4* (*Z-score* = 3,57) e *WNT* (*Z-score* = 3,4).

O processo de desubiquitinação foi compartilhado pelos subtipos LB e HER2, ambos com maior desativação deste processo em relação aos subtipos LA e TN, com o subtipo LB (*Z-score* = -2,82) com a sub expressão das supracitadas subunidades de proteassomos e das DEPs codificadas pelos genes *SNX3*, *RHOA* e *UCHL3*, e o subtipo HER2 (*Z-score* = -1,34) com a sub expressão relativa das DEPs codificadas pelos genes *TRIM25*, *RUVBL1*, *USP7* e *MAT2B*, e maior expressão da DEP *PTRH2*.

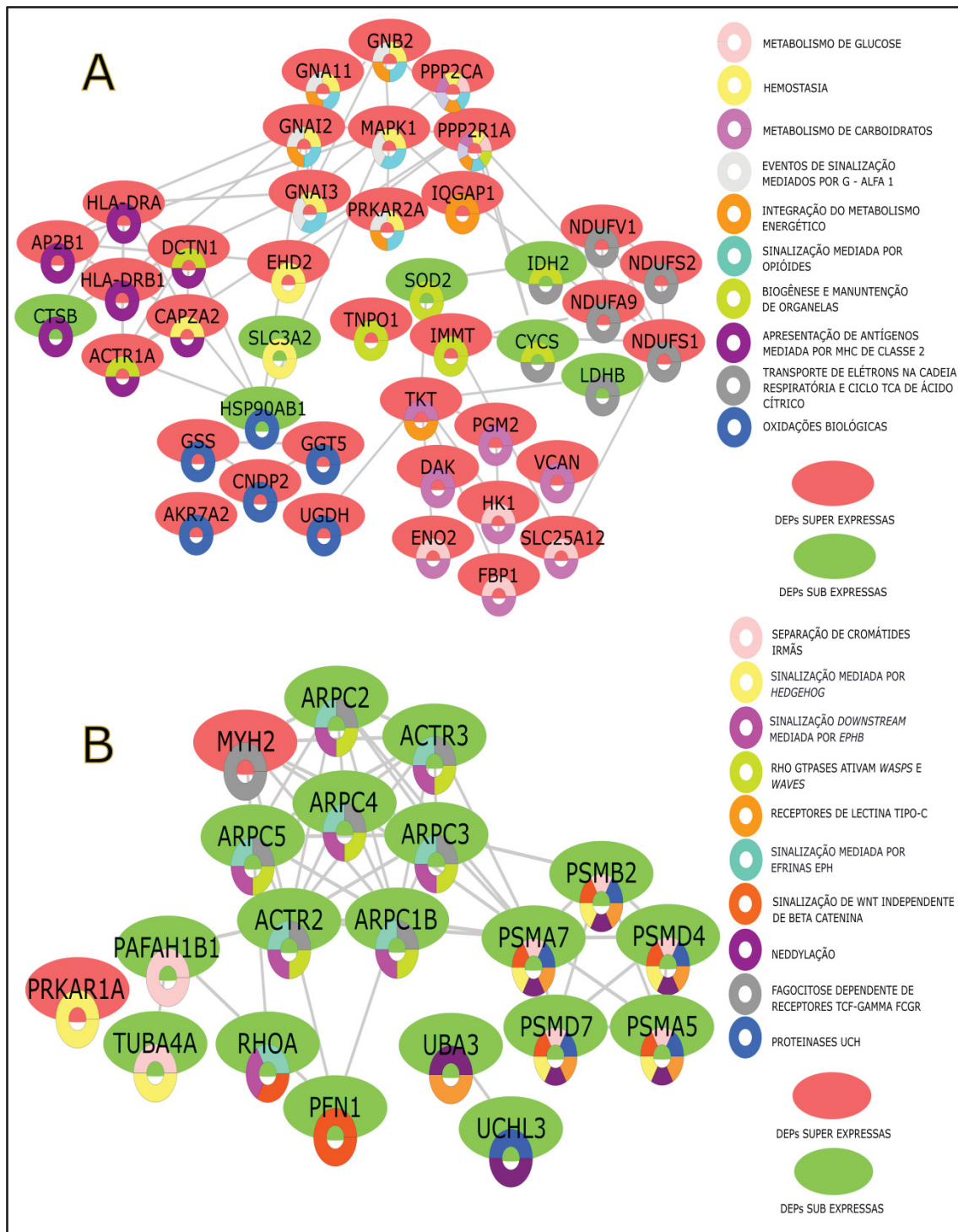
Os subtipos LA e TN obtiveram em comum uma maior ativação relativa do sistema imune adaptativo, resposta às proteínas mal enoveladas, e sinalização por citocinas do sistema imune, bem como desregulação da organização da matriz extracelular. O subtipo LA esteve relacionado à ativação do sistema imune adaptativo (*Z-score* = 1,29) e sinalização por citocinas do sistema imune (*Z-score* = 1,15) através, por exemplo, de maior expressão das DEPs codificadas pelos genes *HLA-DRA*, *HLA-DRB1*, *UFL1*, *PPP2CA* e *PPP2R1A*, e também da desregulação da organização da matriz extracelular (*Z-score* = 0) através da super expressão de *CAPN1*, *VCAN*, *FBN1* e *CAST*, e baixa

expressão de *CTSB*, *P4HB*, *PPIB*, *COLGALT1*. O subtipo TN esteve relacionado com a ativação relativa do sistema imune adaptativo (Z -score = 3,79) e sinalização por citocinas do sistema imune (Z -score = 4,64) através da alta expressão das DEPs codificadas pelos genes *PSMA1*, *PSMA5*, *PSMA7*, *PSMB1*, *PSMB2*, *PSMB9*, *PSMD4*, *PSMD7*, *PSMD9*, *SKP1* e *YWHAZ*, e relacionado também com a diminuição da expressão de DEPs relacionadas com a organização da matriz extracelular (Z -score = -3,12) através da baixa expressão das DEPs codificadas pelos genes *COL6A1*, *COL6A2*, *COL6A3*, *PLEC*, *FBN1*, *DCN* e *LUM*. Outras vias também obtiveram comportamentos discrepantes em dois ou mais subtipos. A tabela suplementar 5 (anexo) apresenta todas as vias/processos alterados, bem como o Z -score de cada uma.

Os processos e vias com enriquecimento exclusivo de cada subtipo foram analisados em termos de interação proteína-proteína, selecionando-se, para cada subtipo, as 10 vias mais significativamente enriquecidas (valor de $P < 0,05$). O termo “exclusivo” não é utilizado no sentido de afirmar que tal via e/ou processo só ocorre em um subtipo ou outro, mas sim que o mecanismo é muito mais ativo ou inativo em um subtipo específico do que nos demais. As vias enriquecidas exclusivamente em luminal A (FIGURA 17A) incluem, por exemplo, o metabolismo de glucose (Z -score = 2,44) e de carboidratos (Z -score = 3,16), biogênese e manutenção de organelas (Z -score = 0,70) e oxidações biológicas (Z -score = 1,63). O subtipo luminal B obteve as vias de sinalização por *hedgehog* (Z -score = -2,64), sinalização mediada por efrinas EPH (Z -score = -2,82), sinalização de *WNT* independente de beta catenina (Z -score = -2,64) e proteinases UHC (Z -score = -2,44), todas significativamente inativadas em relação aos demais subtipos (FIGURA 17B).

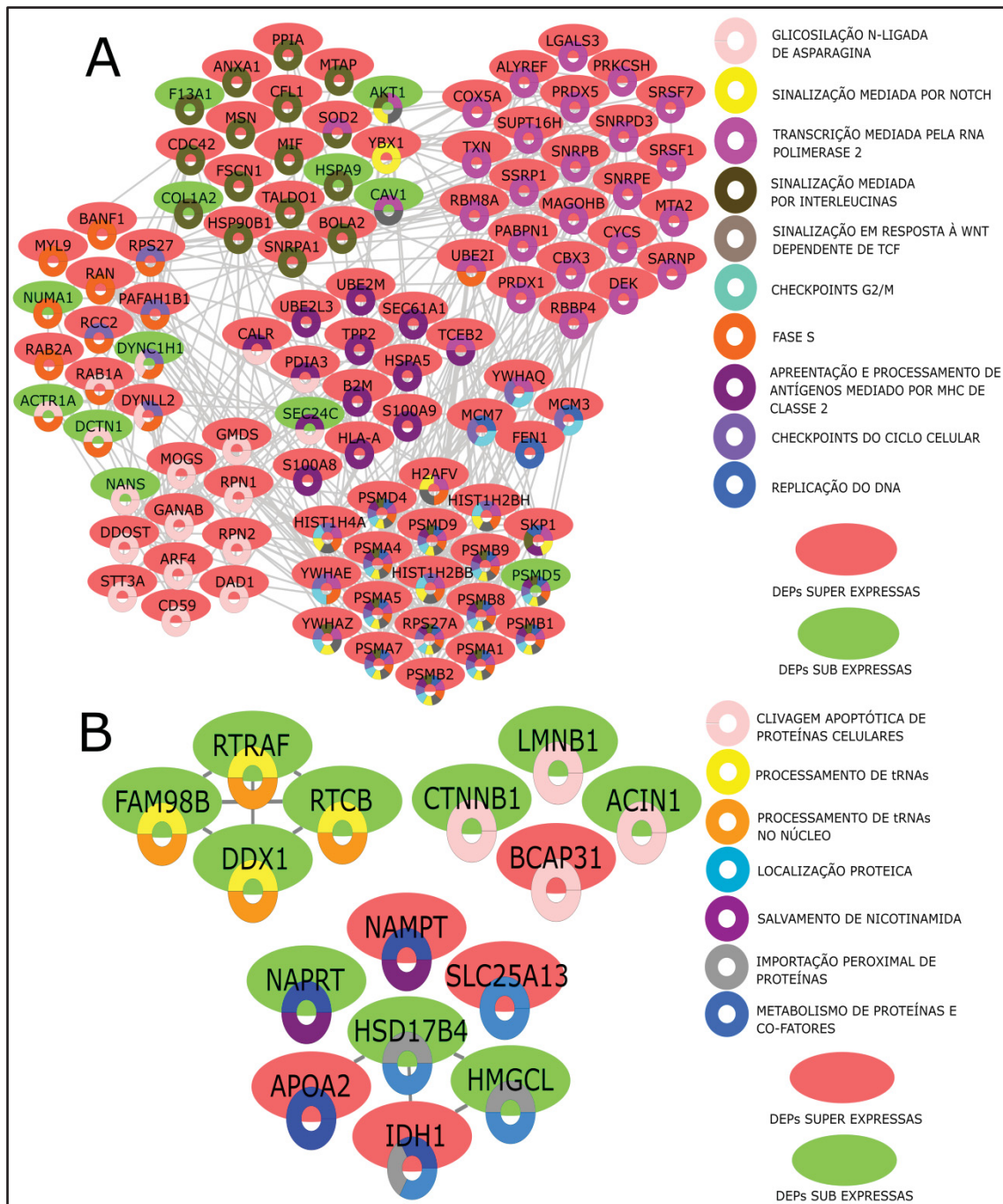
O subtipo triplo negativo obteve como vias e/ou processos significativamente enriquecidos (de modo exclusivo) a ativação das vias de apresentação e processamento de antígenos mediado por MHCs de classe 2 (Z -score = 4,31), sinalização em resposta a *WNT* dependente de TCF (Z -score = 2,98), sinalização mediada por interleucinas (Z -score = 3,77) e *NOTCH* (Z -score = 3,57), glicolisação N-ligada de asparagina (Z -score = 2,40), transcrição

FIGURA 17 – VIAS E PROCESSOS BIOLÓGICOS SIGNIFICATIVAMENTE ENRIQUECIDOS EXCLUSIVAMENTE NOS SUBTIPOS LUMINAIS.



(Legenda = (A) Interações proteína-proteína das DEPs envolvidas nos processos e vias significativamente enriquecidos de modo exclusivo no subtipo LA. (B) Interações proteína-proteína das DEPs envolvidas nos processos e vias significativamente enriquecidos de modo exclusivo no subtipo LB).

FIGURA 18 - VIAS E PROCESSOS BIOLÓGICOS SIGNIFICATIVAMENTE ENRIQUECIDOS EXCLUSIVAMENTE NOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO.



(Legenda: (A) Interações proteína-proteína das DEPs envolvidas nos processos e vias significativamente enriquecidos de modo exclusivo no subtipo TN. (B) Interações proteína-proteína das DEPs envolvidas nos processos e vias significativamente enriquecidos de modo exclusivo no HER2).

mediada pela RNA polimerase 2 (Z -score = 6,06) e outras (FIGURA 18A). Neste contexto, o subtipo HER2 obteve, por exemplo, a clivagem apoptótica de

proteínas celulares (*Z-score* = -1), processamento de tRNA (*Z-score* = -2) e metabolismo de vitaminas e cofatores (*Z-score* = 1) como processos exclusivamente enriquecidos neste subtipo (FIGURA 18B). Os demais processos exclusivamente enriquecidos em cada subtipo estão contidos na tabela suplementar 5 (anexo).

7.5 CONSTRUÇÃO DE PAINÉIS PROTEICOS A PARTIR DE MODELOS DE CLASSIFICAÇÃO BASEADOS EM *MACHINE LEARNING*

Uma vez que o significado e relevância biológica das DEPs foram acessados através das análises de enriquecimento funcional e de predição computacional de interatomos, a metodologia de *machine learning*, suportada pelo algoritmo de classificação *Support Vector Machines* (SVMs), foi utilizada para classificar e selecionar DEPs – ou conjunto de DEPs – que possuem a capacidade de discriminar os tumores de acordo com seus subtipos.

As DEPs utilizadas para a construção destes painéis são as 122 DEPs oriundas do resultado do teste estatístico de ANOVA, e toda a abordagem de *machine learning* foi realizada a partir de funções implementadas no programa Perseus. O processo de determinação de painéis foi realizado para cada subtipo, sendo que cada painel resultante possui o objetivo de complementar a classificação baseada nos marcadores ER, PR, HER2 e Ki-67. Cada painel foi elaborado de forma a permitir que a classificação do tumor em subtipos imunohistoquímicos seja confirmada ou não, ou seja, cada painel irá discriminar se dado tumor é de fato do subtipo apontado, ou não.

A primeira etapa da metodologia de *machine learning* resultou na determinação dos valores dos parâmetros C e Sigma para cada subtipo (tabela 5). Estes valores representam o quão restrigente é a análise, ou seja, qual a margem de erro que será considerada aceitável para cada modelo/painel proposto. Os valores de C e Sigmas selecionados, independentemente do subtipo, representam uma restringência bastante robusta de menos de 15% de erro para cada modelo.

Após a determinação do algoritmo de classificação e dos parâmetros de restringência, 250 repetições da classificação baseada em SVM foram

realizadas em conjuntos de treino (15% das amostras, sorteadas aleatoriamente a cada repetição), culminando no ranqueamento da lista de 122 DEPs, para cada subtipo, segundo a capacidade das DEPs de discriminar os tumores de um dado subtipo em relação aos demais. A tabela suplementar 6 (anexo) apresenta o ranqueamento das DEPs para cada subtipo.

TABELA 5 – VALORES DOS PARÂMETROS C E SIGMA E MARGEM DE ERRO PREVISTA PARA CADA UM DOS MODELOS PROPOSTOS.

Subtipo	Valor de C	Valor de Sigma	Margem de erro prevista
Luminal A	50	7	9,06%
Luminal B	50	7	14,8%
Triplo negativo	50	7	4,53%
HER2+ Enriquecido	50	5	9,86%

As DEPs com melhor ranqueamento não apenas possuem um alto potencial para discriminar subtipos através de seus padrões de expressão, mas também podem ser relevantes candidatas para explicar as diferenças biológicas e de comportamento entre estes subtipos. Através de uma análise de enriquecimento baseada em termos de ontologia gênica, a participação das DEPs do top 30 de cada subtipo foram investigadas em termos de compartimentos celulares onde elas são encontradas, funções celulares desenvolvidas e processos biológicos nas quais as DEPs tomam parte. Para cada categoria de ontologia gênica, um valor de *Z-score* foi calculado, indicando se a tendência do processo é estar ativado ou desativado naquele subtipo, ou no caso das categorias de compartimentos celulares, se as DEPs daquele compartimento tendem a estar sub ou super expressas.

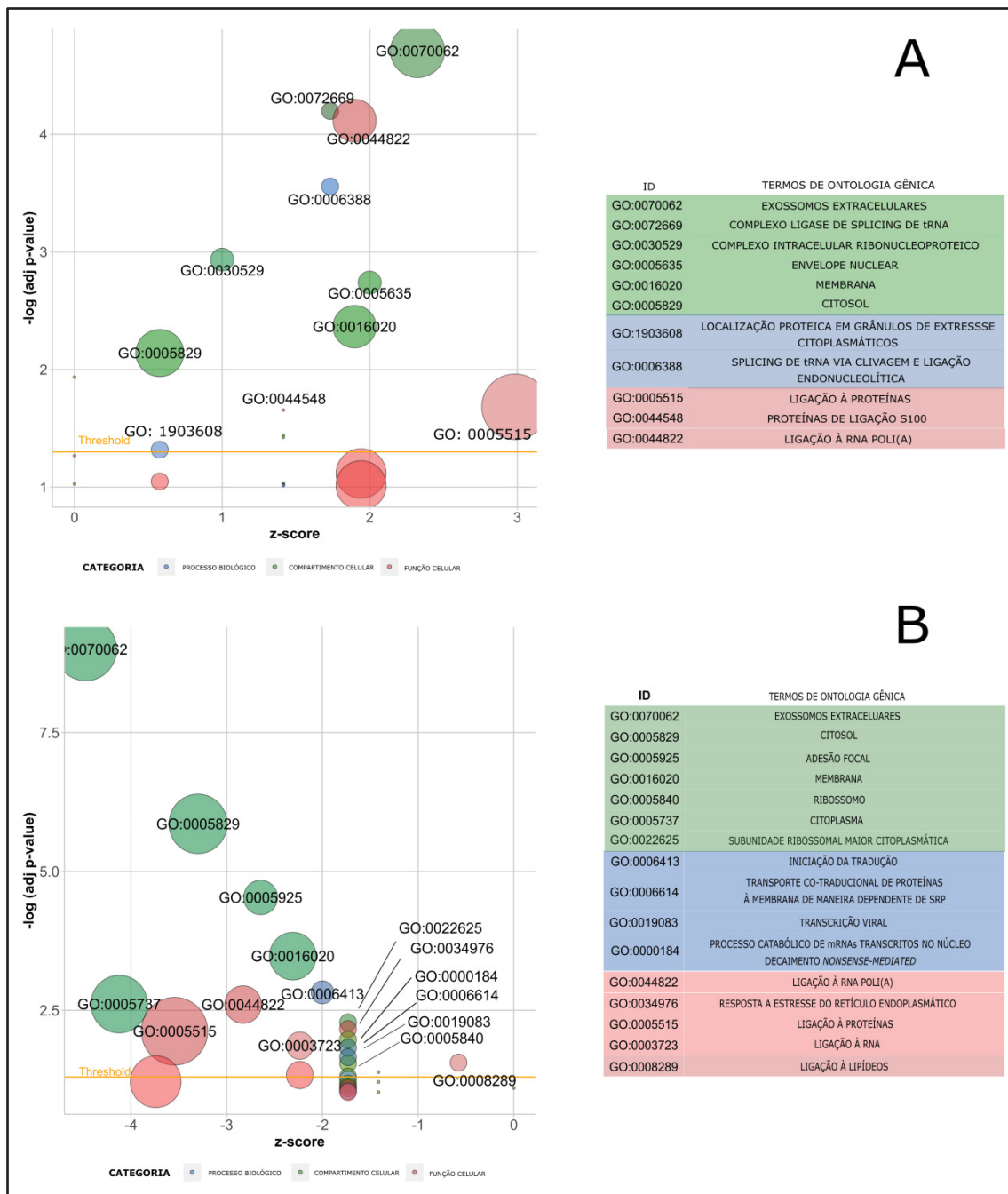
Grande parte das proteínas classificadas dentro do top 30 do subtipo luminal A podem ser encontradas em exossomos, complexos de *splicing* de

tRNA e complexos ribonucleoproteicos, exercendo funções de ligação à proteínas e ligação à cauda poli-A de RNA, indicando a participação destas DEPs na regulação da expressão e atividade de outras moléculas, atuando também em processos biológicos como localização proteica em grânulos de estresse e *splicing* de tRNA (FIGURA 19A).

O subtipo luminal B teve suas DEPs do top 30 sendo principalmente localizadas em exossomos, citosol e em regiões de adesão focal. Estas DEPs estão envolvidas em funções como ligação à cauda poli-A de RNA, resposta a estresse do retículo endoplasmático e ligação a proteínas e RNAs, indicando também o potencial regulatório destas DEPs. Através destas funções, estas DEPs participam de processos como iniciação da tradução e transporte co-transcricional (dependente de *SRPs*) de proteínas à membrana (FIGURA 19B).

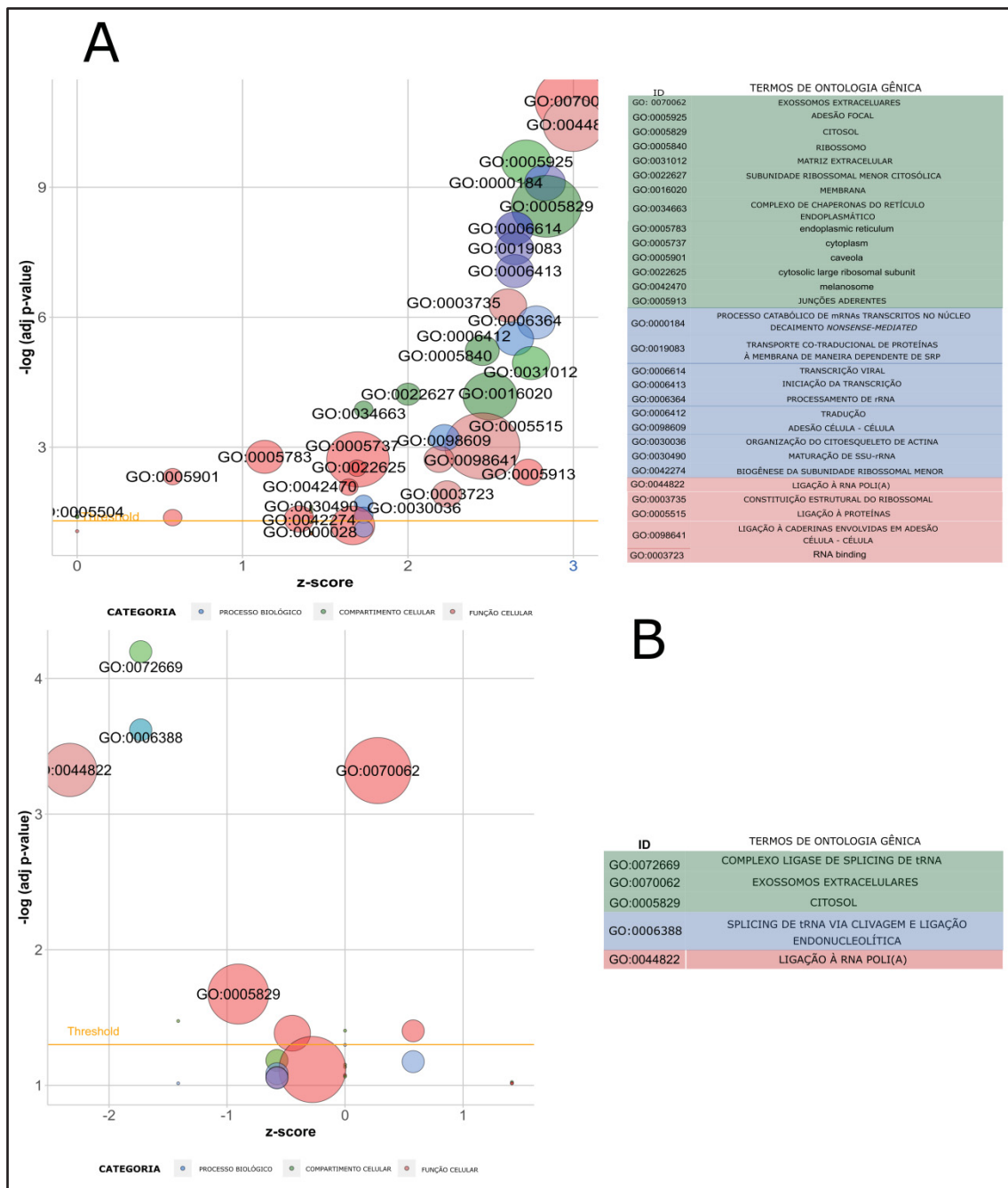
As top 30 DEPs do subtipo TN são proteínas principalmente encontradas em exossomos, regiões de adesão focal, citosol e ribossomos, participando de funções celulares que incluem ligação à proteínas e cauda poli-A de RNAs, constituição estrutural de ribossomos e ligação à caderinas envolvidas em adesão célula-célula. Além disso, estas DEPs participam de processos como catabolismo de RNAs, transporte de proteínas à membrana, iniciação da transcrição, tradução e adesão célula-célula, todos processos relevantes para a tumorigênese mamária (FIGURA 20A). O subtipo HER2, por sua vez, possuiu DEPs em seu top 30 proteínas localizadas em complexos de *splicing* de tRNA, exossomos e citosol, tendo apenas a função de ligação à cauda poli-A de RNAs e o processo de *splicing* de tRNA significativamente enriquecidos neste subtipo (FIGURA 20B).

FIGURA 19 – ANÁLISE DE ENRIQUECIMENTO BASEADA EM ONTOLOGIA GÊNICA PARA AS DEPS DO TOP 30 DA CLASSIFICAÇÃO VIA SUPPORTED VECTOR MACHINES DOS SUBTIPOS LUMINAIS.



(Legenda: O tamanho de cada “bolha” é proporcional ao número de DEPs que se enquadram no processo/via/compartimento celular. (A) Termos de ontologia significativamente enriquecidos para as DEPs do top 30 do subtipo LA. (B) Termos de ontologia significativamente enriquecidos para as DEPs do top 30 do subtipo LB.)

FIGURA 20 – ANÁLISE DE ENRIQUECIMENTO BASEADA EM ONTOLOGIA GÊNICA PARA AS DEPS DO TOP 30 DA CLASSIFICAÇÃO VIA SUPPORTED VECTOR MACHINES DOS SUBTIPOS TRIPLO NEGATIVO E HER2+ ENRIQUECIDO.



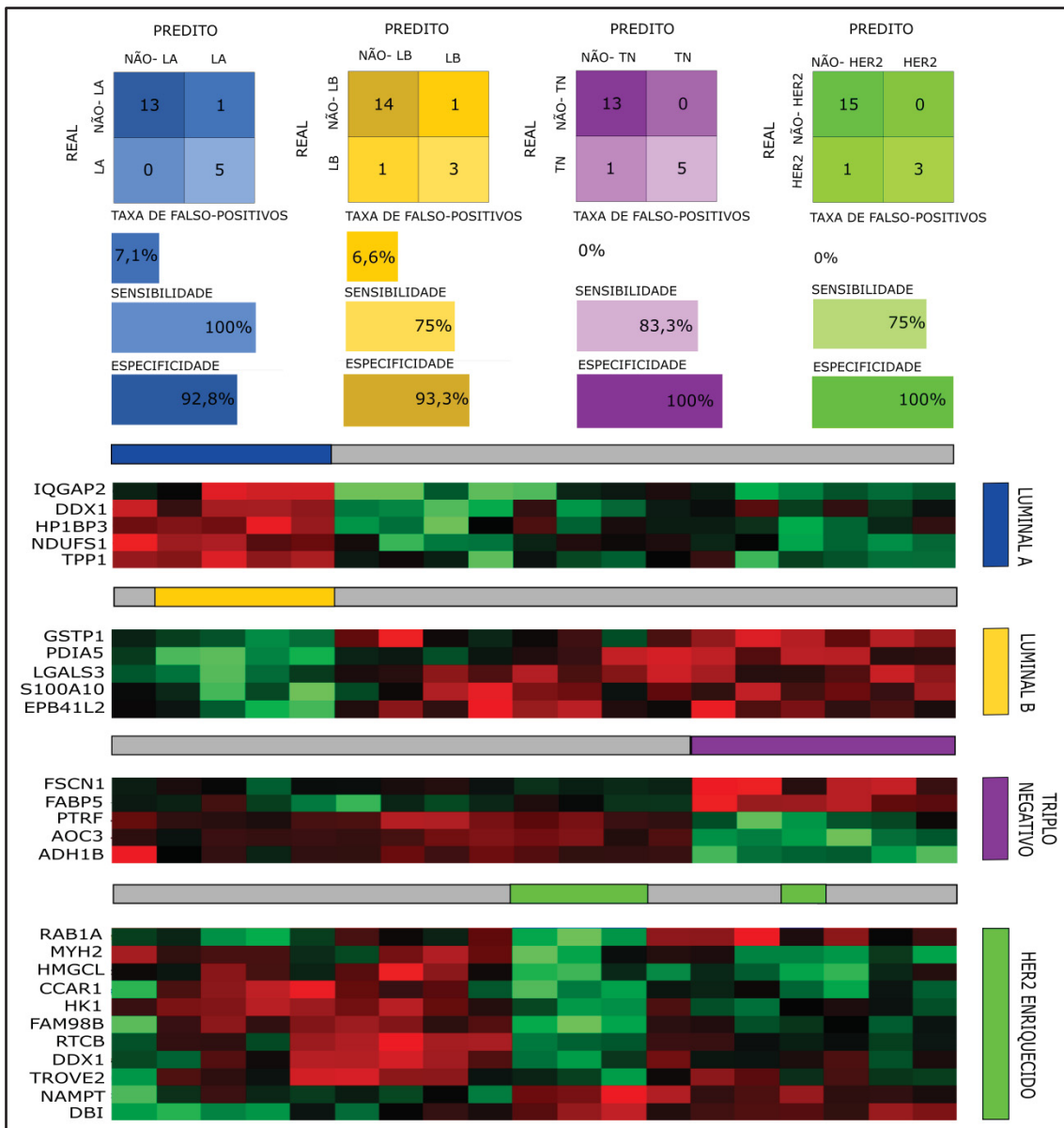
(Legenda: O tamanho de cada “bolha” é proporcional ao número de DEPs que se enquadram no processo/via/compartimento celular. (A) Termos de ontologia significativamente enriquecidos para as DEPs do top 30 do subtipo TN. (B) Termos de ontologia significativamente enriquecidos para as DEPs do top 30 do subtipo HER2+).

A etapa final da metodologia de *machine learning* consistiu na determinação do número mínimo de DEPs necessárias para compor os painéis de cada subtipo e subsequente validação cruzada do modelo no conjunto de validação. O número de DEPs por painel foi determinado tendo em mente selecionar o número mínimo de proteínas necessárias para que o erro intrínseco do modelo não ultrapasse 15%, e para que os valores de sensibilidade e especificidade sejam de, no mínimo, 75%. Se o número de DEPs determinado segundo este padrão fosse de 10, por exemplo, o painel seria formado pelas DEPs do *top* 10 do ranqueamento daquele subtipo.

Para os subtipos LA, LB e TN foi possível padronizar o número de DEPs que constituiu os painéis: cinco DEPs foram utilizadas em cada um dos painéis, acarretando erros de 8,13%, 11,73% e 6,79%, respectivamente, para os modelos destes subtipos. O subtipo HER2 necessitou de mais DEPs para compor um painel com menos de 15% de erro intrínseco, sendo o painel final composto por 11 DEPs, com erro de 10,26%. A validação cruzada dos modelos demonstra, através da mensuração da sensibilidade, especificidade e taxas de falso-positivos, o desempenho de cada modelo. A taxa de falso-positivos não ultrapassou 10% em nenhum dos painéis, com os valores de sensibilidade e especificidade nunca estando abaixo de 75% (FIGURA 21).

A validação cruzada se refere a um processo no qual cada modelo é testado dentro do próprio conjunto de dados utilizados para sua construção, avaliando o desempenho da “máquina” em classificar o subtipo de cada amostra. Uma matriz de confusão é gerada para cada modelo, indicando os acertos e erros de classificação cometidos pela máquina usando cada um dos painéis. Por exemplo, a matriz de confusão do modelo luminal A (FIGURA 21) nos mostra que, das cinco amostras LA (REAL), a máquina acertou a classificação de todas (PREDITO), com uma sensibilidade de 100%. Das quatorze amostras não-LA (REAL), a máquina acertou a classificação de treze (PREDITO), tendo assim uma taxa de falso-positivos de 7,1% e especificidade de 92,8%.

FIGURA 21 – VALIDAÇÃO CRUZADA DOS MODELOS DE CLASSIFICAÇÃO DE SUBTIPOS DE CÂNCER DE MAMA E DEPS QUE COMPÕEM CADA PAINEL.



(Legenda: Azul = Dados referentes ao modelo/painel do subtipo Luminal A. Amarelo = Dados referentes ao modelo/painel do subtipo Luminal B. Roxo = Dados referentes ao modelo/painel do subtipo triplo negativo. Verde = Dados referentes ao modelo/painel do subtipo HER2+ enriquecido).

O painel proposto para o subtipo luminal A foi composto pelas DEPs codificadas pelos genes *IQGAP2*, *DDX1*, *HP1BP3*, *NDUFS1* e *TPP1*, todas marcadoras positivas para este subtipo, ou seja, são DEPs que apresentam expressão elevada neste subtipo em relação aos demais. Em contraste, o painel do subtipo luminal B teve em sua composição cinco marcadoras negativas, em

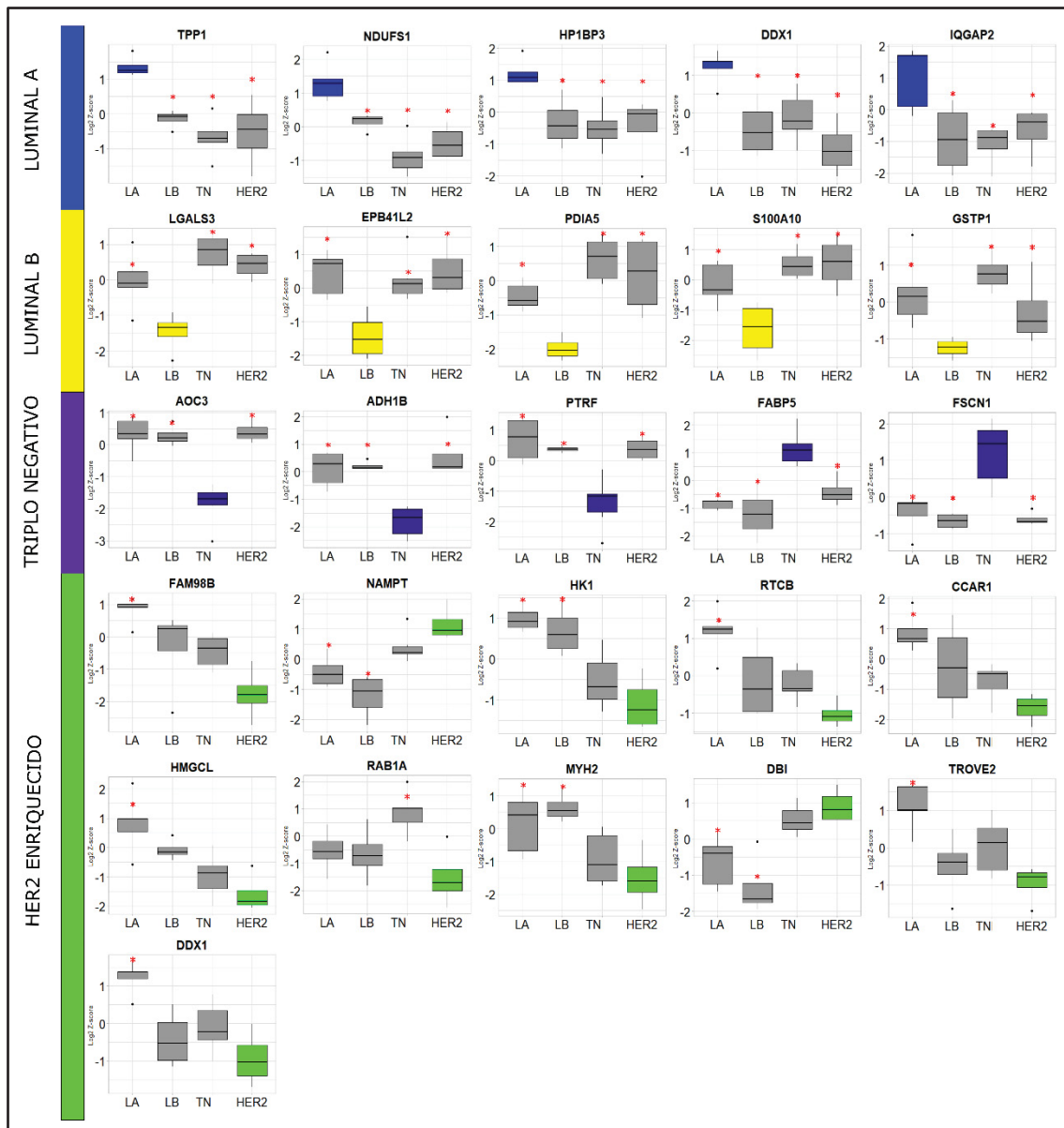
outras palavras, DEPs com baixa expressão neste subtipo em relação aos demais, sendo elas codificadas pelos genes *GSTP1*, *PDIA5*, *LGALS3*, *S100A10* e *EPB41L2*.

O subtipo triplo negativo teve um painel composto por duas marcadoras positivas (*FSCN1* e *FABP5*) e três DEPs marcadoras negativas (*PTRF*, *AOC3* e *ADH1B*), enquanto o subtipo HER2+ enriquecido foi composto por nove marcadoras negativas (*RAB1A*, *MYH2*, *HMGCL*, *CCAR1*, *HK1*, *FAM98B*, *RTCB*, *DDX1* e *TROVE2*) e duas DEPs marcadoras positivas (*NAMPT* e *DBI*). A figura 22 apresenta, na forma de *boxplots*, a expressão de todas as DEPs que compõem os painéis aqui propostos em cada um dos subtipos.

7.6 VALIDAÇÃO DOS PAINÉIS EM UMA NOVA COORTE DE AMOSTRAS

Visando determinar se os quatro painéis proteicos propostos nesta pesquisa são, de fato, robustos e capazes de corroborar a classificação de subtipos imunoistoquímicos de tumores de câncer de mama, os quatro painéis foram aplicados em uma nova coorte de amostras de câncer de mama, e o desempenho dos painéis foi mensurado através de curvas ROC (*Receiver Operating Characteristic Curves*), obtidas por cálculos de regressão logística, sendo utilizado, mais especificamente, o valor de área sob a curva (AUC) para determinação do desempenho dos painéis. Toda esta validação foi realizada através de funções implementadas no programa SPSS.

FIGURA 22 – BOXPLOTS REPRESENTANDO A EXPRESSÃO DAS DEPS QUE COMPÕEM OS PAINÉIS DE CADA SUBTIPO.



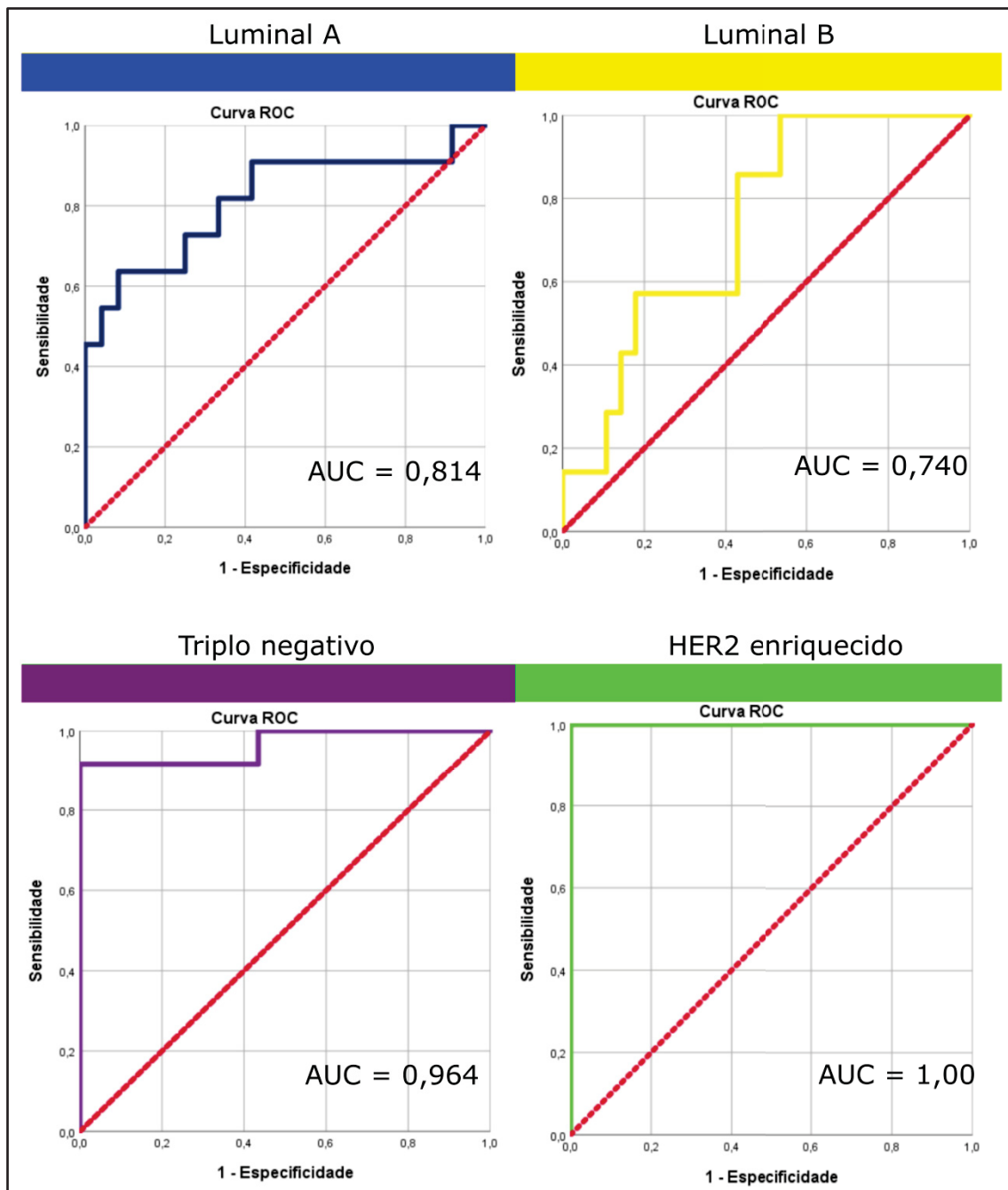
(Legenda: Azul = Dados referentes ao modelo/painel do subtipo Luminal A. Amarelo = Dados referentes ao modelo/painel do subtipo Luminal B. Roxo = Dados referentes ao modelo/painel do subtipo triplamente negativo. Verde = Dados referentes ao modelo/painel do subtipo HER2+ enriquecido. (*) Subtipos nos quais há diferença significativa de expressão das DEPs na comparação entre subtipos, segundo o teste de ANOVA, valor de P-FDR < 0,05).

A nova coorte de tumores mamários foi obtida do trabalho publicado por Krug e colaboradores (2020), sendo todas as amostras classificadas em subtipos imunohistoquímicos através da avaliação do *status* da expressão de ER, PR e HER2.

Os painéis referentes aos subtipos LA, TN e HER2+ obtiveram ótimos valores de AUC quando aplicados à nova coorte de amostras, com AUCs de 0,814, 0,964 e 1,00 respectivamente. O valor de AUC do painel do subtipo HER2 indica, inclusive, que talvez menos de 11 proteínas sejam necessárias para compor este painel caso sejam avaliadas mais amostras. O painel do subtipo LB obteve um valor de AUC relativamente menor em relação aos demais painéis, com um valor de AUC de apenas 0,740. Este valor se justifica devido ao fato de que os tumores do subtipo LB utilizados neste estudo são do subtipo luminal B HER2-, enquanto os tumores oriundos do trabalho de Krug e colaboradores são do subtipo luminal B HER2+, ou seja, apesar de serem do mesmo subtipo, há diferenças entre os tumores LB utilizados para a construção dos painéis em relação aos utilizados para a validação.

No geral, os painéis proteicos construídos nesta pesquisa apresentam um grande potencial para discriminação entre subtipos do câncer de mama, se apresentando como uma opção para a confirmação de diagnósticos realizados a partir dos marcadores clássicos ER, PR, HER2 e Ki-67. Tanto a validação cruzada quanto a validação em uma coorte nova e independente (FIGURA 23) demonstram a alta especificidade e sensibilidade dos painéis, porém devemos considerar que o número reduzido de amostras aqui utilizadas é um fator limitante para a extrapolação da utilidade clínica dos painéis, sendo necessárias mais investigações no sentido de verificar sua verdadeira relevância clínica.

FIGURA 23 - OPERATING CHARACTERISTIC CURVES E VALORES DE AUC OBTIDOS PELOS PAINÉIS PROTEICOS QUANDO APLICADOS A UMA NOVA COORTE DE AMOSTRAS.



(Legenda: Azul = Dados referentes ao modelo/painel do subtipo Luminal A. Amarelo = Dados referentes ao modelo/painel do subtipo Luminal B. Roxo = Dados referentes ao modelo/painel do subtipo triplo negativo. Verde = Dados referentes ao modelo/painel do subtipo HER2+ enriquecido).

8 DISCUSSÃO

O câncer de mama é uma doença complexa e heterogênea que apresenta um grande espectro de fenótipos e comportamentos biológicos e clínicos. Esta heterogeneidade inspirou pesquisadores a buscarem por maneiras de classificar estes tumores em grupos que permitissem aos mesmos serem tratados de maneira mais específica para suas características intrínsecas. Como resultado, a prática clínica atual classifica os tumores mamários em subtipos imunoistoquímicos, onde cada subtipo representa uma unidade patológica que reúne tumores que tendem a apresentar comportamentos e respostas a tratamentos mais uniformes dentro do grupo (GOLDHIRSCH et al., 2013; DAI et al., 2015; FRAGOMENI et al., 2019; NASCIMENTO & OTONI, 2020). Contudo, apesar desta classificação ser altamente relevante, não só as diferenças biológicas entre os subtipos são apenas parcialmente conhecidas, como existem limitações para a classificação baseada apenas nos quatro marcadores clássicos (ER, PR, HER2 e Ki-67). Assim, novas contribuições são necessárias para avançarmos na compreensão das diferenças biológicas entre tumores e subtipos de câncer de mama, assim como novas contribuições no sentido de alcançarmos um padrão mais elevado de biomarcadores de diagnóstico, prognóstico e de discriminação de subtipos, sendo os estudos proteômicos ótimas opções para avançarmos neste objetivo (PANIS et al., 2014; TYANOVA et al., 2016).

Através de tecnologias proteômicas baseadas em espectrometria de massas livre de marcação, esta pesquisa obteve e comparou proteomas de tumores mamários de diferentes subtipos, identificando proteínas que possuem padrões de expressão relacionados com a classificação imunoistoquímica dos tumores. Após a obtenção destas proteínas diferencialmente expressas (DEPs), a relevância biológica das mesmas, no contexto das diferenças clínicas e de comportamento dos subtipos, foi explorada, bem como padrões de interação proteína-proteína. Por fim, através da metodologia de *machine learning*, DEPs com potencial de discriminar os subtipos através de seus padrões de expressão foram selecionadas e painéis proteicos discriminativos de subtipos foram propostos.

A partir da obtenção das DEPs, o primeiro aspecto avaliado foi como os oncogenes e genes supressores de tumor se comportavam em relação à expressão de seus produtos (proteínas) nos diferentes subtipos. Dentre os resultados obtidos, alguns oncogenes e genes supressores de tumor merecem destaque. O oncogene representado pela *Mitogen-Activated Protein Kinase 1 (MAPK1)* foi expresso com muito mais intensidade no subtipo luminal A, enquanto a proteína supressora de tumor *Flap Structure-Specific Endonuclease 1 (FEN1)*, envolvida na manutenção da integridade do genoma (ZHENG et al., 2007) esteve conspicuamente sub expressa neste subtipo em relação aos demais.

A candidata oncoproteína codificada pelo gene *PTPN1* foi encontrada expressa muito mais fortemente em tumores luminais B, concordantemente com outra pesquisa que demonstra que esta DEP é positivamente associada com a expressão de receptor de estrogênio e com tumores do subtipo luminal B (SOYSAL et al., 2013), enquanto a candidata supressora de tumor *RPL22*, uma subunidade ribossomal cuja atividade antitumoral já é determinada em certos tipos de câncer (RAO et al., 2013; CAO et al., 2017), foi encontrada com acentuada sub expressão neste subtipo.

O subtipo triplo negativo foi o que apresentou a maior quantidade de oncogenes e supressores de tumor diferencialmente expressos em relação aos outros subtipos: Os clássicos oncogenes *RHOA*, *CALR* e *IDH2* apresentaram uma expressão mais alta neste subtipo em relação aos demais, juntamente com o candidato oncogene representado pela proteína ribossomal 27 (*RPS27*) que é abordada adiante nesta discussão. Em relação a genes supressores de tumor, apenas genes candidatos foram encontrados com menor expressão neste subtipo em relação aos demais, e mesmo assim com uma margem pequena de sub expressão. O candidato supressor de tumor representado pela proteína Alcohol Dehydrogenase 1B (*ADH1B*) foi o mais sub expresso neste subtipo: *ADH1B* é apresentada por outro trabalho (WU et al., 2012) como sub expressa no câncer de mama, tendo sua baixa expressão associada com pior prognóstico e sobrevida, corroborando com a possibilidade de atuar como supressora de tumor neste subtipo.

O subtipo HER2, por sua vez, apresentou uma acentuada super expressão do oncogene *IDH1*, e dos candidatos a oncogenes *NAMPT*, *DBI* e *APMAP*, este último envolvido em processos inflamatórios em tumores mamários (SUÁREZ-ARROYO et al., 2016), enquanto o candidato supressor de tumor *CCAR1*, associado com indução de apoptose e ativação da proteína supressora de tumor p53 (MUTHU et al., 2015), foi encontrado com evidente menor expressão neste subtipo. Conjuntamente, estes resultados indicam para novas assinaturas proteômicas subtipo-específicas, indicando que alguns oncogenes e gene supressores de tumor podem ter padrões de expressão que contribuem diferencialmente para a tumorigênese mamária dos subtipos.

Através de análises de enriquecimento funcional, as proteínas que possuíam diferenças de expressão nas comparações um-contra-todos foram analisadas em termos de vias e processos biológicos que possuíssem comportamentos característicos em dado subtipo. Assim, vias e/ou processos que possam estar envolvidas nas especificidades biológicas e clínicas dos diferentes subtipos puderam ser indicados. Os subtipos luminais apresentaram em comum a via de endocitose mediada por clatrin, sendo ela ativada no subtipo LA em relação aos subtipos LB, TN e HER2, e desativada no subtipo LB em relação aos demais, indicando que os subtipos luminais, apesar de apresentarem diferentes extremos no comportamento desta via, são mais influenciados pela sua desregulação do que outros subtipos. A via de endocitose mediada por clatrin é uma das mais importantes rotas de endocitose, funcionando através da formação de vesículas revestidas por clatrin que capturam ligantes ou receptores no espaço extra membrana, levando-os para o citoplasma, onde serão degradados ou reciclados, além de estar envolvida na internalização de vesículas extracelulares (SCHMID, 2016). Neste contexto, o subtipo luminal A apresentou, por exemplo, a super expressão da DEP codificada pelo gene *RAB5B*, um importante regulador do processo de endocitose, que em alta expressão contribui também para estimular o citoesqueleto de actina e a motilidade de células tumorais (PALAMIDESSI et al., 2008).

O subtipo luminal B, por sua vez, apresentou uma menor expressão de proteínas relacionadas com o complexo Arp2/3 (*ARPC5*, *ACTR2*, *ARPC3*,

ACTR3 e *ARPC2*, por exemplo), que é necessário para o remodelamento do citoesqueleto de actina e, conseqüentemente, para a viabilização da endocitose dependente de clatrin (PAPALAZAROU & MACHESKY, 2021). A baixa expressão destas DEPs no subtipo LB pode indicar que este subtipo possui uma tendência menor, em relação aos demais subtipos, de apresentar este tipo de endocitose, enquanto a maior ativação relativa da endocitose mediada por clatrin no luminal A pode significar que pacientes com tumores deste subtipo tendem a se responder melhor às drogas inibidoras de endocitose do que tumores de outros subtipos.

O subtipo luminal A, mais especificamente, apresentou uma relação muito íntima com alterações metabólicas como, por exemplo, a ativação do metabolismo de glicose e outros carboidratos, através da maior expressão, em relação aos demais subtipos, das enzimas codificadas pelos genes *ENO2*, *FBP1*, *HK1* e *PGM2*. Alguns autores indicam o subtipo LA como aquele que menos apresenta desregulações metabólicas (CAPPELLETTI et al., 2017; SERRANO-CARBAJAL et al., 2020): Este trabalho, porém, indica que o subtipo luminal A apresenta alterações metabólicas associadas com padrões de expressão proteica que são bastante característicos deste subtipo, como por exemplo, o destas quatro enzimas, contribuindo para a expansão do conhecimento sobre o metabolismo de tumores LA. Outro processo exacerbado neste subtipo em relação aos demais foi o de oxidação biológica, envolvendo a super expressão de enzimas como as codificadas pelos genes *UGDH*, *GSS*, *CNDP2* e *AKR7A2*, todas elas apresentando uma maior expressão no subtipo LA em relação aos demais. A Carnosine Dipeptidase 2 (*CNDP2*) foi encontrada super expressa em câncer de cólon e ovário, atuando como uma indutora de proliferação, crescimento e metástase (CONGLONG et al., 2014; ZHANG et al., 2019), de modo que a super expressão desta proteína pode, desta forma, estar envolvida na tumorigênese de tumores luminal A. A apresentação de antígenos envolvendo MHC de classe 2 também foi uma característica apresentada pelos tumores do subtipo LA, o qual apresentou elevada expressão de moléculas envolvidas neste processo, tais como *HLA-DRA*, *HLA-DRB1*, *DCTN1* e *AP2B1*, e baixa expressão de *CTSB*. Tumores que apresentam expressão de MHC de classe 2 possuem um fenótipo associado com elevada resposta à imunoterapia

contra PD-1 (JOHNSON et al., 2018), indicando a possibilidade de que tumores LA também possuam boa resposta à esta opção de imunoterapia.

O subtipo luminal B, por sua vez, apresentou dois estados alterados bem particulares deste subtipo: o primeiro é a baixa expressão das DEPs codificadas por *ACTR2*, *ACTR3*, *ARPC1B*, *ARPC2*, *ARPC3*, *ARPC4* e *ARPC5*, juntamente com a alta expressão de *MYH2*, todas proteínas relacionadas com o citoesqueleto de actina. Com exceção da cadeia pesada 2 de miosina (*MYH2*), todas estas proteínas participam, direta ou indiretamente, do complexo Arp2/3, uma maquinaria celular conservada evolutivamente que atua ativamente na ramificação, polimerização e movimentação de redes complexas de actina, citosólicas ou associadas a membrana, tomando parte na regulação da protrusão, movimento, cisão e tráfego pela membrana, além de migração celular (MOLINIE & GAUTREAU, 2018). Como consequência da baixa expressão destas proteínas, as vias de sinalização dependentes de efrinas, WASPS e WAVES e fagocitose dependente de receptores TCF-Gama apresentaram uma inativação relativa no subtipo LB. Contudo, isto não significa que tumores LB não apresentam estes processos, mas sim que, em relação a outros subtipos, estas vias parecem ter menor ativação no subtipo LB. O segundo grupo de alterações apresentado por este subtipo foi a menor expressão das DEPs codificadas pelos genes *PSMA5*, *PSMA7*, *PSMB2*, *PSMD4* e *PSMD7*, todas proteínas que compõem subunidades de proteassomos. A alta expressão destas proteínas é frequentemente descrita no câncer e relacionada com piores prognósticos e incidência de metástase (LI et al., 2017; XIA et al., 2019; ZHAO et al., 2020), contudo, o subtipo LB apresenta uma expressão menor destas proteínas em relação a outros subtipos, principalmente em relação ao subtipo TN, indicando que pacientes com tumores do subtipo LB podem não se beneficiar tão conspicuamente de tratamentos inibidores de proteassomos – como o tratamento baseado em tocotrienol gama (RAMDAS et al., 2019).

Em termos funcionais, o subtipo triplo negativo apresentou uma exacerbada ativação de diversas vias em comparação aos demais subtipos. Uma das vias que apresentaram super ativação no subtipo TN foi a transcrição mediada pela RNA polimerase 2, que se apresentou ativada através da super expressão de DEPs como, por exemplo, as codificadas por *MAGOHB*, *SRSF1*,

SRSF7, *SNRPB*, *SARNP* e *SSRP1*, todas envolvidas no *splicing* de RNA. Stricker e colaboradores (2017) demonstraram que o *splicing* pode ocorrer de maneira subtipo-dependente no câncer de mama, com *MAGOHB* tendo uma contribuição importante para a diferenciação de *splicing* em subtipos ER+ em relação ao subtipo TN. A proteína codificada pelo gene *SARNP*, por sua vez, é reconhecidamente super expressa em tumores TN e envolvida no *splicing* e exportação de isoformas, sendo necessária para a manutenção do fenótipo mesenquimal invasor típico destes tumores, sendo inclusive relacionada com supressão da expressão da proteína E-caderina, cuja ausência serve como indicativo de fenótipos invasivos (KANG et al., 2019). Juntas, estas DEPs participam da regulação da transcrição via RNA polimerase 2 e do subsequente *splicing* de mRNAs, contribuindo para o fenótipo agressivo e metastático do subtipo TN.

Outros processos desregulados especialmente em tumores TN incluem os *checkpoints* do ciclo celular, principalmente os da fase G2/M, processos da fase S e de replicação do DNA, indicando que o descontrole do ciclo celular em tumores TN é exacerbado em relação aos demais subtipos. Neste contexto, as proteínas codificadas pelos genes *YWHAE*, *YWHAZ* e *RPS27A* se destacam. Os genes *YWHAE* e *YWHAZ* traduzem proteínas adaptadoras envolvidas, dentre outros processos, no controle do ciclo celular (TENGA & LAZAR, 2013; GAN et al., 2020). Ambas são reconhecidamente super expressas em tumores mamários, participando como indutoras de proliferação através da disfunção dos *checkpoints* do ciclo celular, metástase e, interessantemente, ambas as proteínas são relacionadas com quimioresistência (NISHIMURA et al., 2013; YANG et al., 2018; GAN et al., 2020). A maior expressão destas duas DEPs no subtipo triplo negativo indica uma grande relevância de ambas para o fenótipo mais agressivo e proliferativo deste subtipo, além de implicarem em um dos possíveis mecanismos envolvidos na quimioresistência adquirida por tumores triplo negativos.

A proteína ribossomal S27a (*RPS27A*) atua no câncer não apenas através de suas funções ribossomais, mas também através da regulação de vias relacionadas com o ciclo celular. Em leucemia, a *RPS27A* é capaz de desregular o ciclo celular através da regulação do gene *STAT3*, além de conferir

quimioresistência (WANG et al., 2015), enquanto no câncer de pulmão, esta proteína foi encontrada regulando o ciclo celular, apoptose e proliferação (LI et al., 2020). No câncer de mama, poucos trabalhos exploram o funcionamento desta proteína e sua relação com o ciclo celular (LOPÉZ-CORTÉS et al., 2020), contudo, LIN e colaboradores (2021) destacam a relevância desta proteína em tumores do subtipo TN. Conjuntamente, estes fatos reforçam uma possível alta relevância da *RPS27A* no câncer de mama, mas principalmente para o subtipo triplo negativo.

Por fim, o subtipo HER2+ enriquecido apresentou uma relação estreita com processos como clivagem apoptótica de proteínas, salvamento de nicotinamida e processamento de tRNAs. Foi observada neste subtipo uma menor expressão de *LMNB1*, *CTNNB1* e *ACIN1* e maior expressão de *BCAP31*, todas envolvidas na clivagem apoptótica de proteínas. A laminina 1 (*LMNB1*) é uma proteína que compõem a lâmina nuclear, atuando na manutenção da integridade nuclear, estabilidade cromossômica e apoptose, sendo encontrada sub expressa em tumores mamários, com uma diminuição de expressão que acompanha o ganho de agressividade pelos tumores (WAZIR et al., 2013), enquanto a proteína codificada pelo gene *ACIN1* também possui atividade nuclear associada com integridade da cromatina, sendo necessária para a condensação apoptótica da cromatina (SAHARA et al., 1999). A DEP codificada pelo gene *BCAP31*, pelo contrário, foi encontrada super expressa neste subtipo, concordantemente com estudos anteriores que abordam a super expressão desta proteína no câncer de mama, além de sua participação em processos apoptóticos através da interação com Bcl-2/Bcl-xL, e correlação com pior prognóstico (FU et al., 2019). Segundo nossas análises, estas três DEPs, juntamente com a *CTNNB1*, possuem padrões de expressão típicos para o subtipo HER2, ressaltando a relevância destas DEPs em processos apoptóticos em tumores deste subtipo.

Em relação ao processamento de tRNAs, o subtipo HER2 apresentou uma desregulação mais acentuada deste processo em relação aos demais subtipos, principalmente através da baixa expressão das DEPs codificadas por *DDX1*, *FAM98B* e *RTCB*. Estas três DEPs estão envolvidas no processamento de RNAs, incluindo tRNAs, com alterações na expressão das mesmas podendo

afetar processos apoptóticos e proliferativos (GERMAIN et al., 2011; AKTER et al., 2017; DIREITO et al., 2021), indicando que a desregulação do processamento de tRNAs pode ser um fator relevante para a tumorigênese do subtipo HER2.

Após acessar as diferenças biológicas entre os subtipos, painéis proteicos foram elaborados através de metodologias de *machine learning* baseadas em *supported vector machines* (SVM). Estes painéis possuem como objetivo permitir que a classificação em subtipos imunoistoquímicos seja mais precisa e específica, atuando como modelos de confirmação da classificação baseada nos marcadores clássicos ER, PR, HER2 e Ki-67. O painel proposto para a confirmação da classificação do subtipo LA foi composto pelas DEPs codificadas pelos genes *IQGAP2*, *DDX1*, *HP1BP3*, *NDUFS1* e *TPP1*, todas marcadoras positivas para a classificação de tumores no subtipo luminal A. Estas proteínas não apenas possuem um padrão de expressão típico de tumores deste subtipo, mas também fazem parte da tumorigênese destes tumores. A Tripeptidyl Peptidase 1 (*TPP1*), por exemplo, que foi encontrada com maior expressão no subtipo luminal A do que em qualquer outro subtipo, possui uma relação íntima com a expressão dos receptores ER e PR (JUNAID et al., 2000) que são marcadores altamente expressos no subtipo LA. A aplicação deste painel para classificação do subtipo de nossos tumores mamários obteve uma sensibilidade de 100%, especificidade de 92,8% e taxa de falso-positivos de apenas 7,1%, demonstrando a robustez deste painel para confirmar a classificação de tumores do subtipo LA. Quando aplicado em uma coorte de validação independente, o painel manteve seu bom desempenho, alcançando um valor de AUC de 0,814.

Outro painel proteico formado por cinco DEPs foi proposto para o subtipo luminal B, sendo constituído apenas por marcadoras negativas (*GSTP1*, *PDIA5*, *LGALS3*, *S100A10* e *EPB41L2*) com baixa expressão neste subtipo em relação aos demais. A proteína Glutathione S-Transferase Pi 1 (*GSTP1*), por exemplo, possui na literatura diversos apontamentos de sua menor expressão no subtipo luminal B: *GSTP1* é tipicamente super expressa em tumores ER negativos, como os tumores HER2 ou TN e, imunoistoquimicamente, tumores luminais B demonstram menor positividade para expressão de *GSTP1* do que outros subtipos (MIYAKE et al., 2012). Além disso, o subtipo luminal B é aquele que

apresenta maior hipermetilação no gene *GSTP1*, o que contribui para a supressão da expressão deste gene no subtipo LB e para o fenótipo agressivo deste tumor ER positivo (MIYAKE et al., 2013). Outra proteína deste painel, a proteína S100-A10 (*S100A10*), é descrita como sub expressa em tumores mamários em relação a tecidos não tumorais (CARLSSON et al., 2005), sendo que nossas análises demonstram uma ainda mais acentuada sub expressão desta proteína no subtipo LB. Além disso, Grosset e colaboradores (2016) já haviam indicado o subtipo LB como o subtipo com menor expressão do mRNA de *LGALS3*, sendo que esta pesquisa corrobora estes dados demonstrando uma menor expressão proteica da mesma no subtipo LB.

Este painel proteico obteve um ótimo desempenho para discriminar tumores LB de tumores dos demais subtipos, alcançando uma especificidade de 93,3% e sensibilidade de 75%, com apenas 6,8% de taxa de falso-positivos. Quando este painel foi aplicado a uma coorte independente de amostras, a capacidade de discriminação do painel se mostrou reduzida, alcançando apenas 0,740 como valor de AUC. Contudo, vale destacar que os tumores utilizados para a montagem deste painel eram do subtipo luminal B HER2-, enquanto os tumores da coorte independente eram tumores luminal B HER2+, o que pode explicar o menor desempenho do painel nesta coorte independente.

Composto por duas DEPs marcadoras positivas (*FSCN1* e *FABP5*) e três marcadoras negativas (*PTRF*, *AOC3* e *ADH1B*), o painel proposto para discriminação do subtipo triplo negativo foi o que apresentou o melhor desempenho, seja na validação cruzada realizada nas amostras deste estudo, seja na validação realizada na coorte independente. A taxa de falso-positivos deste painel foi de 0%, com especificidade de 100% e sensibilidade de 83,3%, sendo o valor de AUC da validação na coorte independente de 0,964. Assim como ocorrido para os outros painéis, as DEPs que compõem o painel do subtipo TN apresentam uma relação estreita com a biologia destes tumores. A proteína fascina (*FSCN1*), por exemplo, é uma proteína de ligação ao citoesqueleto de actina cuja alta expressão está envolvida com ativação de movimentos celulares e metástase, além de ser reconhecidamente (WANG et al., 2017) associada com tumores mamários ER/PR negativos, fenótipos mais agressivos e, em populações africanas e chinesas, sua alta expressão já foi associada

diretamente com tumores TN, sendo inclusive sugerida como marcadora de diagnóstico (WANG et al., 2016).

A proteína *Fatty acid binding protein 5 (FABP5)*, uma proteína intracelular de ligação à ácidos graxos, assim como a fascina, é reconhecida por sua associação com tumores TN e piores prognósticos, atuando no estímulo de crescimento tumoral e metástase através de sua ação como estabilizadora do receptor *EGFR* (POWELL et al., 2015). Liu e colaboradores (2011) inclusive demonstram que esta proteína é preferencialmente expressa em tumores ER/PR negativos, sendo relacionada com tumores de alto grau histológico e pior prognóstico. A proteína caveolina 1 (*PTRF/CAVIN1*) também constitui o painel deste subtipo, apresentando uma acentuada baixa expressão no subtipo TN. A caveolina 1 é reconhecida como uma supressora de tumor em câncer de próstata (AN et al., 2020), e alguns estudos já apontam sua baixa expressão em tumores e linhagens de câncer de mama, bem como seu potencial como alvo terapêutico (BAI et al., 2011; DÍAZ et al., 2020).

O subtipo HER2+ enriquecido se apresentou, durante toda esta pesquisa, como o subtipo com proteoma mais intermediário em relação aos demais subtipos, tendo em seu proteoma o menor número de proteínas com diferença de expressão na comparação com outros subtipos, e também o menor número de vias/processos diferencialmente enriquecidos, enfatizando o fato de que este subtipo possui um fenótipo com características intermediárias em relação aos demais subtipos. Esta observação é reforçada pelo fato que o painel proposto para este subtipo necessitou de 11 DEPs para alcançar um desempenho aceitável em seu objetivo de discriminar tumores deste subtipo dos demais tumores, enquanto para os outros painéis cinco DEPs foram o suficiente. O painel deste subtipo obteve um bom desempenho, com 0% de taxa de falso-positivos, sensibilidade de 75% e especificidade de 100% e, na validação na coorte independente, o valor de AUC foi de 1,00, o que pode indicar que, na possibilidade de contar com mais amostras, este painel pode até mesmo ser reduzido a um menor número de DEPs.

As DEPs que compõem este painel não discriminam, individualmente, tumores HER2 dos demais, uma vez que nenhuma DEP do subtipo HER2

apresentou diferença de expressão simultânea a todos os outros subtipos, ou seja, algumas DEPs discriminam tumores HER2 de tumores LA, outras discriminam tumores HER2 de tumores TN e assim por diante, sendo os padrões de expressão combinados destas 11 DEPs necessários para discriminar o subtipo HER2 dos demais. Nove marcadoras negativas compõem este painel (*RAB1A*, *MYH2*, *HMGCL*, *CCAR1*, *HK1*, *FAM98B*, *RTCB*, *DDX1* e *TROVE2*) juntamente com duas marcadoras positivas (*NAMPT* e *DBI*). A proteína *Nicotinamide Phosphoribosyltransferase (NAMPT)* apresentou, segundo nossas análises de enriquecimento, uma relação com o subtipo HER2 através da indução de uma maior intensidade da via de salvamento de nicotinamida neste subtipo. Tendo maior expressão neste subtipo em relação aos outros (mas não significativamente em relação ao subtipo TN), *NAMPT* foi uma marcadora positiva do subtipo HER2, sendo inclusive anteriormente associada positivamente com a expressão de HER2 e se apresentando como um potencial alvo terapêutico (ZHU et al., 2015; HESKE, 2020).

9 CONCLUSÃO

Através da espectrometria de massas, este estudo quantificou e identificou as proteínas que compõem o proteoma de tumores de diferentes subtipos do câncer de mama, apresentando proteínas diferencialmente expressas entre estes subtipos, as quais obtiveram sua significância biológica determinada através de análises de bioinformática. Além disso, painéis proteicos com alto potencial para discriminação de subtipos foram propostos para os subtipos luminais A e B, triplo negativo e HER2+ enriquecido.

Os resultados obtidos através de análises de enriquecimento funcional nos permitiram avaliar vias e processos biológicos diferencialmente ativados em cada subtipo, além de DEPs que aparentemente possuem funções mais relevantes e impactantes em determinados subtipos. Cada subtipo do câncer de mama possui um comportamento biológico e clínico diferenciado, e através de nossas análises, novos elementos envolvidos nesta diferenciação puderam ser revelados e avaliados.

A tecnologia de *machine learning* pautada em *Support Vector Machines* (SVMs) foi utilizada para o ranqueamento das DEPs de cada subtipo segundo seu potencial discriminativo entre subtipos, resultando no apontamento de quatro painéis proteicos – um para cada subtipo – os quais obtiveram um desempenho satisfatório na validação cruzada dos modelos de classificação e também na validação em uma coorte independente, constituindo-se de boas opções de painéis proteicos de diagnóstico para o câncer de mama, apesar de futuros estudos confirmatórios ainda serem necessários.

10 ANEXOS

TABELA SUPLEMENTAR 1 – PARÂMETROS UTILIZADOS PARA EXECUÇÃO DAS CORRIDAS DE LC-MS/MS

Ion Source
Positive Ion (V): 2300
Ion Transfer Tube Temp (°C): 175
Master Scan (MS1)
Detector Type: Orbitrap
Orbitrap Resolution: 120000
Use Quadrupole Isolation: True
Scan Range (m/z): 300-1500
RF Lens (%): 30
AGC Target: standard
Maximum Injection Time (ms): 50
Microscans: 1
Data Type: Profile
Polarity: Positive
Use EASY-IC™ (Internal calibration): True
Advanced Peak Determination: True
Intensity
Filter Type: Intensity Threshold
Intensity Threshold for MS2: 2.0e4
Charge State
Include charge state(s): 2-7
Include undetermined charge states: False
Include charge states 25 and higher: False
Dynamic Exclusion
Exclude after n times: 1
Exclusion duration (s): 60
Mass Tolerance: ppm
Low: 10

High: 10
Exclude Isotopes: True
Data Dependent
Data Dependent Mode: Cycle Time
Time between Master Scans (sec): 2
ddMS² OT HCD (MS2)
Isolation Mode: Quadrupole
Isolation Window (m/z): 1.6
Isolation Offset: Off
Activation Type: HCD
HCD Collision Energy (%): 30
Detector Type: Orbitrap
Scan Range Mode: Auto: m/z Normal
Orbitrap Resolution: 15000
AGC Target: standard
Maximum Injection Time (ms): 22
Microscans: 1
Data Type: Centroid

TABELA SUPLEMENTAR 2 – DEPS RESULTANTES DA ANÁLISE DE ANOVA NAS COMPARAÇÕES ENTRE SUBTIPOS, E IDENTIFICAÇÃO DOS PARES ONDE A DIFERENÇA DE EXPRESSÃO É SIGNIFICATIVA ATRAVÉS DO PÓS-TESTE DE TUKEY.

Símbolo do gene	Nome da proteína	Valor de P-FDR	Pares de subtipos com expressão diferencial
<i>ABCD3</i>	ATP-binding cassette sub-family D member 3	0,0377412	T LA_T TN; T LB_T TN
<i>ABCF1</i>	ATP-binding cassette sub-family F member 1	0,0288	T TN_T LB; T TN_T LA; T TN_T HER2
<i>ADH1B</i>	Alcohol dehydrogenase 1B	0,019	T LA_T TN; T LB_T TN; T HER2_T TN
<i>AHNAK</i>	Neuroblast differentiation-associated protein AHNAK	0,0409796	T LA_T TN
<i>ANP32E</i>	Acidic leucine-rich nuclear phosphoprotein 32 family member E	0,018	T TN_T LB; T TN_T HER2; T TN_T LA
<i>ANXA11</i>	Annexin A11	0,0421947	T LA_T TN
<i>ANXA7</i>	Annexin A7	0,0269333	T LA_T HER2; T LA_T TN
<i>AOC3</i>	Membrane primary amine oxidase	> 0,0001	T LB_T TN; T LA_T TN; T HER2_T TN
<i>APEH</i>	Acylamino-acid-releasing enzyme	0,0247	T LB_T TN; T LA_T TN
<i>ARFIP1</i>	Arfaptin-1	0,0428	T LA_T TN; T LB_T TN
<i>ARMT1</i>	Protein-glutamate O-methyltransferase	0,0196364	T LB_T TN; T LA_T TN; T LB_T HER2; T LA_T HER2
<i>ARPC5</i>	Actin-related protein 2/3 complex subunit 5	0,0371818	T TN_T LB
<i>CALR</i>	Calreticulin	0,0265143	T TN_T LB; T TN_T LA
<i>CAV1</i>	Caveolin-1	0,0262069	T LB_T TN; T HER2_T TN; T LA_T TN
<i>CCAR1</i>	Cell division cycle and apoptosis regulator protein 1	0,0408247	T LA_T HER2; T LA_T TN
<i>CD59</i>	CD59 glycoprotein	0,0185882	T TN_T LB; T TN_T LA; T TN_T HER2
<i>COLGALT1</i>	Procollagen galactosyltransferase 1	0,0375802	T TN_T LA; T TN_T LB
<i>CSTB</i>	Cystatin-B	0,0293939	T TN_T LB; T TN_T LA
<i>DAD1</i>	Dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit DAD1	0,0367568	T TN_T LB; T TN_T LA

<i>DAK</i>	Bifunctional ATP-dependent dihydroxyacetone kinase/FAD-AMP lyase (cyclizing)	0,0275918	T LB_T TN; T LA_T TN
<i>DBI</i>	Acyl-CoA-binding protein	0,0222857	T TN_T LB; T HER2_T LB; T TN_T LA; T HER2_T LA
<i>DCPS</i>	m7GpppX diphosphatase	0,0383297	T TN_T LB
<i>DDX1</i>	ATP-dependent RNA helicase DDX1	0,0264516	T LA_T HER2; T LA_T LB; T LA_T TN
<i>EEF1A1</i>	Elongation factor 1-alpha 1	0,0275238	T TN_T LA; T TN_T HER2; T TN_T LB
<i>EEF1D</i>	Elongation factor 1-delta	0,0424299	T TN_T LA
<i>EEF1G</i>	Elongation factor 1-gamma	0,0387391	T TN_T LA; T TN_T LB
<i>EHD2</i>	EH domain-containing protein 2	0,0186667	T HER2_T TN; T LB_T TN; T LA_T TN
<i>EIF4A1</i>	Eukaryotic initiation factor 4A-I	0,0382857	T TN_T LA; T TN_T LB
<i>EIF4H</i>	Eukaryotic translation initiation factor 4H	0,04072	T TN_T LB
<i>EIF5A</i>	Eukaryotic translation initiation factor 5A-1	0,0329412	T TN_T LB; T TN_T LA
<i>EPB41L2</i>	Band 4.1-like protein 2	0,0380952	T TN_T LB; T LA_T LB; T HER2_T LB
<i>ESYT1</i>	Extended synaptotagmin-1	0,0382169	T LA_T TN; T LB_T TN
<i>FABP4</i>	Fatty acid-binding protein, adipocyte	0,0286809	T LA_T TN; T LB_T TN; T HER2_T TN
<i>FABP5</i>	Fatty acid-binding protein, epidermal	0,0184	T TN_T LB; T TN_T LA; T TN_T HER2
<i>FAM98B</i>	Protein FAM98B	0,029194	T LA_T HER2
<i>FBP1</i>	Fructose-1,6-bisphosphatase 1	0,0257778	T LB_T TN; T LA_T TN
<i>FKBP1A</i>	Peptidyl-prolyl cis-trans isomerase FKBP1A; Peptidyl-prolyl cis-trans isomerase	0,0263396	T TN_T LB; T TN_T LA
<i>FSCN1</i>	Fascin	0,01975	T TN_T LB; T TN_T HER2; T TN_T LA
<i>GSS</i>	Glutathione synthetase	0,0288387	T LB_T TN; T LA_T TN
<i>GSTP1</i>	Glutathione S-transferase P	0,0404356	T LA_T LB; T TN_T LB
<i>H2AFY</i>	Core histone macro-H2A.1	0,0379024	T LA_T TN; T LA_T HER2
<i>HEBP2</i>	Heme-binding protein 2	0,0487541	T TN_T LB; T TN_T LA

<i>HK1</i>	Hexokinase-1	0,0196667	T LB_T HER2; T LA_T HER2; T LB_T TN; T LA_T TN
<i>HMGCL</i>	Hydroxymethylglutaryl-CoA lyase, mitochondrial	0,0255789	T LA_T HER2; T LA_T TN
<i>HP1BP3</i>	Heterochromatin protein 1-binding protein 3	0,0409293	T LA_T TN; T LA_T HER2; T LA_T LB
<i>HSPE1</i>	10 kDa heat shock protein, mitochondrial	0,0269508	T TN_T LA; T TN_T LB
<i>HYOU1</i>	Hypoxia up-regulated protein 1	0,0251429	T TN_T LB; T TN_T LA; T TN_T HER2
<i>IDH2</i>	Isocitrate dehydrogenase [NADP], mitochondrial	0,0374382	T TN_T LA; T TN_T LB
<i>IFI16</i>	Gamma-interferon-inducible protein 16	0,0372889	T TN_T LB; T TN_T LA
<i>IQGAP2</i>	Ras GTPase-activating-like protein IQGAP2	0,047453	T LA_T TN; T LA_T LB; T LA_T HER2
<i>LDHB</i>	L-lactate dehydrogenase B chain;L-lactate dehydrogenase	0,0275714	T TN_T LA; T TN_T LB
<i>LGALS3</i>	Galectin-3;Galectin	0,0177143	T LA_T LB; T HER2_T LB; T TN_T LB
<i>LMNB2</i>	Lamin-B2	0,02528	T LB_T TN; T LA_T TN; T LA_T HER2
<i>LRP1</i>	Prolow-density lipoprotein receptor-related protein 1	0,0425714	T HER2_T TN; T LA_T TN
<i>MANF</i>	Mesencephalic astrocyte-derived neurotrophic factor	0,0427027	T TN_T LB; T TN_T LA
<i>MOGS</i>	Mannosyl-oligosaccharide glucosidase	0,0387234	T TN_T LB; T TN_T LA
<i>MVP</i>	Major vault protein	0,0235	T LA_T TN; T LA_T LB
<i>MYH2</i>	Myosin-2	0,0478655	T LA_T HER2; T LB_T HER2; T LB_T TN
<i>MYOF</i>	Myoferlin	0,0178947	T LB_T TN; T LA_T TN; T LA_T HER2
<i>NAMPT</i>	Nicotinamide phosphoribosyltransferase	0,025	T TN_T LB; T HER2_T LB; T HER2_T LA
<i>NDUFS1</i>	NADH-ubiquinone oxidoreductase 75 kDa subunit, mitochondrial	0,016	T LB_T TN; T LA_T TN; T LA_T HER2; T LA_T LB
<i>PDIA4</i>	Protein disulfide-isomerase A4	0,0387368	T TN_T LB; T TN_T LA
<i>PDIA5</i>	Protein disulfide-isomerase A5	0,0210667	T LA_T LB; T HER2_T LB; T TN_T LB
<i>PDIA6</i>	Protein disulfide-isomerase A6	0,0258824	T TN_T LA; T TN_T LB; T TN_T HER2

<i>PFKP</i>	ATP-dependent 6-phosphofructokinase, platelet type	0,0252973	T HER2_T LB; T TN_T LB; T TN_T LA
<i>PFN1</i>	Profilin-1	0,0278462	T TN_T LB; T TN_T LA; T TN_T HER2
<i>PGRMC1</i>	Membrane-associated progesterone receptor component 1	0,0379747	T HER2_T LB; T TN_T LB
<i>PHGDH</i>	D-3-phosphoglycerate dehydrogenase	0,0254634	T TN_T LB; T TN_T LA
<i>PLEC</i>	Plectin	0,0280833	T HER2_T TN; T LA_T TN; T LB_T TN
<i>PPIB</i>	Peptidyl-prolyl cis-trans isomerase B	0,0216	T TN_T LA; T TN_T LB; T TN_T HER2
<i>PRDX4</i>	Peroxiredoxin-4	0,0251282	T TN_T LB; T TN_T LA
<i>PRKAR2A</i>	cAMP-dependent protein kinase type II-alpha regulatory subunit	0,0386237	T LA_T TN; T LA_T LB
<i>PSMD4</i>	26S proteasome non-ATPase regulatory subunit 4	0,0275556	T LA_T LB; T TN_T LB; T TN_T HER2
<i>PTRF</i>	Polymerase I and transcript release factor	0,0190769	T LB_T TN; T HER2_T TN; T LA_T TN
<i>RAB1A</i>	Ras-related protein Rab-1A	0,0466207	T TN_T HER2
<i>RAN</i>	GTP-binding nuclear protein Ran	0,0328116	T TN_T LA; T TN_T LB; T TN_T HER2
<i>RBM8A</i>	RNA-binding protein 8A	0,0333143	T TN_T LB; T TN_T LA; T TN_T HER2
<i>RNPEP</i>	Aminopeptidase B	0,0266207	T LB_T TN; T LA_T TN; T LA_T HER2
<i>RPL11</i>	60S ribosomal protein L11	0,0379487	T TN_T LA; T TN_T LB
<i>RPL13A</i>	60S ribosomal protein L13a	0,0427925	T TN_T LB; T TN_T LA
<i>RPL22</i>	60S ribosomal protein L22	0,0266667	T HER2_T LB; T TN_T LB; T TN_T LA; T TN_T HER2
<i>RPL23</i>	60S ribosomal protein L23	0,0266667	T TN_T LB; T TN_T LA; T TN_T HER2
<i>RPL23A</i>	60S ribosomal protein L23a	0,0437913	T TN_T LB; T TN_T LA
<i>RPL27</i>	60S ribosomal protein L27	0,0284444	T TN_T LB; T TN_T LA
<i>RPL3</i>	60S ribosomal protein L3	0,0482333	T TN_T LB; T TN_T LA
<i>RPL30</i>	60S ribosomal protein L30	0,02712	T TN_T LA; T TN_T LB; T TN_T HER2
<i>RPL32</i>	60S ribosomal protein L32	0,0326111	T TN_T LB; T TN_T LA

<i>RPL35</i>	60S ribosomal protein L35	0,0406667	T TN_T LA; T TN_T LB
<i>RPL35A</i>	60S ribosomal protein L35a	0,048595	T TN_T LB; T TN_T LA
<i>RPL38</i>	60S ribosomal protein L38	0,0430459	T TN_T LB; T TN_T LA
<i>RPS12</i>	40S ribosomal protein S12	0,0261818	T TN_T LB; T TN_T LA
<i>RPS13</i>	40S ribosomal protein S13	0,0272	T TN_T LB; T TN_T LA; T TN_T HER2
<i>RPS15A</i>	40S ribosomal protein S15a	0,0268462	T TN_T LB; T TN_T LA
<i>RPS19</i>	40S ribosomal protein S19	0,0285217	T TN_T LA; T TN_T LB; T TN_T HER2
<i>RPS20</i>	40S ribosomal protein S20	0,028	T TN_T LA; T TN_T LB
<i>RPS21</i>	40S ribosomal protein S21	0,0424762	T TN_T LB; T TN_T LA
<i>RPS25</i>	40S ribosomal protein S25	0,0404632	T TN_T LB; T TN_T LA
<i>RPS26</i>	40S ribosomal protein S26; Putative 40S ribosomal protein S26-like 1	0,0481017	T TN_T LA; T TN_T LB
<i>RPS27</i>	40S ribosomal protein S27	0,0373333	T TN_T LA; T TN_T LB
<i>RPS28</i>	40S ribosomal protein S28	0,0271429	T TN_T LB; T TN_T LA; T TN_T HER2
<i>RTCB</i>	tRNA-splicing ligase RtcB homolog	0,0265882	T LA_T HER2; T LA_T TN; T LA_T LB
<i>S100A10</i>	Protein S100-A10	0,0260339	T LA_T LB; T TN_T LB; T HER2_T LB
<i>S100A11</i>	Protein S100-A11	0,0174	T HER2_T LB; T TN_T LB; T TN_T LA
<i>S100A6</i>	Protein S100-A6; Protein S100	0,0424815	T TN_T LB; T TN_T LA
<i>S100A9</i>	Protein S100-A9	0,0341918	T HER2_T LA; T TN_T LA
<i>SEC61A1</i>	Protein transport protein Sec61 subunit alpha isoform 1	0,0424231	T TN_T LA; T TN_T LB
<i>SNRPD1</i>	Small nuclear ribonucleoprotein Sm D1	0,0266667	T TN_T LB; T TN_T HER2; T TN_T LA
<i>SNX3</i>	Sorting nexin-3	0,0173333	T HER2_T LB; T TN_T LB; T TN_T LA; T TN_T HER2
<i>SNX9</i>	Sorting nexin-9	0,0426316	T LA_T TN; T LB_T TN
<i>SRPRB</i>	Signal recognition particle receptor subunit beta	0,0405	T TN_T LA; T TN_T LB
<i>SSR1</i>	Translocon-associated protein subunit alpha	0,0404272	T TN_T LB; T TN_T LA
<i>SSR4</i>	Translocon-associated protein subunit delta	0,0267333	T TN_T LB; T TN_T LA; T TN_T HER2

<i>TAGLN2</i>	Transgelin-2	0,02625	T TN_T LB; T TN_T LA; T TN_T HER2
<i>TPD52L2</i>	Tumor protein D54	0,03785	T TN_T LB;T TN_T LA
<i>TPP1</i>	Tripeptidyl-peptidase 1	0,0231111	T LA_T TN;T LA_T HER2;T LA_T LB
<i>TROVE2</i>	60 kDa SS-A/Ro ribonucleoprotein	0,0373953	T LA_T HER2;T LA_T LB
<i>TSN</i>	Translin	0,0236522	T TN_T LB;T TN_T HER2;T TN_T LA
<i>UBE2I</i>	SUMO-conjugating enzyme UBC9	0,0373333	T TN_T LB;T TN_T LA
<i>UFC1</i>	Ubiquitin-fold modifier-conjugating enzyme 1	0,0275455	T TN_T LB;T TN_T LA
<i>UGP2</i>	UTP--glucose-1-phosphate uridylyltransferase	0,0265926	T TN_T LB;T TN_T LA
<i>VCAN</i>	Versican core protein	0,0273488	T LA_T TN

TABELA SUPLEMENTAR 3 – DEPS IDENTIFICADAS, PARA CADA SUBTIPO, ATRAVÉS DA ABORDAGEM DE UM-CONTRA-TODOS. (A) DEPS DA COMPARAÇÃO DE TUMORES LUMINAL A X NÃO-LUMINAL A. (B) DEPS DA COMPARAÇÃO DE TUMORES LUMINAL B X NÃO-LUMINAL B. (C) DEPS DA COMPARAÇÃO DE TUMORES TRIPLO NEGATIVO X NÃO-TRIPLO NEGATIVO. (D) DEPS DA COMPARAÇÃO DE TUMORES HER2+ ENRIQUECIDO X NÃO-HER2+ ENRIQUECIDO.

TABELA SUPLEMENTAR 3A

Símbolo dos genes	Média de expressão nos tumores LA (<i>Log-Z-score LQF</i>)	Média de expressão nos tumores não - LA (<i>Log-Z-score LQF</i>)	Log-Fold Change	Valor de P
<i>ACTR1A</i>	0,785469	-0,280525	1,06599	0,0365146
<i>ADD1</i>	0,930008	-0,425442	1,35545	0,00627842
<i>AHNAK</i>	0,984224	-0,351509	1,33573	0,00613602
<i>AKR7A2</i>	0,934066	-0,333595	1,26766	0,0102514
<i>ALDH16A1</i>	0,596299	-0,869086	1,46539	0,0136693
<i>ALDH18A1</i>	-0,955825	0,341366	-1,29719	0,00825503
<i>ANXA11</i>	1,00746	-0,359806	1,36726	0,00475337
<i>ANXA6</i>	0,76502	-0,273221	1,03824	0,042443
<i>ANXA7</i>	1,07777	-0,384919	1,46269	0,00202868
<i>AP2B1</i>	0,744461	-0,265879	1,01034	0,0491215
<i>APEH</i>	1,03124	-0,368301	1,39954	0,00361379
<i>APPL1</i>	0,678349	-0,563022	1,24137	0,0224382
<i>ARMT1</i>	0,903917	-0,572664	1,47658	0,00541024
<i>ASAH1</i>	0,765683	-0,273458	1,03914	0,0422399
<i>ATP6V1C1</i>	0,599905	-0,593544	1,19345	0,0385617
<i>BCL2L13</i>	0,558307	-0,736905	1,29521	0,0329146
<i>CAPN1</i>	0,746984	-0,26678	1,01376	0,0482612
<i>CAPZA2</i>	0,856884	-0,30603	1,16291	0,0206751
<i>CAST</i>	0,71249	-0,534722	1,24721	0,0299698
<i>CCAR1</i>	0,87709	-0,865811	1,7429	0,00227963
<i>CNDP2</i>	0,753839	-0,269228	1,02307	0,0459823
<i>COASY</i>	0,841465	-0,606332	1,4478	0,0119175
<i>COLGALT1</i>	-1,46289	-0,140257	-1,32263	0,0295434
<i>COPS2</i>	0,741569	-0,566101	1,30767	0,0249043
<i>COPS4</i>	1,10346	-0,394092	1,49755	0,00143715
<i>CORO7</i>	0,852191	-0,621048	1,47324	0,00517608
<i>CRYZ</i>	0,876793	-0,313141	1,18993	0,0174137
<i>CTSB</i>	-0,889568	0,317703	-1,20727	0,0155466
<i>CYCS</i>	-0,810661	0,289522	-1,10018	0,0301154
<i>CYFIP1</i>	0,878259	-0,313664	1,19192	0,0171909
<i>DAK</i>	0,892031	-0,43586	1,32789	0,00979038

<i>DBNL</i>	0,777261	-0,277593	1,05485	0,0388117
<i>DCTN1</i>	0,94739	-0,338354	1,28574	0,00898759
<i>DDX1</i>	1,22496	-0,437486	1,66245	0,000203839
<i>DDX21</i>	-0,981776	0,209763	-1,19154	0,0281983
<i>DNAJC8</i>	0,431537	-0,762287	1,19382	0,0438855
<i>DPP3</i>	0,875855	-0,312805	1,18866	0,0175576
<i>DPP7</i>	0,693089	-0,436055	1,12914	0,0309102
<i>DPYSL3</i>	0,775722	-0,277044	1,05277	0,0392547
<i>EEF1D</i>	-1,05788	0,377813	-1,43569	0,00261507
<i>EEF1G</i>	-0,744708	0,265967	-1,01068	0,0490366
<i>EHD2</i>	0,792302	-0,282965	1,07527	0,0346835
<i>EIF3I</i>	-1,471	0,221147	-1,69215	0,00206026
<i>ENO2</i>	0,618883	-0,561142	1,18002	0,0343084
<i>EZR</i>	0,753824	-0,269223	1,02305	0,0459872
<i>FAH</i>	0,724307	-0,382872	1,10718	0,0390019
<i>FAM120A</i>	0,710623	-0,35952	1,07014	0,0421509
<i>FAM98B</i>	0,829973	-0,789228	1,6192	0,00382349
<i>FBN1</i>	0,523086	-0,631428	1,15451	0,0444633
<i>FBP1</i>	0,851759	-0,3042	1,15596	0,0215882
<i>FEN1</i>	-1,27128	-0,0296154	-1,24166	0,0317095
<i>FLOT1</i>	0,888319	-0,317257	1,20558	0,0157217
<i>FLOT2</i>	1,03647	-0,660647	1,69712	0,00163095
<i>GDI1</i>	0,790331	-0,282261	1,07259	0,0352043
<i>GGT5</i>	0,7961	-0,529113	1,32521	0,0157848
<i>GLOD4</i>	0,847817	-0,302792	1,15061	0,0223121
<i>GNA11</i>	0,475982	-0,826468	1,30245	0,0336218
<i>GNAI2</i>	1,00289	-0,358175	1,36106	0,0050028
<i>GNAI3</i>	0,740979	-0,465836	1,20682	0,043397
<i>GNB2</i>	0,871673	-0,439853	1,31153	0,0122774
<i>GPS1</i>	0,91211	-0,579577	1,49169	0,00523555
<i>GSS</i>	0,802243	-0,384197	1,18644	0,0208198
<i>H1FO</i>	1,04287	-0,372453	1,41532	0,00314497
<i>H2AFY</i>	1,16583	-0,416368	1,5822	0,000568182
<i>HDHD2</i>	0,734632	-0,960133	1,69477	0,0055537
<i>HK1</i>	0,978351	-0,349411	1,32776	0,00653307
<i>HLA-A</i>	-0,764853	0,273162	-1,03801	0,0424941
<i>HLA-DRA</i>	1,02352	-0,365544	1,38907	0,0039558
<i>HLA-DRB1</i>	0,840518	-0,527549	1,36807	0,0094907
<i>HMGCL</i>	0,826221	-0,892181	1,7184	0,00151255
<i>HNRNPH2</i>	0,881177	-0,654204	1,53538	0,00342953
<i>HP1BP3</i>	1,23331	-0,440469	1,67378	0,000173948
<i>HSDL2</i>	0,635297	-0,750614	1,38591	0,0183618
<i>HSP90AB1</i>	-0,824418	0,294435	-1,11885	0,027011
<i>HSPA4</i>	1,06574	-0,38062	1,44636	0,0023686
<i>HSPD1</i>	-0,802718	0,286685	-1,0894	0,0320304

<i>HSPE1</i>	-0,772697	0,275963	-1,04866	0,0401366
<i>IDH2</i>	-0,91578	0,327064	-1,24284	0,0122152
<i>IMMT</i>	0,87053	-0,310904	1,18143	0,0183922
<i>IQGAP1</i>	0,822125	-0,293616	1,11574	0,0275102
<i>IQGAP2</i>	1,03387	-0,914067	1,94793	0,000303979
<i>KCTD12</i>	0,902985	-0,322495	1,22548	0,0137611
<i>KPNB1</i>	0,79227	-0,282954	1,07522	0,034692
<i>LANCL1</i>	0,760573	-0,422411	1,18298	0,032012
<i>LDHB</i>	-0,753014	0,268934	-1,02195	0,046252
<i>LMNB2</i>	1,06483	-0,380298	1,44513	0,0023959
<i>LRP1</i>	0,713362	-0,414493	1,12786	0,0452011
<i>LRRC47</i>	0,915068	-0,32681	1,24188	0,0122974
<i>LRRC59</i>	-0,815464	0,291237	-1,1067	0,0290015
<i>LTA4H</i>	0,831134	-0,296834	1,12797	0,0255896
<i>MAPK1</i>	0,799915	-0,285684	1,0856	0,0327283
<i>MATR3</i>	0,857047	-0,306088	1,16314	0,0206465
<i>MCCC2</i>	0,43694	-0,843014	1,27995	0,0411569
<i>MVP</i>	1,15351	-0,411969	1,56548	0,000690162
<i>MYO1C</i>	0,922968	-0,329631	1,2526	0,0114101
<i>MYOF</i>	1,037	-0,370356	1,40735	0,00337499
<i>NAPRT</i>	0,514323	-0,675358	1,18968	0,0418494
<i>NDUFA9</i>	0,94733	-0,338332	1,28566	0,00899299
<i>NDUFS1</i>	1,31929	-0,471174	1,79046	2,63E-05
<i>NDUFS2</i>	0,994035	-0,355012	1,34905	0,00551683
<i>NDUFV1</i>	0,934494	-0,333748	1,26824	0,0102086
<i>NIT2</i>	0,743283	-0,556175	1,29946	0,0239143
<i>NME1</i>	-0,806319	0,287971	-1,09429	0,0311507
<i>NUMA1</i>	0,766295	-0,273677	1,03997	0,0420529
<i>OXSR1</i>	0,752316	-0,268684	1,021	0,0464814
<i>P4HB</i>	-1,07831	0,385111	-1,46342	0,00201449
<i>PCNA</i>	-0,793933	0,283548	-1,07748	0,0342571
<i>PDCD4</i>	0,774409	-0,276575	1,05098	0,0396355
<i>PDCD6IP</i>	1,00388	-0,358528	1,36241	0,00494793
<i>PDIA6</i>	-0,748445	0,267302	-1,01575	0,0477684
<i>PDS5A</i>	0,849372	-0,570099	1,41947	0,0103952
<i>PGM2</i>	0,866732	-0,309547	1,17628	0,0190064
<i>PIIB</i>	-0,780892	0,27889	-1,05978	0,0377823
<i>PPP1R7</i>	0,788144	-0,41314	1,20128	0,0247622
<i>PPP2CA</i>	0,862558	-0,308056	1,17061	0,0197002
<i>PPP2R1A</i>	0,751859	-0,268521	1,02038	0,046632
<i>PPP2R2A</i>	0,649784	-0,881643	1,53143	0,00976709
<i>PPT1</i>	-1,07573	0,164416	-1,24015	0,0199278
<i>PRDX4</i>	-0,753265	0,269023	-1,02229	0,0461698
<i>PRKAR2A</i>	0,960575	-0,608872	1,56945	0,00320763
<i>PSMD1</i>	0,934354	-0,333698	1,26805	0,0102225

<i>PYCR1</i>	-1,3424	0,210371	-1,55278	0,0039609
<i>RAB5B</i>	0,475614	-0,718886	1,1945	0,0361338
<i>RAB6A</i>	-0,92182	0,329222	-1,25104	0,0115357
<i>RAD50</i>	0,795009	-0,401184	1,19619	0,0227118
<i>RBM39</i>	0,875079	-0,312528	1,18761	0,0176773
<i>RCN1</i>	-0,876888	0,233916	-1,1108	0,0291698
<i>RNH1</i>	0,806049	-0,287875	1,09392	0,0312161
<i>RNPEP</i>	1,01268	-0,36167	1,37435	0,004481
<i>RPA1</i>	0,898109	-0,479027	1,37714	0,0111672
<i>RPL11</i>	-0,775638	0,277014	-1,05265	0,0392788
<i>RPL35</i>	-0,859497	0,306963	-1,16646	0,0202215
<i>RPS11</i>	-0,761301	0,271893	-1,03319	0,043596
<i>RPS14</i>	-0,749405	0,267645	-1,01705	0,0474468
<i>RPS20</i>	-0,858538	0,306621	-1,16516	0,020387
<i>RTCB</i>	1,17414	-0,419336	1,59348	0,000496593
<i>S100A9</i>	-1,17275	0,317256	-1,49	0,00221503
<i>SAFB</i>	0,678235	-0,626514	1,30475	0,0239958
<i>SCFD1</i>	0,741052	-0,421307	1,16236	0,0372931
<i>SCP2</i>	-0,75352	0,269114	-1,02263	0,0460865
<i>SEC61A1</i>	-0,928237	0,331513	-1,25975	0,0108475
<i>SELENBP1</i>	0,811295	-0,289748	1,10104	0,0299665
<i>SHMT2</i>	-1,06437	0,380131	-1,4445	0,00241013
<i>SLC25A12</i>	0,935066	-0,940945	1,87601	0,00148432
<i>SLC25A24</i>	0,764684	-0,273101	1,03779	0,0425461
<i>SLC25A6</i>	-0,892531	0,318761	-1,21129	0,0151372
<i>SLC3A2</i>	-1,1208	0,400287	-1,52109	0,00112551
<i>SNX1</i>	0,805132	-0,287547	1,09268	0,0314385
<i>SOD2</i>	-0,747761	0,267058	-1,01482	0,0479986
<i>SRPRB</i>	-0,808556	0,28877	-1,09733	0,0306138
<i>STAU1</i>	-1,00484	0,110564	-1,11541	0,0463406
<i>STT3A</i>	-1,07995	0,219466	-1,29942	0,0195429
<i>TKT</i>	0,931922	-0,332829	1,26475	0,0104675
<i>TMOD3</i>	0,823558	-0,452209	1,27577	0,020596
<i>TNPO1</i>	0,784145	-0,280052	1,0642	0,0368778
<i>TPP1</i>	1,35241	-0,483003	1,83541	1,08E-05
<i>TROVE2</i>	1,09767	-0,392026	1,4897	0,00155588
<i>UBE2M</i>	-1,38516	-0,129771	-1,25539	0,039243
<i>UFL1</i>	0,780895	-0,436028	1,21692	0,028238
<i>UGDH</i>	0,862761	-0,308129	1,17089	0,0196659
<i>VCAN</i>	1,02754	-0,496246	1,52379	0,00246928
<i>YBX1</i>	-1,00177	0,357774	-1,35954	0,00506563

TABELA SUPLEMENTAR 3B

Símbolo dos genes	Média de expressão nos tumores LB (<i>Log-Z-score LQF</i>)	Média de expressão nos tumores não – LB (<i>Log-Z-score LQF</i>)	Log-Fold Change	Valor de P
<i>ABCD3</i>	0,947197	-0,362098	1,3093	0,0206722
<i>ABCF1</i>	-1,0931	0,0890402	-1,18214	0,0438048
<i>ACTR2</i>	-0,976806	0,260482	-1,23729	0,0230201
<i>ACTR3</i>	-0,969877	0,258634	-1,22851	0,0241839
<i>AK3</i>	-1,51569	0,135862	-1,65155	0,00587816
<i>AP1M1</i>	-1,47578	-0,0618622	-1,41392	0,0244205
<i>ARF4</i>	-1,06499	0,283998	-1,34899	0,0117085
<i>ARL6IP5</i>	-0,946317	0,252351	-1,19867	0,0284914
<i>ARPC1B</i>	-1,23485	0,204364	-1,43921	0,0113084
<i>ARPC2</i>	-0,97049	0,258797	-1,22929	0,0240792
<i>ARPC3</i>	-0,915316	0,244084	-1,1594	0,0350471
<i>ARPC4</i>	-1,06739	0,284638	-1,35203	0,0114796
<i>ARPC5</i>	-1,47208	0,0553113	-1,52739	0,0101038
<i>BANF1</i>	-1,13936	0,107801	-1,24716	0,0308808
<i>CALM2</i>	-1,04073	0,277528	-1,31826	0,0142351
<i>CALU</i>	-1,16557	0,0694347	-1,23501	0,0439725
<i>CAMK2D</i>	-1,46453	0,0375735	-1,5021	0,0132948
<i>CAST</i>	-1,23854	0,0687013	-1,30725	0,0360292
<i>CBR3</i>	-1,42309	-0,07775	-1,34534	0,0317097
<i>CD59</i>	-1,37577	0,0890264	-1,4648	0,0185998
<i>CHMP4B</i>	-1,18022	0,182657	-1,36288	0,0190275
<i>CMPK1</i>	-0,88884	0,237024	-1,12586	0,0415221
<i>CNN3</i>	-1,02695	0,273853	-1,3008	0,0158535
<i>CPNE3</i>	0,921285	-0,245676	1,16696	0,0337019
<i>DAD1</i>	-1,44245	-0,0967317	-1,34572	0,0393791
<i>DBI</i>	-1,3321	0,247452	-1,57955	0,00341912
<i>DCPS</i>	-1,65493	-0,0551954	-1,59973	0,00578152
<i>DHRS4</i>	0,836443	-0,469446	1,30589	0,0320127
<i>EEF1B2</i>	-0,954815	0,254617	-1,20943	0,0268738
<i>EHD4</i>	0,897812	-0,239417	1,13723	0,0392326
<i>EIF4H</i>	-1,55102	0,0358492	-1,58687	0,009952
<i>EIF5A</i>	-1,03019	0,274718	-1,30491	0,0154601
<i>EIF5B</i>	-1,24157	0,209064	-1,45063	0,0101818
<i>EPB41L2</i>	-1,43505	0,382679	-1,81773	0,000152718
<i>ERP44</i>	-0,94331	0,251549	-1,19486	0,0290818
<i>ESD</i>	-1,01038	0,269435	-1,27982	0,0179901
<i>ESYT1</i>	0,976747	-0,260466	1,23721	0,0230297

<i>FABP5</i>	-1,20244	0,065996	-1,26843	0,0411683
<i>FBLN2</i>	-1,18042	0,161045	-1,34147	0,0264461
<i>FKBP4</i>	1,11009	-0,296023	1,40611	0,00797377
<i>FKBP9</i>	-1,70146	-0,162955	-1,53851	0,0196963
<i>GANAB</i>	-0,866676	0,231114	-1,09779	0,0476179
<i>GNS</i>	0,859819	-0,845496	1,70531	0,00699305
<i>GSN</i>	-1,24833	0,0455447	-1,29388	0,044163
<i>GSTK1</i>	-0,8709	0,23224	-1,10314	0,0464067
<i>GSTP1</i>	-1,24071	0,330857	-1,57157	0,00218349
<i>HDGF</i>	-0,955236	0,25473	-1,20997	0,0267954
<i>HSPA9</i>	1,05821	-0,282189	1,3404	0,0123752
<i>HYOU1</i>	-1,03982	0,277285	-1,3171	0,0143378
<i>IFI16</i>	-1,27183	0,109151	-1,38098	0,0203483
<i>IVD</i>	1,04731	-0,367508	1,41482	0,00893614
<i>KRT18</i>	1,00663	-0,268435	1,27506	0,0185046
<i>LAMB1</i>	-1,01166	0,153354	-1,16501	0,0452597
<i>LGALS3</i>	-1,46949	0,391864	-1,86135	8,30E-05
<i>LGMN</i>	-1,67959	-0,0444478	-1,63514	0,00980327
<i>LUC7L2</i>	-1,39186	0,141646	-1,53351	0,0088581
<i>MDH1</i>	-0,877251	0,233934	-1,11118	0,0446296
<i>MOGS</i>	-0,981889	0,261837	-1,24373	0,0221947
<i>MYH2</i>	0,64094	-0,705087	1,34603	0,0279284
<i>MYO1F</i>	-1,62061	-0,20824	-1,41237	0,0403413
<i>NAMPT</i>	-1,22327	0,326206	-1,54948	0,00264293
<i>NMT1</i>	-1,3148	-0,0129641	-1,30184	0,0386098
<i>NUCB1</i>	-1,48815	0,16404	-1,65219	0,003572
<i>OAT</i>	-1,08527	0,177552	-1,26282	0,0270669
<i>OXCT1</i>	-1,47582	-0,106103	-1,36971	0,0258247
<i>PAFAH1B1</i>	-1,66364	-0,197418	-1,46622	0,0289789
<i>PDIA4</i>	-0,938115	0,250164	-1,18828	0,0301241
<i>PDIA5</i>	-1,97981	0,143608	-2,12342	0,000160516
<i>PDXK</i>	-0,925131	0,246702	-1,17183	0,0328565
<i>PFKP</i>	-1,3359	0,032329	-1,36822	0,0224448
<i>PFN1</i>	-0,91057	0,242819	-1,15339	0,0361462
<i>PGRMC1</i>	-1,52811	0,172177	-1,70029	0,00279369
<i>PHGDH</i>	-0,97002	0,258672	-1,22869	0,0241595
<i>PLEC</i>	0,908716	-0,242324	1,15104	0,0365826
<i>PRKAR1A</i>	1,01779	-0,271411	1,2892	0,0170083
<i>PRPS2</i>	-1,24905	0,19627	-1,44533	0,0125315
<i>PSMA5</i>	-0,861303	0,229681	-1,09098	0,0491936
<i>PSMA7</i>	-0,872252	0,2326	-1,10485	0,046024
<i>PSMB2</i>	-1,19678	0,0700866	-1,26686	0,0389778
<i>PSMD4</i>	-1,15443	0,307849	-1,46228	0,00530816

<i>PSMD7</i>	-0,862158	0,229909	-1,09207	0,0489402
<i>PTPN1</i>	1,18131	-0,315017	1,49633	0,0040839
<i>PYGB</i>	-0,920788	0,245543	-1,16633	0,0338125
<i>RAB7A</i>	-0,962758	0,256735	-1,21949	0,0254276
<i>RHOA</i>	-0,88176	0,235136	-1,1169	0,0434001
<i>RIC8A</i>	-1,4122	0,126943	-1,53915	0,00932621
<i>RNH1</i>	-0,936191	0,249651	-1,18584	0,0305173
<i>RPL13A</i>	-0,943367	0,251564	-1,19493	0,0290706
<i>RPL21</i>	-0,973965	0,259724	-1,23369	0,0234918
<i>RPL22</i>	-1,07105	0,285613	-1,35666	0,0111378
<i>RPL23A</i>	-0,908264	0,242204	-1,15047	0,0366897
<i>RPL3</i>	-0,993016	0,264804	-1,25782	0,0204693
<i>RPL31</i>	-0,89835	0,23956	-1,13791	0,0390984
<i>RPL35A</i>	-1,35412	-0,0284248	-1,3257	0,0432672
<i>S100A10</i>	-1,64021	0,29687	-1,93708	0,000234648
<i>S100A11</i>	-1,07608	0,286954	-1,36303	0,0106811
<i>S100A13</i>	-1,57465	-0,0415852	-1,53306	0,0130872
<i>S100A4</i>	-1,19142	0,157495	-1,34891	0,0271933
<i>SAR1A</i>	-0,966353	0,257694	-1,22405	0,0247934
<i>SEC11A</i>	-1,72574	-0,163994	-1,56174	0,0179297
<i>SNRNP200</i>	-0,970998	0,258933	-1,22993	0,0239927
<i>SNRPD1</i>	-0,882716	0,235391	-1,11811	0,0431429
<i>SNX3</i>	-1,34375	0,210797	-1,55454	0,00734504
<i>SEM</i>	-1,29569	-0,0517541	-1,24393	0,0384902
<i>SRPR</i>	-1,53467	-0,0270255	-1,50764	0,0116963
<i>STAT3</i>	-1,12064	0,298838	-1,41948	0,00725767
<i>SUGT1</i>	-1,02867	0,274311	-1,30298	0,0156441
<i>TAGLN2</i>	-0,899521	0,239872	-1,13939	0,0388078
<i>TMED10</i>	-1,03184	0,275157	-1,30699	0,0152636
<i>TMEM43</i>	-1,00047	0,266792	-1,26726	0,0193741
<i>TPT1</i>	-1,07935	0,182426	-1,26177	0,0258661
<i>TSN</i>	-1,11754	0,298012	-1,41556	0,00746222
<i>TUBA4A</i>	-1,30365	0,073833	-1,37748	0,0276647
<i>TXNDC17</i>	-1,46521	-0,0440174	-1,4212	0,0348728
<i>TXNDC5</i>	-0,901856	0,240495	-1,14235	0,038233
<i>UBA3</i>	-1,29575	0,0716997	-1,36745	0,0291249
<i>UBE2I</i>	-1,02935	0,164261	-1,19361	0,0378877
<i>UCHL3</i>	-1,31816	0,0831526	-1,40132	0,0236205
<i>UFC1</i>	-1,8696	-0,0977233	-1,77188	0,0044118
<i>UGP2</i>	-1,03784	0,276758	-1,3146	0,0145627

TABELA SUPLEMENTAR 3C

Símbolo dos genes	Média de expressão nos tumores TN (Log-Z-score LQF)	Média de expressão nos tumores não – TN (Log-Z-score LQF)	Log-Fold Change	Valor de P
<i>2-Sep</i>	-0,654991	0,302303	-0,957294	0,0490766
<i>ABCD3</i>	-1,08455	0,374202	-1,45876	0,00175978
<i>ABCF1</i>	0,986034	-0,688693	1,67473	0,000190187
<i>ACAA2</i>	0,662191	-0,305627	0,967818	0,0463299
<i>ACO1</i>	-0,662949	0,305976	-0,968925	0,046048
<i>ACP1</i>	0,488146	-0,772307	1,26045	0,0235314
<i>ACTR1A</i>	-0,998497	0,460845	-1,45934	0,000913345
<i>ADH1B</i>	-1,80579	0,292506	-2,0983	2,33E-06
<i>ADPRHL2</i>	0,569308	-0,928449	1,49776	0,0120096
<i>ADSL</i>	0,715922	-0,330426	1,04635	0,0293702
<i>AGR2</i>	-1,36617	-0,140347	-1,22582	0,0440228
<i>AHNAK</i>	-0,954288	0,44044	-1,39473	0,00184892
<i>AHSA1</i>	0,875283	-0,662139	1,53742	0,00142703
<i>AIFM1</i>	-0,794396	0,366644	-1,16104	0,0137507
<i>AK2</i>	0,660634	-0,304908	0,965541	0,0469139
<i>AKT1</i>	-0,979091	0,144022	-1,12311	0,0256099
<i>ALAD</i>	-1,02025	0,0709047	-1,09116	0,0456087
<i>ALDH1A1</i>	-0,892758	0,412042	-1,3048	0,00436969
<i>ALDH3A2</i>	-1,08336	0,115857	-1,19922	0,0241727
<i>ALDH4A1</i>	-1,18814	-0,0550102	-1,13313	0,0376025
<i>ALYREF</i>	0,962043	-0,558356	1,5204	0,000748861
<i>ANP32B</i>	0,682582	-0,315038	0,997621	0,0391866
<i>ANP32E</i>	1,03111	-0,885911	1,91702	3,79E-05
<i>ANXA1</i>	0,7105	-0,327923	1,03842	0,03082
<i>ANXA11</i>	-0,870071	0,401571	-1,27164	0,00582196
<i>ANXA3</i>	0,611908	-0,668876	1,28078	0,0155408
<i>ANXA6</i>	-0,764847	0,353006	-1,11785	0,018561
<i>ANXA7</i>	-0,674651	0,311377	-0,986028	0,0418564
<i>AOC3</i>	-1,83803	0,3279	-2,16593	5,48E-08
<i>AP2B1</i>	-0,884946	0,408437	-1,29338	0,00483159
<i>AP3B1</i>	-0,911052	0,420485	-1,33154	0,00342842
<i>AP3D1</i>	-0,893429	0,412352	-1,30578	0,00433174
<i>APCS</i>	-0,65495	0,302285	-0,957235	0,0490923
<i>APEH</i>	-1,00229	0,462595	-1,46488	0,000856354
<i>API5</i>	0,883241	-0,40765	1,29089	0,00493758
<i>APMAP</i>	-0,723108	0,333742	-1,05685	0,0275318
<i>APOA4</i>	-0,66927	0,308894	-0,978165	0,0437455
<i>ARF4</i>	0,655097	-0,302353	0,95745	0,049035
<i>ARFIP1</i>	-1,0455	0,482539	-1,52804	0,00039067
<i>ARHGDI1A</i>	0,738634	-0,340908	1,07954	0,0238671
<i>ARMT1</i>	-1,17209	0,271909	-1,444	0,00375515
<i>ARPC3</i>	0,788653	-0,363994	1,15265	0,0145972
<i>ARPC4</i>	0,810023	-0,373857	1,18388	0,0116448

<i>ARPC5</i>	0,792368	-0,754834	1,5472	0,00195681
<i>ASNA1</i>	-0,672185	0,310239	-0,982424	0,0427143
<i>ASPN</i>	-1,00411	0,463434	-1,46754	0,000830101
<i>ATP1A1</i>	0,782645	-0,361221	1,14387	0,0155268
<i>ATP1B1</i>	0,496658	-0,625178	1,12184	0,0387165
<i>ATP5A1</i>	-0,666899	0,307799	-0,974698	0,0445987
<i>ATP5D</i>	0,614182	-0,581351	1,19553	0,0283988
<i>ATP6V1C1</i>	-1,22325	0,156108	-1,37936	0,00881266
<i>B2M</i>	0,959055	-0,442641	1,4017	0,00172009
<i>BANF1</i>	0,694319	-0,546642	1,24096	0,0121846
<i>BGN</i>	-1,04637	0,482939	-1,52931	0,000384178
<i>BLVRA</i>	-0,729132	0,336522	-1,06565	0,026061
<i>BLVRB</i>	-0,880188	0,406241	-1,28643	0,00513203
<i>BOLA2</i>	0,399674	-0,919617	1,31929	0,0246725
<i>C1QBP</i>	0,847454	-0,391133	1,23859	0,00764051
<i>CACYBP</i>	0,904738	-0,417571	1,32231	0,00373221
<i>CALM2</i>	0,768917	-0,354885	1,1238	0,0178289
<i>CALR</i>	1,02229	-0,471825	1,49411	0,000602708
<i>CALU</i>	0,542749	-0,529021	1,07177	0,0465523
<i>CAPN1</i>	-0,739127	0,341136	-1,08026	0,0237574
<i>CARS</i>	0,663597	-0,566328	1,22992	0,0164668
<i>CAV1</i>	-1,32182	0,444323	-1,76615	0,000102226
<i>CBR1</i>	0,764791	-0,35298	1,11777	0,0185713
<i>CBX3</i>	0,688339	-0,317695	1,00603	0,0373326
<i>CD59</i>	1,06049	-0,810049	1,87054	9,07E-05
<i>CDC42</i>	0,703037	-0,324479	1,02752	0,0329056
<i>CFL1</i>	0,823925	-0,380273	1,2042	0,00999711
<i>CISD2</i>	0,467774	-0,626668	1,09444	0,0469968
<i>CLIC1</i>	0,856792	-0,395442	1,25223	0,00684064
<i>CLIC4</i>	0,690124	-0,427321	1,11744	0,0222192
<i>CMA1</i>	-1,10537	0,113767	-1,21914	0,0123659
<i>CMPK1</i>	0,868278	-0,400744	1,26902	0,00595174
<i>CNN3</i>	0,812194	-0,374859	1,18705	0,011374
<i>COL14A1</i>	-0,774699	0,357553	-1,13225	0,0168283
<i>COL1A2</i>	-0,818584	0,377808	-1,19639	0,0106062
<i>COL6A1</i>	-0,87373	0,40326	-1,27699	0,00556415
<i>COL6A2</i>	-0,920552	0,42487	-1,34542	0,00300995
<i>COL6A3</i>	-0,896333	0,413692	-1,31003	0,00417073
<i>COLGALT1</i>	0,503716	-0,946181	1,4499	0,00958616
<i>CORO1C</i>	0,934493	-0,431304	1,3658	0,002473
<i>COTL1</i>	0,76188	-0,351637	1,11352	0,0191097
<i>COX5A</i>	0,411648	-0,870691	1,28234	0,0220313
<i>CPA3</i>	-1,16142	0,034882	-1,1963	0,0254822
<i>CRYAB</i>	0,724929	-0,767108	1,49204	0,00501347
<i>CSTB</i>	0,998434	-0,460816	1,45925	0,00091431
<i>CTSB</i>	0,907351	-0,418777	1,32613	0,00360389
<i>CYCS</i>	0,969031	-0,447245	1,41628	0,00147459
<i>DAD1</i>	0,727593	-0,891256	1,61885	0,0024403
<i>DAK</i>	-1,09189	0,37765	-1,46954	0,00157742

<i>DCN</i>	-0,922837	0,425925	-1,34876	0,00291587
<i>DCPS</i>	0,629149	-0,863272	1,49242	0,00278623
<i>DCTN1</i>	-0,80907	0,373417	-1,18249	0,0117652
<i>DDAH2</i>	0,970772	-0,448048	1,41882	0,00143491
<i>DDOST</i>	0,752444	-0,347282	1,09973	0,0209405
<i>DDX21</i>	0,632859	-0,443796	1,07665	0,0376973
<i>DEK</i>	0,754001	-0,503296	1,2573	0,0130888
<i>DNM1L</i>	0,681336	-0,314463	0,995798	0,0395974
<i>DNPEP</i>	-0,743774	0,343281	-1,08705	0,0227424
<i>DPT</i>	-1,13418	0,299854	-1,43403	0,00331777
<i>DSG2</i>	0,839394	-0,387413	1,22681	0,00839049
<i>DSP</i>	0,696074	-0,321265	1,01734	0,034949
<i>DSTN</i>	0,668493	-0,308535	0,977029	0,0440237
<i>DYNC1H1</i>	-0,701136	0,323601	-1,02474	0,033454
<i>DYNLL2</i>	0,657212	-0,73298	1,39019	0,010038
<i>EEA1</i>	-0,908136	0,109667	-1,0178	0,0487218
<i>EEF1A1</i>	1,1405	-0,526386	1,66689	4,57E-05
<i>EEF1B2</i>	0,822748	-0,37973	1,20248	0,0101289
<i>EEF1D</i>	0,862748	-0,398191	1,26094	0,00636716
<i>EEF1G</i>	1,01473	-0,468335	1,48306	0,000689917
<i>EHD2</i>	-1,17696	0,54321	-1,72017	1,61E-05
<i>EIF2AK2</i>	0,691457	-0,319134	1,01059	0,0363571
<i>EIF2S1</i>	0,944948	-0,43613	1,38108	0,00212464
<i>EIF2S3</i>	0,875967	-0,404293	1,28026	0,00541122
<i>EIF4A1</i>	1,06495	-0,491515	1,55646	0,000265304
<i>EIF4E</i>	0,750052	-0,911943	1,662	0,00153234
<i>EIF4H</i>	0,8047	-0,807273	1,61197	0,00183613
<i>EIF5A</i>	1,02409	-0,472656	1,49675	0,000583379
<i>EIF5B</i>	0,818346	-0,51849	1,33684	0,00627142
<i>EIF6</i>	0,710875	-0,328096	1,03897	0,0307177
<i>ELAVL1</i>	0,684653	-0,315994	1,00065	0,0385116
<i>EMILIN1</i>	-1,15055	0,320023	-1,47057	0,00589464
<i>ENO1</i>	0,815149	-0,376222	1,19137	0,0110136
<i>EPHX1</i>	-0,978934	0,451816	-1,43075	0,00126053
<i>ERGIC1</i>	-1,42123	0,243589	-1,66482	0,000831848
<i>ERH</i>	0,686969	-0,734142	1,42111	0,00704803
<i>ERO1L</i>	0,888222	-0,409949	1,29817	0,00463324
<i>ERP29</i>	0,903168	-0,416847	1,32001	0,00381111
<i>ERP44</i>	0,758248	-0,349961	1,10821	0,0197987
<i>ESYT1</i>	-1,01805	0,469868	-1,48792	0,000650385
<i>F13A1</i>	-1,22407	0,285857	-1,50992	0,00242323
<i>FABP4</i>	-1,56084	0,309497	-1,87034	7,38E-05
<i>FABP5</i>	1,15011	-0,824653	1,97477	7,38E-06
<i>FAH</i>	-0,873231	0,269286	-1,14252	0,0226634
<i>FASN</i>	-0,858285	0,396131	-1,25442	0,00671937
<i>FBN1</i>	-1,28259	0,113152	-1,39574	0,00728131
<i>FBP1</i>	-1,07126	0,494429	-1,56569	0,000232795
<i>FEN1</i>	0,565685	-0,781931	1,34762	0,0116096
<i>FKBP1A</i>	0,803262	-1,08583	1,88909	0,000212754

<i>FKBP9</i>	0,490272	-0,93783	1,4281	0,0125483
<i>FSCN1</i>	1,19965	-0,553683	1,75333	7,76E-06
<i>FTH1</i>	0,723102	-0,333739	1,05684	0,0275333
<i>FTL</i>	0,670826	-0,309612	0,980439	0,0431926
<i>FXR1</i>	0,544871	-0,842247	1,38712	0,0063789
<i>GANAB</i>	0,779404	-0,359725	1,13913	0,0160476
<i>GAPDH</i>	0,734264	-0,338891	1,07316	0,0248572
<i>GARS</i>	0,756447	-0,349129	1,10558	0,0201476
<i>GART</i>	0,992189	-0,457933	1,45012	0,0010152
<i>GBAS</i>	0,449306	-0,818033	1,26734	0,0306827
<i>GBE1</i>	-0,780749	0,360346	-1,1411	0,0158298
<i>GDI1</i>	-0,698504	0,322387	-1,02089	0,0342249
<i>GGT5</i>	-0,960066	0,179486	-1,13955	0,0312449
<i>GLIPR2</i>	0,376582	-0,770987	1,14757	0,0308751
<i>GLRX3</i>	0,779432	-0,359738	1,13917	0,0160431
<i>GLUD1</i>	-0,823683	0,380161	-1,20384	0,010024
<i>GMDS</i>	0,834133	-0,726815	1,56095	0,00353733
<i>GMPS</i>	0,749074	-0,345727	1,0948	0,0216269
<i>GNB2L1</i>	0,787672	-0,363541	1,15121	0,014746
<i>GOLT1B</i>	0,413492	-0,717261	1,13075	0,0422904
<i>GSS</i>	-1,14655	0,423982	-1,57053	0,000342426
<i>GSTK1</i>	0,687053	-0,317101	1,00415	0,0377407
<i>GSTM3</i>	-0,823804	0,264236	-1,08804	0,0278896
<i>GSTP1</i>	0,775592	-0,357965	1,13356	0,016678
<i>H2AFV</i>	0,696383	-0,321407	1,01779	0,0348562
<i>H2AFY</i>	-0,763883	0,352562	-1,11644	0,0187379
<i>HDGF</i>	0,946301	-0,436754	1,38306	0,00208271
<i>HEBP2</i>	0,852592	-0,660553	1,51314	0,00201275
<i>HIST1H2BB</i>	0,833805	-0,384833	1,21864	0,00894483
<i>HIST1H2BH</i>	0,780431	-0,360199	1,14063	0,0158811
<i>HIST1H4A</i>	0,655653	-0,302609	0,958262	0,048819
<i>HLA-A</i>	0,72258	-0,333498	1,05608	0,0276636
<i>HNRNPH3</i>	0,692375	-0,319558	1,01193	0,0360736
<i>HPRT1</i>	0,804514	-0,371314	1,17583	0,0123553
<i>HSD17B12</i>	0,703167	-0,324539	1,02771	0,0328682
<i>HSDL2</i>	-1,20121	-0,00960659	-1,1916	0,0352224
<i>HSP90AB2P</i>	0,831047	-0,38356	1,21461	0,00922904
<i>HSP90B1</i>	0,921742	-0,425419	1,34716	0,00296066
<i>HSPA5</i>	0,814037	-0,37571	1,18975	0,011148
<i>HSPA9</i>	-0,732809	0,338219	-1,07103	0,025194
<i>HSPB1</i>	-0,694026	0,32032	-1,01435	0,0355683
<i>HSPE1</i>	0,99448	-0,458991	1,45347	0,000977163
<i>HSPG2</i>	-0,658554	0,303948	-0,962502	0,0477023
<i>HYOU1</i>	1,11363	-0,513983	1,62761	9,00E-05
<i>IDH2</i>	0,893542	-0,412404	1,30595	0,0043254
<i>IFI16</i>	0,928366	-0,693864	1,62223	0,00072661
<i>IGLV3-9</i>	0,858962	-0,396444	1,25541	0,00666489
<i>ILK</i>	-0,683867	0,315631	-0,999498	0,0387669
<i>IPO5</i>	0,92232	-0,425686	1,34801	0,00293696

<i>ITIH1</i>	-0,687314	0,317222	-1,00454	0,0376577
<i>IVD</i>	-0,810541	0,272298	-1,08284	0,0264082
<i>KARS</i>	0,73156	-0,337643	1,0692	0,0254858
<i>KHDRBS1</i>	0,70202	-0,324009	1,02603	0,033198
<i>KRT18</i>	-0,775658	0,357996	-1,13365	0,0166668
<i>LAMB2</i>	-1,43474	0,0800079	-1,51474	0,0068752
<i>LDHA</i>	0,696436	-0,321432	1,01787	0,0348404
<i>LDHB</i>	1,10099	-0,50815	1,60914	0,00012125
<i>LGALS1</i>	0,83955	-0,387485	1,22703	0,00837544
<i>LGALS3</i>	0,736127	-0,339751	1,07588	0,0244311
<i>LGMN</i>	0,62035	-0,854398	1,47475	0,0074965
<i>LIMA1</i>	-0,999373	0,16188	-1,16125	0,0319894
<i>LMNB2</i>	-0,753491	0,347765	-1,10126	0,0207308
<i>LRP1</i>	-1,24361	0,401969	-1,64558	0,000578178
<i>LTA4H</i>	-0,815399	0,376338	-1,19174	0,0109835
<i>LUC7L2</i>	0,572919	-0,529251	1,10217	0,0387294
<i>LUM</i>	-0,893151	0,412224	-1,30537	0,00434742
<i>MAGOHB</i>	1,03981	-0,47991	1,51972	0,000435715
<i>MANF</i>	0,755756	-0,928047	1,6838	0,00138383
<i>MCCC2</i>	-1,37705	-0,104247	-1,2728	0,0308692
<i>MCM3</i>	0,909681	-0,842616	1,7523	0,000394955
<i>MCM7</i>	0,669438	-1,00352	1,67296	0,0015426
<i>MFAP2</i>	-1,29839	-0,0724636	-1,22593	0,0306406
<i>MIF</i>	0,676666	-0,312307	0,988973	0,0411654
<i>MOGS</i>	0,987433	-0,455738	1,44317	0,00109816
<i>MSN</i>	0,665929	-0,307352	0,973281	0,0449513
<i>MTA2</i>	0,513552	-0,575913	1,08946	0,0345752
<i>MTAP</i>	0,64029	-0,452791	1,09308	0,0358582
<i>MTPN</i>	0,569797	-0,802465	1,37226	0,0120004
<i>MVP</i>	-0,950339	0,438618	-1,38896	0,00196162
<i>MX1</i>	0,554863	-0,648647	1,20351	0,0239411
<i>MYBBP1A</i>	0,55166	-0,603773	1,15543	0,0260034
<i>MYDGF</i>	0,954716	-0,440638	1,39535	0,00183703
<i>MYH11</i>	-0,688729	0,317875	-1,0066	0,0372095
<i>MYL12B</i>	0,965512	-0,445621	1,41113	0,00155762
<i>MYL6</i>	0,87831	-0,405374	1,28368	0,00525477
<i>MYL9</i>	0,422953	-0,770263	1,19322	0,0371024
<i>MYO1C</i>	-0,730305	0,337064	-1,06737	0,0257819
<i>MYO1D</i>	-0,657409	0,30342	-0,960829	0,0481406
<i>MYOF</i>	-0,961419	0,443732	-1,40515	0,00165905
<i>NACA</i>	0,833316	-0,384607	1,21792	0,00899468
<i>NANS</i>	-0,876116	0,271153	-1,14727	0,0219182
<i>NASP</i>	0,583073	-0,693883	1,27696	0,0188379
<i>NCL</i>	0,837201	-0,386401	1,2236	0,00860455
<i>NDUFA5</i>	0,489089	-0,625059	1,11415	0,040687
<i>NDUFS1</i>	-0,881251	0,406731	-1,28798	0,00506362
<i>NHP2L1</i>	0,723292	-0,333827	1,05712	0,0274858
<i>NIPSNAP3A</i>	0,598309	-0,608852	1,20716	0,0315253
<i>NMT1</i>	0,512613	-0,656103	1,16872	0,0336415

<i>NUDT5</i>	0,768213	-0,35456	1,12277	0,0179538
<i>NUMA1</i>	-0,654166	0,301923	-0,956088	0,049399
<i>NUTF2</i>	0,880801	-0,694251	1,57505	0,00167503
<i>OGDH</i>	-0,866717	0,400023	-1,26674	0,00606672
<i>OGN</i>	-0,875236	0,403955	-1,27919	0,00546083
<i>PABPN1</i>	0,474495	-0,61018	1,08467	0,0449201
<i>PAFAH1B1</i>	0,333831	-0,893755	1,22759	0,0384466
<i>PAFAH1B2</i>	0,84523	-0,519518	1,36475	0,00435601
<i>PAPSS1</i>	0,990346	-0,457083	1,44743	0,0010467
<i>PCBP1</i>	0,729291	-0,336596	1,06589	0,026023
<i>PCMT1</i>	0,694425	-0,320504	1,01493	0,0354469
<i>PCYOX1</i>	-0,909383	0,419715	-1,3291	0,00350665
<i>PDIA3</i>	0,80307	-0,370648	1,17372	0,0125471
<i>PDIA4</i>	0,92328	-0,426129	1,34941	0,00289791
<i>PDIA5</i>	0,630045	-0,734262	1,36431	0,0133179
<i>PDIA6</i>	1,13347	-0,523139	1,65661	5,49E-05
<i>PDXDC1</i>	-1,04741	0,0900517	-1,13746	0,0346456
<i>PDXK</i>	0,94472	-0,436025	1,38075	0,00213177
<i>PEA15</i>	0,631387	-0,81579	1,44718	0,00579353
<i>PFKP</i>	0,696462	-0,695186	1,39165	0,00625833
<i>PFN1</i>	1,1251	-0,519277	1,64438	6,79E-05
<i>PGRMC1</i>	0,599892	-0,548395	1,14829	0,0301564
<i>PHGDH</i>	0,967333	-0,446461	1,41379	0,00151417
<i>PITPNB</i>	0,580135	-0,931347	1,51148	0,00473764
<i>PKP3</i>	0,607443	-0,523273	1,13072	0,0269178
<i>PLEC</i>	-1,08644	0,501433	-1,58787	0,000168188
<i>PLIN4</i>	-1,66898	0,172339	-1,84131	0,000374187
<i>PLOD1</i>	0,42867	-0,746531	1,1752	0,0359814
<i>POSTN</i>	-0,806649	0,3723	-1,17895	0,0120759
<i>PPA1</i>	0,733327	-0,338459	1,07179	0,0250737
<i>PPIA</i>	0,836891	-0,386257	1,22315	0,00863518
<i>PPIB</i>	1,13245	-0,522669	1,65512	5,64E-05
<i>PRDX1</i>	0,664385	-0,306639	0,971023	0,045517
<i>PRDX3</i>	-0,759488	0,350533	-1,11002	0,0195613
<i>PRDX4</i>	0,977882	-0,45133	1,42921	0,00128194
<i>PRDX5</i>	0,754961	-0,348444	1,1034	0,0204392
<i>PRDX6</i>	0,866965	-0,400138	1,2671	0,00604831
<i>PRELP</i>	-1,11643	0,374454	-1,49088	0,0016062
<i>PREP</i>	0,927211	-0,427943	1,35515	0,00274257
<i>PRKAR2A</i>	-1,00248	0,176428	-1,17891	0,0267517
<i>PRKCSH</i>	0,677321	-0,31261	0,989931	0,0409425
<i>PRPF19</i>	0,73758	-0,340421	1,078	0,0241031
<i>PSIP1</i>	0,986927	-0,636664	1,62359	0,000904578
<i>PSMA1</i>	0,723097	-0,333737	1,05683	0,0275345
<i>PSMA4</i>	0,8038	-0,370984	1,17478	0,0124499
<i>PSMA5</i>	0,955208	-0,440865	1,39607	0,00182345
<i>PSMA7</i>	0,782033	-0,360938	1,14297	0,0156242
<i>PSMB1</i>	0,816344	-0,376774	1,19312	0,0108704
<i>PSMB2</i>	0,63124	-0,578711	1,20995	0,0226947

<i>PSMB8</i>	0,554316	-0,495242	1,04956	0,0427731
<i>PSMB9</i>	0,614867	-0,392687	1,00755	0,0425039
<i>PSMD4</i>	0,803718	-0,370947	1,17466	0,0124608
<i>PSMD5</i>	-0,691915	0,319345	-1,01126	0,0362155
<i>PSMD9</i>	0,683965	-0,592814	1,27678	0,0138596
<i>PTRF</i>	-1,37354	0,491655	-1,86519	6,81E-06
<i>RAB1A</i>	0,863673	-0,878556	1,74223	0,001143
<i>RAB2A</i>	1,02722	-0,474103	1,50133	0,000550976
<i>RAN</i>	1,11206	-0,51326	1,62532	9,35E-05
<i>RANBP1</i>	1,01566	-0,810592	1,82626	0,000271838
<i>RBBP4</i>	0,887248	-0,409499	1,29675	0,0046915
<i>RBM8A</i>	0,915152	-1,02665	1,9418	0,000145695
<i>RBMX</i>	0,763363	-0,352321	1,11568	0,0188339
<i>RCC2</i>	0,697333	-0,321846	1,01918	0,0345724
<i>RCN1</i>	0,84146	-0,473721	1,31518	0,00420382
<i>RHOA</i>	0,751454	-0,346825	1,09828	0,0211404
<i>RNPEP</i>	-0,823658	0,38015	-1,20381	0,0100268
<i>RPL10</i>	0,739279	-0,341206	1,08048	0,0237236
<i>RPL10A</i>	0,859785	-0,396824	1,25661	0,00659926
<i>RPL11</i>	1,00467	-0,463694	1,46836	0,000822106
<i>RPL12</i>	0,991657	-0,457688	1,44934	0,00102422
<i>RPL13</i>	0,992041	-0,457865	1,44991	0,00101771
<i>RPL13A</i>	1,01085	-0,466545	1,47739	0,000738601
<i>RPL14</i>	0,970856	-0,448087	1,41894	0,00143301
<i>RPL15</i>	0,780546	-0,360252	1,1408	0,0158626
<i>RPL17</i>	1,00806	-0,465257	1,47331	0,0007754
<i>RPL18</i>	0,856864	-0,395476	1,25234	0,00683473
<i>RPL18A</i>	0,846148	-0,39053	1,23668	0,00775815
<i>RPL21</i>	0,841964	-0,388599	1,23056	0,00814514
<i>RPL22</i>	1,10876	-0,511738	1,6205	0,000101128
<i>RPL23</i>	1,10557	-0,510264	1,61583	0,000109019
<i>RPL23A</i>	0,95854	-0,442403	1,40094	0,00173365
<i>RPL24</i>	0,869673	-0,518959	1,38863	0,00304854
<i>RPL26</i>	0,87792	-0,537162	1,41508	0,00289192
<i>RPL27</i>	1,04873	-0,48403	1,53276	0,000366914
<i>RPL27A</i>	0,969033	-0,447246	1,41628	0,00147455
<i>RPL28</i>	0,814454	-0,480408	1,29486	0,00596544
<i>RPL3</i>	0,878413	-0,405421	1,28383	0,00524801
<i>RPL30</i>	1,1115	-0,513001	1,6245	9,48E-05
<i>RPL31</i>	0,930739	-0,429572	1,36031	0,00260909
<i>RPL32</i>	0,977843	-0,583916	1,56176	0,00064571
<i>RPL34</i>	0,537339	-0,937719	1,47506	0,00790648
<i>RPL35</i>	0,983167	-0,453769	1,43694	0,00117732
<i>RPL35A</i>	0,659552	-0,753859	1,41341	0,0110061
<i>RPL36</i>	0,879726	-0,573185	1,45291	0,00347056
<i>RPL38</i>	1,05607	-0,487418	1,54349	0,00031746
<i>RPL4</i>	0,720979	-0,332759	1,05374	0,0280668
<i>RPL5</i>	0,970337	-0,447848	1,41818	0,00144474
<i>RPL7</i>	0,802556	-0,370411	1,17297	0,012616

<i>RPL7A</i>	0,705509	-0,32562	1,03113	0,0322029
<i>RPL9</i>	0,880524	-0,406396	1,28692	0,00511036
<i>RPLP2</i>	0,812334	-0,374923	1,18726	0,0113567
<i>RPN1</i>	0,834347	-0,385083	1,21943	0,00888979
<i>RPN2</i>	0,664459	-0,306673	0,971132	0,0454898
<i>RPS10</i>	1,04528	-0,482437	1,52772	0,000392346
<i>RPS11</i>	0,852976	-0,393681	1,24666	0,00715886
<i>RPS12</i>	1,08331	-0,49999	1,5833	0,000180074
<i>RPS13</i>	1,10187	-0,508554	1,61042	0,000118825
<i>RPS14</i>	0,983436	-0,453894	1,43733	0,00117218
<i>RPS15</i>	1,00241	-0,64033	1,64274	0,000681473
<i>RPS15A</i>	1,05124	-0,485189	1,53643	0,000349311
<i>RPS16</i>	1,01413	-0,468062	1,4822	0,000697153
<i>RPS17</i>	0,844066	-0,389569	1,23364	0,00794878
<i>RPS18</i>	0,850503	-0,39254	1,24304	0,00737138
<i>RPS19</i>	1,12058	-0,517193	1,63778	7,60E-05
<i>RPS2</i>	0,753536	-0,347786	1,10132	0,0207219
<i>RPS20</i>	1,02678	-0,473897	1,50067	0,000555493
<i>RPS21</i>	0,848585	-0,934373	1,78296	0,000304403
<i>RPS23</i>	0,921116	-0,425131	1,34625	0,00298648
<i>RPS25</i>	0,952777	-0,439743	1,39252	0,00189139
<i>RPS26</i>	0,972291	-0,44875	1,42104	0,001401
<i>RPS27</i>	0,889271	-0,88513	1,7744	0,000694908
<i>RPS27A</i>	0,808022	-0,372933	1,18095	0,011899
<i>RPS28</i>	0,898403	-0,94621	1,84461	0,000117635
<i>RPS3</i>	0,822157	-0,379457	1,20161	0,0101955
<i>RPS3A</i>	0,86624	-0,399803	1,26604	0,00610214
<i>RPS4X</i>	0,729522	-0,336702	1,06622	0,0259679
<i>RPS5</i>	0,933181	-0,430699	1,36388	0,00251985
<i>RPS6</i>	0,836048	-0,385868	1,22192	0,00871888
<i>RPS7</i>	0,990914	-0,457345	1,44826	0,00103691
<i>RPS8</i>	0,896844	-0,413928	1,31077	0,0041429
<i>RPS9</i>	0,835628	-0,385675	1,2213	0,0087608
<i>RPSA</i>	0,748103	-0,345278	1,09338	0,0218281
<i>RSL1D1</i>	0,806903	-0,372417	1,17932	0,012043
<i>RSU1</i>	0,656434	-0,302969	0,959403	0,0485163
<i>RTN4</i>	-0,768612	0,354744	-1,12336	0,017883
<i>S100A11</i>	1,00608	-0,464346	1,47043	0,000802352
<i>S100A4</i>	0,703821	-0,509705	1,21353	0,0228835
<i>S100A6</i>	0,899786	-0,584511	1,4843	0,00273495
<i>S100A8</i>	0,747771	-0,784628	1,5324	0,00393045
<i>S100A9</i>	0,820967	-0,488304	1,30927	0,00555544
<i>SARNP</i>	0,759213	-0,489286	1,2485	0,0118223
<i>SCARB2</i>	-0,737437	0,340355	-1,07779	0,0241352
<i>SCP2</i>	0,675337	-0,311694	0,987031	0,04162
<i>SEC11A</i>	0,595233	-0,994942	1,59018	0,00441272
<i>SEC23B</i>	-1,24261	-0,0289802	-1,21363	0,0382178
<i>SEC24C</i>	-0,815098	0,376199	-1,1913	0,0110197
<i>SEC61A1</i>	0,84343	-0,389275	1,2327	0,00800784

<i>SEC63</i>	0,597614	-0,997222	1,59484	0,0037538
<i>SELENBP1</i>	-0,677168	0,312539	-0,989708	0,0409943
<i>SEPHS1</i>	0,589742	-0,751067	1,34081	0,0191352
<i>SH3BGRL3</i>	0,575478	-0,680414	1,25589	0,0201429
<i>SHMT2</i>	0,695035	-0,320785	1,01582	0,0352621
<i>SKP1</i>	0,882178	-0,407159	1,28934	0,00500456
<i>SLC25A1</i>	-0,705782	0,325746	-1,03153	0,032126
<i>SLC9A3R1</i>	-0,719453	0,332055	-1,05151	0,0284553
<i>SNRPA1</i>	0,663153	-0,438761	1,10191	0,028697
<i>SNRPB</i>	0,931296	-0,429829	1,36113	0,00258849
<i>SNRPD1</i>	1,12325	-0,518425	1,64168	7,11E-05
<i>SNRPD2</i>	1,00281	-0,462838	1,46565	0,000848682
<i>SNRPD3</i>	1,02739	-0,474181	1,50157	0,000549259
<i>SNRPE</i>	0,598071	-0,933515	1,53159	0,00353337
<i>SNX1</i>	-0,877723	0,405103	-1,28283	0,00529361
<i>SNX3</i>	1,10815	-0,681685	1,78983	8,32E-05
<i>SNX9</i>	-1,15138	0,38905	-1,54043	0,000987764
<i>SOD2</i>	1,00147	-0,462215	1,46368	0,00086845
<i>SPR</i>	-0,944569	0,289559	-1,23413	0,0140155
<i>SEM</i>	0,553072	-0,713653	1,26673	0,0136632
<i>SRP14</i>	0,781119	-0,360516	1,14164	0,0157704
<i>SRPR</i>	0,626585	-0,792582	1,41917	0,00590424
<i>SRPRB</i>	0,979488	-0,452071	1,43156	0,00124936
<i>SRSF1</i>	0,771456	-0,356057	1,12751	0,0173838
<i>SRSF7</i>	0,711869	-0,328555	1,04042	0,0304487
<i>SSB</i>	0,676524	-0,312242	0,988766	0,0412136
<i>SSBP1</i>	1,01834	-0,679395	1,69774	0,000809505
<i>SSR1</i>	1,00174	-0,462341	1,46408	0,000864412
<i>SSR4</i>	1,10435	-0,5097	1,61405	0,000112171
<i>SSRP1</i>	0,675675	-0,31185	0,987525	0,041504
<i>STIP1</i>	0,662212	-0,305637	0,967849	0,046322
<i>STK24</i>	0,792412	-0,528435	1,32085	0,00913145
<i>STOML2</i>	0,773514	-0,357006	1,13052	0,0170297
<i>STT3A</i>	0,615729	-0,4632	1,07893	0,045012
<i>SUGT1</i>	0,816446	-0,376821	1,19327	0,0108583
<i>SUMO2</i>	0,864021	-0,584835	1,44886	0,00477719
<i>SUPT16H</i>	0,576492	-0,542794	1,11929	0,0349873
<i>SYNCRIP</i>	0,97713	-0,450983	1,42811	0,00129746
<i>TAGLN2</i>	1,10375	-0,509421	1,61317	0,000113763
<i>TALDO1</i>	0,719739	-0,332187	1,05193	0,0283821
<i>TCEB2</i>	0,686926	-0,643758	1,33068	0,0148057
<i>TF</i>	-0,654167	0,301923	-0,95609	0,0493984
<i>TPD52L2</i>	1,05167	-0,659976	1,71164	0,000280545
<i>TPI1</i>	0,676161	-0,312074	0,988236	0,0413375
<i>TPM2</i>	0,789706	-0,36448	1,15419	0,0144389
<i>TPM3</i>	0,835697	-0,385706	1,2214	0,00875393
<i>TPM3</i>	0,77636	-0,35832	1,13468	0,0165493
<i>TPM4</i>	0,664617	-0,306746	0,971363	0,0454316
<i>TPP1</i>	-0,67263	0,310445	-0,983075	0,0425583

<i>TPP2</i>	0,71384	-0,63818	1,35202	0,0111707
<i>TPSAB1</i>	-1,20513	0,23027	-1,4354	0,00776241
<i>TPT1</i>	0,729736	-0,458417	1,18815	0,0153586
<i>TRA2B</i>	0,734253	-0,338886	1,07314	0,0248599
<i>TSN</i>	1,06733	-0,492612	1,55994	0,000252635
<i>TXN</i>	0,818232	-0,377646	1,19588	0,0106473
<i>TXNDC17</i>	0,653494	-0,803237	1,45673	0,0111786
<i>TXNDC5</i>	0,757603	-0,349663	1,10727	0,0199231
<i>UBA2</i>	0,721271	-0,332894	1,05417	0,0279929
<i>UBE2I</i>	0,989453	-0,583864	1,57332	0,000505774
<i>UBE2L3</i>	0,893276	-0,412281	1,30556	0,00434037
<i>UBE2M</i>	0,57458	-0,937699	1,51228	0,00599012
<i>UCHL3</i>	0,603551	-0,588205	1,19176	0,029009
<i>UFC1</i>	0,696258	-1,00937	1,70563	0,00125858
<i>UGDH</i>	-0,89015	0,410838	-1,30099	0,00451966
<i>UGP2</i>	0,97129	-0,448288	1,41958	0,00142326
<i>UQCRC2</i>	-0,709715	0,327561	-1,03728	0,0310343
<i>USP39</i>	0,808804	-0,523937	1,33274	0,00726777
<i>VASP</i>	0,75908	-0,350345	1,10942	0,0196391
<i>VAT1</i>	-0,869968	0,401524	-1,27149	0,00582931
<i>VCAN</i>	-1,14097	0,387391	-1,52836	0,0010611
<i>VDAC3</i>	0,656118	-0,302824	0,958942	0,0486384
<i>VPS29</i>	0,651566	-0,843059	1,49462	0,00517464
<i>XPO5</i>	0,433907	-0,823325	1,25723	0,0195795
<i>YARS</i>	0,870229	-0,401644	1,27187	0,0058106
<i>YBX1</i>	0,950864	-0,43886	1,38972	0,0019463
<i>YWHAE</i>	0,790499	-0,364846	1,15534	0,0143207
<i>YWHAQ</i>	0,670997	-0,309691	0,980688	0,0431324
<i>YWHAZ</i>	0,741805	-0,342372	1,08418	0,0231681

TABELA SUPLEMENTAR 3D

Símbolo dos genes	Média de expressão nos tumores HER2 (Log-Z-score LQF)	Média de expressão nos tumores não – HER2 (Log-Z-score LQF)	Log-Fold Change	Valor de P
<i>ACIN1</i>	-0,937646	0,250039	-1,18769	0,0302196
<i>ANP32A</i>	-1,15955	0,309213	-1,46876	0,00505456
<i>ANP32B</i>	-1,05345	0,28092	-1,33437	0,0128611
<i>ANPEP</i>	-0,87951	0,234536	-1,11405	0,0440103
<i>APMAP</i>	1,26025	-0,336067	1,59632	0,00174957
<i>APOA2</i>	0,677127	-0,832762	1,50989	0,0245915
<i>APOD</i>	1,2318	-0,441614	1,67341	0,00171606
<i>ARF6</i>	-1,04826	0,279537	-1,3278	0,0134077
<i>BCAP31</i>	0,799399	-0,467652	1,26705	0,0406558
<i>CAPG</i>	-1,1251	0,300028	-1,42513	0,00697108
<i>CAPRIN1</i>	-0,97695	0,169936	-1,14689	0,0419605
<i>CBX3</i>	-0,871653	0,232441	-1,10409	0,0461931
<i>CCAR1</i>	-1,63568	-0,0795447	-1,55614	0,016422
<i>CCAR2</i>	-0,877422	0,233979	-1,1114	0,0445825
<i>COPG1</i>	0,939963	-0,250657	1,19062	0,0297499
<i>CPSF6</i>	-1,11961	0,0745888	-1,1942	0,0464362
<i>CSK</i>	-0,948615	0,252964	-1,20158	0,0280467
<i>CTNNB1</i>	-1,039	0,164283	-1,20329	0,0368028
<i>DBI</i>	0,91085	-0,350667	1,26152	0,026386
<i>DDX1</i>	-0,943442	0,251585	-1,19503	0,0290556
<i>DDX39B</i>	-0,914405	0,243841	-1,15825	0,0352562
<i>DDX42</i>	-1,1658	0,0924097	-1,25821	0,0353626
<i>DNAJA2</i>	0,825584	-0,473623	1,29921	0,0403068
<i>ECM29</i>	-1,27213	0,181672	-1,4538	0,0153215
<i>EFTUD2</i>	-0,943435	0,251583	-1,19502	0,0290571
<i>F13A1</i>	0,881838	-0,477041	1,35888	0,023845
<i>FAM98B</i>	-1,76319	0,0102294	-1,77342	0,00322402
<i>G6PD</i>	0,86018	-0,229381	1,08956	0,0495276
<i>GBE1</i>	1,06224	-0,283265	1,34551	0,0119749
<i>GPX1</i>	-0,898302	0,239547	-1,13785	0,0391105
<i>HK1</i>	-1,09126	0,291002	-1,38226	0,00939352
<i>HMGCL</i>	-1,59159	-0,13287	-1,45872	0,0194082
<i>HNRNPD</i>	-0,973158	0,259509	-1,23267	0,0236272
<i>HNRNPM</i>	-0,903028	0,240807	-1,14383	0,0379472
<i>HNRNPR</i>	-1,17016	0,312042	-1,4822	0,0045602
<i>HNRNPUL2</i>	-0,931584	0,248422	-1,18001	0,0314755
<i>HSD17B4</i>	-1,06519	0,284051	-1,34924	0,0116894
<i>HSPA1A</i>	-0,886744	0,236465	-1,12321	0,0420715
<i>IDH1</i>	1,14687	-0,305833	1,45271	0,00570196
<i>IQGAP1</i>	-0,986824	0,263153	-1,24998	0,0214158
<i>KPNB1</i>	-0,976221	0,260326	-1,23655	0,0231167
<i>LMNB1</i>	-0,907906	0,242108	-1,15001	0,0367745
<i>LRRC47</i>	-0,906374	0,2417	-1,14807	0,0371398

<i>MAT2B</i>	-1,10495	0,294653	-1,3996	0,00834261
<i>MYH2</i>	-1,50284	-0,133412	-1,36943	0,0249017
<i>NAMPT</i>	1,15812	-0,308831	1,46695	0,00512448
<i>NAPRT</i>	-1,45734	-0,0702682	-1,38708	0,0262692
<i>NDRG1</i>	0,687109	-0,772014	1,45912	0,0413138
<i>NIT2</i>	-1,27527	0,0687371	-1,34401	0,0317611
<i>PDS5A</i>	-1,29805	0,0971787	-1,39523	0,022147
<i>PKP3</i>	-1,08298	0,078268	-1,16125	0,0496338
<i>POR</i>	0,935373	-0,818152	1,75352	0,0032868
<i>PTBP1</i>	-1,11789	0,298104	-1,41599	0,0074391
<i>PTRH2</i>	0,666622	-0,693025	1,35965	0,0404271
<i>RAB1A</i>	-1,51239	-0,0126408	-1,49975	0,0227674
<i>RPS15A</i>	0,979053	-0,448151	1,4272	0,0101729
<i>RTCB</i>	-1,02393	0,273049	-1,29698	0,0162268
<i>RTRAF</i>	-1,02013	0,272034	-1,29216	0,0167075
<i>RUVBL1</i>	-1,02524	0,273397	-1,29864	0,0160642
<i>RUVBL2</i>	-1,06757	0,284686	-1,35226	0,0114624
<i>SF3A1</i>	-0,892113	0,237897	-1,13001	0,0406752
<i>SLC25A13</i>	0,880419	-0,234778	1,1152	0,043763
<i>SRP14</i>	-0,972456	0,259322	-1,23178	0,0237455
<i>SRRT</i>	-0,943989	0,25173	-1,19572	0,0289476
<i>SURF4</i>	0,927588	-0,613724	1,54131	0,0118108
<i>TRIM25</i>	-0,879757	0,234602	-1,11436	0,0439431
<i>TROVE2</i>	-0,956286	0,25501	-1,2113	0,0266012
<i>TSTA3</i>	0,877113	-0,233897	1,11101	0,0446678
<i>TXNL1</i>	-1,08352	0,288939	-1,37246	0,0100331
<i>USP7</i>	-1,0355	0,276133	-1,31163	0,0148333
<i>XPNPEP1</i>	-0,92302	0,246139	-1,16916	0,0333186
<i>XRCC6</i>	-0,886441	0,236384	-1,12283	0,0421513

TABELA SUPLEMENTAR 4 – ONCOGENES E GENES SUPRESSORES DE TUMOR (CANÔNICOS E CANDIDATOS) IDENTIFICADOS DENTRE AS DEPS DE CADA SUBTIPO. (A) ONCOGENES E GENES SUPRESSORES DE TUMOR IDENTIFICADOS DENTRE AS DEPS DA COMPARAÇÃO DE TUMORES LA X NÃO-LA. (B) ONCOGENES E GENES SUPRESSORES DE TUMOR IDENTIFICADOS DENTRE AS DEPS DA COMPARAÇÃO DE TUMORES LB X NÃO-LB. (C) ONCOGENES E GENES SUPRESSORES DE TUMOR IDENTIFICADOS DENTRE AS DEPS DA COMPARAÇÃO DE TUMORES TNBC X NÃO-TNBC. (D) ONCOGENES E GENES SUPRESSORES DE TUMOR IDENTIFICADOS DENTRE AS DEPS DA COMPARAÇÃO DE TUMORES HER2+ X NÃO-HER2+.

TABELA SUPLEMENTAR 4A

DEP (Símbolo do gene)	Classificação NCG	Média de expressão em amostras LA [log(Z-score)]	Média de expressão em amostras não-LA [log(Z-score)]
<i>FEN1</i>	Supressor de tumor	-1,27128	-0,0296154
<i>PPP2R1A</i>	Oncogene e relacionado com saúde	0,751859	-0,268521
<i>GNA11</i>	Oncogene	0,475982	-0,826468
<i>IDH2</i>	Oncogene	-0,91578	0,327064
<i>MAPK1</i>	Oncogene	0,799915	-0,285684
<i>CYFIP1</i>	Candidato oncogene	0,878259	-0,313664
<i>DPP7</i>	Candidato oncogene	0,693089	-0,436055
<i>EEF1D</i>	Candidato oncogene	-1,05788	0,377813
<i>GPS1</i>	Candidato oncogene	0,91211	-0,579577
<i>HLA-A</i>	Candidato oncogene	-0,764853	0,273162
<i>HSP90AB1</i>	Candidato oncogene	-0,824418	0,294435
<i>IQGAP1</i>	Candidato oncogene	0,822125	-0,293616
<i>KPNB1</i>	Candidato oncogene	0,79227	-0,282954
<i>LTA4H</i>	Candidato oncogene	0,831134	-0,296834
<i>NUMA1</i>	Candidato oncogene	0,766295	-0,273677
<i>RBM39</i>	Candidato oncogene	0,875079	-0,312528
<i>TROVE2</i>	Candidato oncogene	1,09767	-0,392026
<i>AHNAK</i>	Candidato supressor de tumor	0,984224	-0,351509
<i>CAST</i>	Candidato supressor de tumor	0,71249	-0,534722
<i>CCAR1</i>	Candidato supressor de tumor	0,87709	-0,865811
<i>CNDP2</i>	Candidato supressor de tumor	0,753839	-0,269228
<i>DCTN1</i>	Candidato supressor de tumor	0,94739	-0,338354
<i>EZR</i>	Candidato supressor de tumor	0,753824	-0,269223
<i>GNAI2</i>	Candidato supressor de tumor	1,00289	-0,358175
<i>HLA-DRB1</i>	Candidato supressor de tumor	0,840518	-0,527549

<i>HSPA4</i>	Candidato supressor de tumor	1,06574	-0,38062
<i>IQGAP2</i>	Candidato supressor de tumor	1,03387	-0,914067
<i>LRP1</i>	Candidato supressor de tumor	0,713362	-0,414493
<i>LRRC47</i>	Candidato supressor de tumor	0,915068	-0,32681
<i>PDS5A</i>	Candidato supressor de tumor	0,849372	-0,570099
<i>TNPO1</i>	Candidato supressor de tumor	0,784145	-0,280052
<i>VCAN</i>	Candidato supressor de tumor	1,02754	-0,496246

TABELA SUPLEMENTAR 4B

DEP (Símbolo do gene)	Classificação NCG	Média de expressão em amostras LB [log(Z-score)]	Média de expressão em amostras não-LB [log(Z-score)]
<i>FBLN2</i>	Supressor de tumor	-1,18042	0,161045
<i>RPL22</i>	Supressor de tumor	-1,07105	0,285613
<i>RHOA</i>	Oncogene e relacionado com saúde	-0,88176	0,235136
<i>STAT3</i>	Oncogene	-1,12064	0,298838
<i>FKBP9</i>	Oncogene	-1,70146	-0,162955
<i>PRKAR1A</i>	Função ambígua no câncer	1,01779	-0,271411
<i>ABCF1</i>	Candidato oncogene	-1,0931	0,0890402
<i>ARPC1B</i>	Candidato oncogene	-1,23485	0,204364
<i>DBI</i>	Candidato oncogene	-1,3321	0,247452
<i>LAMB1</i>	Candidato oncogene	-1,01166	0,153354
<i>LUC7L2</i>	Candidato oncogene	-1,39186	0,141646
<i>NAMPT</i>	Candidato oncogene	-1,22327	0,326206
<i>PDIA4</i>	Candidato oncogene	-0,938115	0,250164
<i>PFKP</i>	Candidato oncogene	-1,3359	0,032329
<i>PHGDH</i>	Candidato oncogene	-0,97002	0,258672
<i>PLEC</i>	Candidato oncogene	0,908716	-0,242324
<i>PTPN1</i>	Candidato oncogene	1,18131	-0,315017
<i>RPL13A</i>	Candidato oncogene	-0,943367	0,251564
<i>RPL31</i>	Candidato oncogene	-0,89835	0,23956
<i>CAST</i>	Candidato supressor de tumor	-1,23854	0,0687013
<i>EIF5A</i>	Candidato supressor de tumor	-1,03019	0,274718
<i>MYH2</i>	Candidato supressor de tumor	0,64094	-0,705087
<i>SNRNP200</i>	Candidato supressor de tumor	-0,970998	0,258933

TABELA SUPLEMENTAR 4C

DEP (Símbolo do gene)	Classificação NCG	Média de expressão em amostras TNBC [log(Z-score)]	Média de expressão em amostras não-TNBC [log(Z-score)]
<i>B2M</i>	Supressor de tumor	0,959055	-0,442641
<i>CARS</i>	Supressor de tumor	0,663597	-0,566328
<i>FEN1</i>	Supressor de tumor	0,565685	-0,781931
<i>RPL10</i>	Supressor de tumor	0,739279	-0,341206
<i>RPL22</i>	Supressor de tumor	1,10876	-0,511738
<i>RPL5</i>	Supressor de tumor	0,970337	-0,447848
<i>TPM3</i>	Supressor de tumor	0,77636	-0,35832
<i>YWHAE</i>	Supressor de tumor	0,790499	-0,364846
<i>RHOA</i>	Oncogene e relacionado com saúde	0,751454	-0,346825
<i>CALR</i>	Oncogene	1,02229	-0,471825
<i>DEK</i>	Oncogene	0,754001	-0,503296
<i>IDH2</i>	Oncogene	0,893542	-0,412404
<i>PSIP1</i>	Oncogene	0,986927	-0,636664
<i>FKBP9</i>	Oncogene	0,490272	-0,93783
<i>API5</i>	Função ambígua no câncer	0,883241	-0,40765
<i>ATP1A1</i>	Função ambígua no câncer	0,782645	-0,361221
<i>PCBP1</i>	Função ambígua no câncer	0,729291	-0,336596
<i>AKT1</i>	Candidato oncogene	-0,979091	0,144022
<i>ABCF1</i>	Candidato oncogene	0,986034	-0,688693
<i>ACO1</i>	Candidato oncogene	-0,662949	0,305976
<i>AHSA1</i>	Candidato oncogene	0,875283	-0,662139
<i>ALYREF</i>	Candidato oncogene	0,962043	-0,558356
<i>APMAP</i>	Candidato oncogene	-0,723108	0,333742
<i>COL1A2</i>	Candidato oncogene	-0,818584	0,377808
<i>EEF1A1</i>	Candidato oncogene	1,1405	-0,526386
<i>EEF1D</i>	Candidato oncogene	0,862748	-0,398191
<i>ELAVL1</i>	Candidato oncogene	0,684653	-0,315994
<i>FTH1</i>	Candidato oncogene	0,723102	-0,333739
<i>FXR1</i>	Candidato oncogene	0,544871	-0,842247
<i>GLIPR2</i>	Candidato oncogene	0,376582	-0,770987
<i>H2AFV</i>	Candidato oncogene	0,696383	-0,321407
<i>HIST1H2BB</i>	Candidato oncogene	0,833805	-0,384833
<i>HLA-A</i>	Candidato oncogene	0,72258	-0,333498
<i>HNRNPH3</i>	Candidato oncogene	0,692375	-0,319558
<i>IPO5</i>	Candidato oncogene	0,92232	-0,425686
<i>LGALS1</i>	Candidato oncogene	0,83955	-0,387485
<i>LTA4H</i>	Candidato oncogene	-0,815399	0,376338
<i>LUC7L2</i>	Candidato oncogene	0,572919	-0,529251
<i>MCM7</i>	Candidato oncogene	0,669438	-1,00352
<i>MSN</i>	Candidato oncogene	0,665929	-0,307352
<i>NDUFA5</i>	Candidato oncogene	0,489089	-0,625059
<i>NUMA1</i>	Candidato oncogene	-0,654166	0,301923
<i>PDIA4</i>	Candidato oncogene	0,92328	-0,426129
<i>PDXDC1</i>	Candidato oncogene	-1,04741	0,0900517

<i>PFKP</i>	Candidato oncogene	0,696462	-0,695186
<i>PHGDH</i>	Candidato oncogene	0,967333	-0,446461
<i>PLEC</i>	Candidato oncogene	-1,08644	0,501433
<i>RBM8A</i>	Candidato oncogene	0,915152	-1,02665
<i>RBMX</i>	Candidato oncogene	0,763363	-0,352321
<i>RPL13A</i>	Candidato oncogene	1,01085	-0,466545
<i>RPL18A</i>	Candidato oncogene	0,846148	-0,39053
<i>RPL28</i>	Candidato oncogene	0,814454	-0,480408
<i>RPL31</i>	Candidato oncogene	0,930739	-0,429572
<i>RPL34</i>	Candidato oncogene	0,537339	-0,937719
<i>RPS13</i>	Candidato oncogene	1,10187	-0,508554
<i>RPS15</i>	Candidato oncogene	1,00241	-0,64033
<i>RPS2</i>	Candidato oncogene	0,753536	-0,347786
<i>RPS27</i>	Candidato oncogene	0,889271	-0,88513
<i>RPS27A</i>	Candidato oncogene	0,808022	-0,372933
<i>RPS3A</i>	Candidato oncogene	0,86624	-0,399803
<i>RTN4</i>	Candidato oncogene	-0,768612	0,354744
<i>SLC9A3R1</i>	Candidato oncogene	-0,719453	0,332055
<i>SNRPD2</i>	Candidato oncogene	1,00281	-0,462838
<i>SNRPD3</i>	Candidato oncogene	1,02739	-0,474181
<i>SRSF1</i>	Candidato oncogene	0,771456	-0,356057
<i>SSB</i>	Candidato oncogene	0,676524	-0,312242
<i>SSBP1</i>	Candidato oncogene	1,01834	-0,679395
<i>TPSAB1</i>	Candidato oncogene	-1,20513	0,23027
<i>UBA2</i>	Candidato oncogene	0,721271	-0,332894
<i>ADH1B</i>	Candidato supressor de tumor	-1,80579	0,292506
<i>AHNAK</i>	Candidato supressor de tumor	-0,954288	0,44044
<i>AP3D1</i>	Candidato supressor de tumor	-0,893429	0,412352
<i>CDC42</i>	Candidato supressor de tumor	0,703037	-0,324479
<i>COL14A1</i>	Candidato supressor de tumor	-0,774699	0,357553
<i>COL6A3</i>	Candidato supressor de tumor	-0,896333	0,413692
<i>DCTN1</i>	Candidato supressor de tumor	-0,80907	0,373417
<i>DSP</i>	Candidato supressor de tumor	0,696074	-0,321265
<i>EIF4A1</i>	Candidato supressor de tumor	1,06495	-0,491515
<i>EIF5A</i>	Candidato supressor de tumor	1,02409	-0,472656
<i>GBE1</i>	Candidato supressor de tumor	-0,780749	0,360346
<i>HSPG2</i>	Candidato supressor de tumor	-0,658554	0,303948
<i>ITIH1</i>	Candidato supressor de tumor	-0,687314	0,317222
<i>KARS</i>	Candidato supressor de tumor	0,73156	-0,337643

<i>LRP1</i>	Candidato supressor de tumor	-1,24361	0,401969
<i>LUM</i>	Candidato supressor de tumor	-0,893151	0,412224
<i>S100A8</i>	Candidato supressor de tumor	0,747771	-0,784628
<i>SYNCRIP</i>	Candidato supressor de tumor	0,97713	-0,450983
<i>VCAN</i>	Candidato supressor de tumor	-1,14097	0,387391

TABELA SUPLEMENTAR 4D

DEP (Símbolo do gene)	Classificação NCG	Média de expressão em amostras HER2+ [log(Z-score)]	Média de expressão em amostras não-HER2+ [log(Z-score)]
<i>CTNNB1</i>	Oncogene e relacionado com saúde	-1,039	0,164283
<i>IDH1</i>	Oncogene	1,14687	-0,305833
<i>APMAP</i>	Candidato oncogene	1,26025	-0,336067
<i>DBI</i>	Candidato oncogene	0,91085	-0,350667
<i>GPX1</i>	Candidato oncogene	-0,898302	0,239547
<i>IQGAP1</i>	Candidato oncogene	-0,986824	0,263153
<i>NAMPT</i>	Candidato oncogene	1,15812	-0,308831
<i>SRRT</i>	Candidato oncogene	-0,943989	0,25173
<i>TROVE2</i>	Candidato oncogene	-0,956286	0,25501
<i>USP7</i>	Candidato oncogene	-1,0355	0,276133
<i>NDRG1</i>	Candidato supressor de tumor	0,687109	-0,772014
<i>ANPEP</i>	Candidato supressor de tumor	-0,87951	0,234536
<i>CCAR1</i>	Candidato supressor de tumor	-1,63568	-0,0795447
<i>EFTUD2</i>	Candidato supressor de tumor	-0,943435	0,251583
<i>GBE1</i>	Candidato supressor de tumor	1,06224	-0,283265
<i>HNRNPD</i>	Candidato supressor de tumor	-0,973158	0,259509
<i>KPNB1</i>	Candidato supressor de tumor	-0,976221	0,260326
<i>LRRC47</i>	Candidato supressor de tumor	-0,906374	0,2417
<i>MYH2</i>	Candidato supressor de tumor	-1,50284	-0,133412
<i>PDS5A</i>	Candidato supressor de tumor	-1,29805	0,0971787

TABELA SUPLEMENTAR 5 – VIAS E PROCESSOS IDENTIFICADOS COMO DIFERENCIALMENTE ATIVADOS PARA CADA SUBTIPO. (A) VIAS/PROCESSOS IDENTIFICADOS PARA O SUBTIPO LUMINAL A. (B) VIAS/PROCESSOS IDENTIFICADOS PARA O SUBTIPO LUMINAL B. (C) VIAS/PROCESSOS IDENTIFICADOS PARA O SUBTIPO TRIPLO NEGATIVO. (D) VIAS/PROCESSOS IDENTIFICADOS PARA O SUBTIPO HER2+ ENRIQUECIDO.

TABELA SUPLEMENTAR 5A – Luminal A

<i>Gene set</i> (Via ou processo biológico)	DEPs em <i>Overlap</i>	DEPs super expressas nesta via/processo	DEPs sub expressas nesta via/processo	Z-score	Valor de P-FDR
ACTIVATION OF THE mRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S	4	0	4	-2	2.26E-3
ADAPTIVE IMMUNE SYSTEM	15	10	5	1,290994449	8.81E-5
ADP SIGNALLING THROUGH P2Y PURINOCEPTOR 12	3	3	0	1,732050808	2.1E-3
APOPTOSIS	8	7	1	2,121320344	5.44E-5
APOPTOTIC EXECUTION PHASE	4	4	0	2	1.63E-3
BIOLOGICAL OXIDATIONS	6	5	1	1,632993162	4.93E-3
CELL CYCLE	16	13	3	2,5	3.36E-6
CELL CYCLE MITOTIC	15	12	3	2,323790008	1.76E-6
CELLULAR RESPONSE TO STARVATION	6	1	5	-1,632993162	1.32E-3
CELLULAR RESPONSES TO EXTERNAL STIMULI	19	10	9	0,229415734	3.38E-8
CLATHRIN MEDIATED ENDOCYTOSIS	5	5	0	2,236067977	5.21E-3
COMPLEX I BIOGENESIS	4	4	0	2	2.01E-3
COPI INDEPENDENT GOLGI TO ER RETROGRADE TRAFFIC	4	3	1	1	1.66E-3
COSTIMULATION BY THE CD28 FAMILY	4	4	0	2	4.19E-3
CYCLIN A B1 B2 ASSOCIATED EVENTS DURING G2 M TRANSITION	3	3	0	1,732050808	2.82E-3
CYTOKINE SIGNALING IN IMMUNE SYSTEM	12	8	4	1,154700538	1.17E-3
DARPP 32 EVENTS	3	3	0	1,732050808	2.56E-3

DEREGULATED CDK5 TRIGGERS MULTIPLE NEURODEGENERATIVE PATHWAYS IN ALZHEIMER S DISEASE MODELS	3	2	1	0,577350269	2.1E-3
DNA STRAND ELONGATION	3	1	2	-0,577350269	4.93E-3
ERK MAPK TARGETS	3	3	0	1,732050808	2.1E-3
ERKS ARE INACTIVATED	3	3	0	1,732050808	6.77E-4
EUKARYOTIC TRANSLATION ELONGATION	7	0	7	-2,645751311	1.14E-5
EUKARYOTIC TRANSLATION INITIATION	6	0	6	-2,449489743	4.4E-4
EXTRA NUCLEAR ESTROGEN SIGNALING	4	4	0	2	4.36E-3
EXTRACELLULAR MATRIX ORGANIZATION	8	4	4	0	1.1E-3
G ALPHA I SIGNALLING EVENTS	8	8	0	2,828427125	1.32E-3
G PROTEIN MEDIATED EVENTS	5	5	0	2,236067977	1.52E-4
GLOBAL GENOME NUCLEOTIDE EXCISION REPAIR GG NER	5	4	1	1,341640786	8.78E-4
GLUCONEOGENESIS	3	3	0	1,732050808	5.52E-3
GLUCOSE METABOLISM	6	6	0	2,449489743	1.22E-4
GLUTATHIONE SYNTHESIS AND RECYCLING	3	3	0	1,732050808	5.92E-4
GLYCOLYSIS	4	4	0	2	3.94E-3
HEMOSTASIS	11	10	1	2,713602101	2.26E-3
HIV INFECTION	7	-	-	-	1.32E-3
HOST INTERACTIONS OF HIV FACTORS	5	-	-	-	3.72E-3
HSP90 CHAPERONE CYCLE FOR STEROID HORMONE RECEPTORS SHR	4	3	1	1	2.01E-3
INFECTIOUS DISEASE	22	-	-	-	1.39E-8
INFLUENZA INFECTION	7	-	-	-	2.02E-4
INITIATION OF NUCLEAR ENVELOPE NE REFORMATION	4	4	0	2	6.48E-5
INNATE IMMUNE SYSTEM	30	23	7	2,921186973	1.31E-13
INTEGRATION OF ENERGY METABOLISM	8	8	0	2,828427125	2.3E-6
IRE1ALPHA ACTIVATES CHAPERONES	5	3	2	0,447213595	1.11E-4

LAGGING STRAND SYNTHESIS	3	1	2	-0,577350269	1.79E-3
LEISHMANIA INFECTION	9	-	-	#VALOR!	2.48E-4
M PHASE	11	11	0	3,31662479	7.52E-5
MAPK TARGETS NUCLEAR EVENTS MEDIATED BY MAP KINASES	3	3	0	1,732050808	4.6E-3
MEMBRANE TRAFFICKING	12	11	1	2,886751346	4.49E-4
METABOLISM OF AMINO ACIDS AND DERIVATIVES	12	4	8	-1,154700538	4.71E-6
METABOLISM OF CARBOHYDRATES	10	10	0	3,16227766	2.95E-5
METABOLISM OF RNA	18	11	7	0,942809042	9.43E-8
MHC CLASS II ANTIGEN PRESENTATION	7	6	1	1,889822365	6.13E-5
MITOTIC G1 PHASE AND G1 S TRANSITION	6	5	1	1,632993162	1.1E-3
MITOTIC G2 G2 M PHASES	7	6	1	1,889822365	6.77E-4
MITOTIC METAPHASE AND ANAPHASE	7	7	0	2,645751311	1.46E-3
MITOTIC PROMETAPHASE	6	6	0	2,449489743	3.57E-3
NEGATIVE REGULATION OF FGFR1 SIGNALING	3	3	0	1,732050808	5.26E-3
NEGATIVE REGULATION OF FGFR2 SIGNALING	3	3	0	1,732050808	5.52E-3
NEGATIVE REGULATION OF FGFR3 SIGNALING	3	3	0	1,732050808	3.97E-3
NEGATIVE REGULATION OF FGFR4 SIGNALING	3	3	0	1,732050808	4.6E-3
NERVOUS SYSTEM DEVELOPMENT	12	-	-	-	2.39E-4
NEURONAL SYSTEM	9	-	-	-	1.34E-3
NEUROTRANSMITTER RECEPTORS AND POSTSYNAPTIC SIGNAL TRANSMISSION	6	-	-	-	3.68E-3
NEUTROPHIL DEGRANULATION	23	17	6	2,293658555	5.41E-15
NONSENSE MEDIATED DECAY NMD	8	3	5	-0,707106781	3.36E-6
NUCLEAR ENVELOPE NE REASSEMBLY	5	5	0	2,236067977	6.1E-4
NUCLEOTIDE EXCISION REPAIR	5	4	1	1,341640786	2.02E-3
OPIOID SIGNALLING	8	8	0	2,828427125	7.68E-7
ORGANELLE BIOGENESIS AND MAINTENANCE	8	5	3	0,707106781	1.02E-3

PCNA DEPENDENT LONG PATCH BASE EXCISION REPAIR	3	1	2	-0,577350269	1.98E-3
POST TRANSLATIONAL PROTEIN MODIFICATION	21	13	8	1,091089451	2.95E-5
PP2A MEDIATED DEPHOSPHORYLATION OF KEY METABOLIC FACTORS	2	2	0	1,414213562	5.44E-3
PROCESSIVE SYNTHESIS ON THE C STRAND OF THE TELOMERE	3	1	2	-0,577350269	1.63E-3
PROCESSIVE SYNTHESIS ON THE LAGGING STRAND	3	1	2	-0,577350269	1.02E-3
PROGRAMMED CELL DEATH	11	10	1	2,713602101	2.56E-7
REGULATED NECROSIS	4	3	1	1	1.97E-3
REGULATION OF EXPRESSION OF SLITS AND ROBOS	6	1	5	-1,632993162	1.86E-3
REGULATION OF INSULIN LIKE GROWTH FACTOR IGF TRANSPORT AND UPTAKE BY INSULIN LIKE GROWTH FACTOR BINDING PROTEINS IGFBPS	5	2	3	-0,447213595	3.06E-3
REGULATION OF INSULIN SECRETION	5	5	0	2,236067977	6.58E-4
RESOLUTION OF AP SITES VIA THE MULTIPLE NUCLEOTIDE PATCH REPLACEMENT PATHWAY	3	1	2	-0,577350269	2.82E-3
RESPIRATORY ELECTRON TRANSPORT	5	4	1	1,341640786	1.66E-3
RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING AND HEAT PRODUCTION BY UNCOUPLING PROTEINS	5	4	1	1,341640786	3.38E-3
RESPONSE OF EIF2AK4 GCN2 TO AMINO ACID DEFICIENCY	5	0	5	-2,236067977	1.63E-3
RHO GTPASE CYCLE	13	10	3	1,941450687	4.12E-6
RHOBTB GTPASE CYCLE	5	4	1	1,341640786	2.74E-5
RHOBTB1 GTPASE CYCLE	3	3	0	1,732050808	2.31E-3
RHOG GTPASE CYCLE	4	2	2	0	4.19E-3

RIPK1 MEDIATED REGULATED NECROSIS	3	3	0	1,732050808	3.97E-3
RRNA PROCESSING	6	0	6	-2,449489743	3.68E-3
SELENOAMINO ACID METABOLISM	5	0	5	-2,236067977	2.56E-3
SIGNAL AMPLIFICATION	4	4	0	2	4.4E-4
SIGNALING BY RHO GTPASES MIRO GTPASES AND RHOBTB3	17	13	4	2,182820625	1.16E-6
SIGNALING BY ROBO RECEPTORS	7	2	5	-1,133893419	1.07E-3
SPRY REGULATION OF FGF SIGNALING	3	3	0	1,732050808	1.12E-3
SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	7	0	7	-2,645751311	3.22E-5
THE CITRIC ACID TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT	7	4	3	0,377964473	4.05E-4
THROMBIN SIGNALLING THROUGH PROTEINASE ACTIVATED RECEPTORS PARS	3	3	0	1,732050808	4.93E-3
TRANSCRIPTION COUPLED NUCLEOTIDE EXCISION REPAIR TC NER	5	4	1	1,341640786	6.58E-4
TRANSLATION	11	1	10	-2,713602101	4.12E-6
UNFOLDED PROTEIN RESPONSE UPR	5	3	2	0,447213595	1.17E-3
VESICLE MEDIATED TRANSPORT	13	12	1	3,050851079	4.05E-4

TABELA SUPLEMENTAR 5B – Luminal B

Gene set (Via ou processo biológico)	DEPs em Overlap	DEPs super expressas nesta via/processo	DEPs sub expressas nesta via/processo	Z-score	Valor de P-FDR
ABC FAMILY PROTEINS MEDIATED TRANSPORT	7	1	6	-1,889822365	1.45E-6
ABC TRANSPORTER DISORDERS	5	0	5	-2,236067977	7.45E-5
ANTIGEN PROCESSING CROSS PRESENTATION	5	0	5	-2,236067977	2.75E-4
APC C CDH1 MEDIATED DEGRADATION OF CDC20 AND OTHER APC C CDH1 TARGETED PROTEINS IN LATE MITOSIS EARLY G1	5	0	5	-2,236067977	6.37E-5
APC C MEDIATED DEGRADATION OF CELL CYCLE PROTEINS	5	0	5	-2,236067977	1.23E-4
APOPTOSIS	9	1	8	-2,333333333	2.51E-7
ASSEMBLY OF THE PRE REPLICATIVE COMPLEX	5	0	5	-2,236067977	4.62E-5
ASYMMETRIC LOCALIZATION OF PCP PROTEINS	5	0	5	-2,236067977	3.87E-5
AUF1 HNRNP D0 BINDS AND DESTABILIZES mRNA	5	0	5	-2,236067977	2.01E-5
BETA CATENIN INDEPENDENT WNT SIGNALING	7	0	7	-2,645751311	1.2E-5
C TYPE LECTIN RECEPTORS CLRS	6	0	6	-2,449489743	8.44E-5
CELL CYCLE MITOTIC	10	0	10	-3,16227766	1.3E-4
CELLULAR RESPONSE TO CHEMICAL STRESS	7	0	7	-2,645751311	1.94E-5
CELLULAR RESPONSE TO HYPOXIA	5	0	5	-2,236067977	6.72E-5
CELLULAR RESPONSE TO STARVATION	7	0	7	-2,645751311	1.78E-5
CELLULAR RESPONSES TO EXTERNAL STIMULI	18	2	16	-3,299831646	6.42E-10
CLATHRIN MEDIATED ENDOCYTOSIS	6	0	6	-2,449489743	9.96E-5
CLEC7A DECTIN 1 SIGNALING	6	0	6	-2,449489743	1.99E-5
CROSS PRESENTATION OF SOLUBLE EXOGENOUS ANTIGENS ENDOSOMES	5	0	5	-2,236067977	1.56E-5

CYCLIN A CDK2 ASSOCIATED EVENTS AT S PHASE ENTRY	5	0	5	-2,236067977	1.09E-4
CYTOPROTECTION BY HMOX1	6	0	6	-2,449489743	4.82E-5
DECTIN 1 MEDIATED NONCANONICAL NF KB SIGNALING	6	0	6	-2,449489743	1.61E-6
DEFECTIVE CFTR CAUSES CYSTIC FIBROSIS	5	-	-	-	3.09E-5
DEGRADATION OF AXIN	5	0	5	-2,236067977	2.01E-5
DEGRADATION OF BETA CATENIN BY THE DESTRUCTION COMPLEX	5	0	5	-2,236067977	1.09E-4
DEGRADATION OF DVL	5	0	5	-2,236067977	2.32E-5
DEGRADATION OF GLI1 BY THE PROTEASOME	5	0	5	-2,236067977	2.9E-5
DEUBIQUITINATION	8	0	8	-2,828427125	7.03E-5
DEVELOPMENTAL BIOLOGY	25	1	24	-4,6	1.37E-12
DNA REPLICATION PRE INITIATION	5	0	5	-2,236067977	1.09E-4
DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR	5	0	5	-2,236067977	9.21E-5
EPH EPHRIN SIGNALING	8	0	8	-2,828427125	3.28E-8
EPHB MEDIATED FORWARD SIGNALING	8	0	8	-2,828427125	9.4E-11
EUKARYOTIC TRANSLATION ELONGATION	8	0	8	-2,828427125	3.69E-8
EUKARYOTIC TRANSLATION INITIATION	9	0	9	-3	1.08E-8
FCGAMMA RECEPTOR FCGR DEPENDENT PHAGOCYTOSIS	8	1	7	-2,121320344	7.35E-7
G1 S DNA DAMAGE CHECKPOINTS	5	0	5	-2,236067977	4.62E-5
HEDGEHOG LIGAND BIOGENESIS	5	0	5	-2,236067977	4.05E-5
HEDGEHOG OFF STATE	7	1	6	-1,889822365	2.37E-6
HEDGEHOG ON STATE	5	0	5	-2,236067977	1.14E-4
HIV INFECTION	9	-	-	-	1.62E-6
HOST INTERACTIONS OF HIV FACTORS	7	-	-	-	5.87E-6
INFECTIOUS DISEASE	33	-	-	-	3.89E-23
INFLUENZA INFECTION	7	-	-	-	1.78E-5
INNATE IMMUNE SYSTEM	33	3	30	-4,700096711	7.44E-21
INTERLEUKIN 1 FAMILY SIGNALING	6	0	6	-2,449489743	8.44E-5

INTERLEUKIN 1 SIGNALING	5	0	5	-2,236067977	2.35E-4
IRE1ALPHA ACTIVATES CHAPERONES	4	0	4	-2	2.51E-4
LEISHMANIA INFECTION	10	-	-	-	1.61E-6
M PHASE	10	0	10	-3,16227766	1.9E-5
MAPK6 MAPK4 SIGNALING	5	0	5	-2,236067977	1.55E-4
MEMBRANE TRAFFICKING	16	1	15	-3,5	8.13E-9
METABOLISM OF AMINO ACIDS AND DERIVATIVES	16	1	15	-3,5	6.49E-12
METABOLISM OF POLYAMINES	6	0	6	-2,449489743	1.39E-6
METABOLISM OF RNA	15	0	15	-3,872983346	1.31E-7
MITOTIC G2 G2 M PHASES	7	0	7	-2,645751311	5.57E-5
MITOTIC METAPHASE AND ANAPHASE	10	0	10	-3,16227766	1.8E-7
NEDDYLATION	7	0	7	-2,645751311	1.23E-4
NEGATIVE REGULATION OF NOTCH4 SIGNALING	5	0	5	-2,236067977	2.01E-5
NERVOUS SYSTEM DEVELOPMENT	23	-	-	-	1.1E-16
NEUTROPHIL DEGRANULATION	20	2	18	-3,577708764	8.4E-15
NONSENSE MEDIATED DECAY NMD	7	0	7	-2,645751311	2.76E-6
ORC1 REMOVAL FROM CHROMATIN	5	0	5	-2,236067977	5.48E-5
PARASITE INFECTION	8	-	-	-	1.72E-7
PCP CE PATHWAY	7	0	7	-2,645751311	7.35E-7
POST TRANSLATIONAL PROTEIN MODIFICATION	22	0	22	-4,69041576	2.45E-8
PROGRAMMED CELL DEATH	10	1	9	-2,529822128	6.12E-8
REGULATION OF EXPRESSION OF SLITS AND ROBOS	12	0	12	-3,464101615	3.5E-11
REGULATION OF HMOX1 EXPRESSION AND ACTIVITY	5	0	5	-2,236067977	4.05E-5
REGULATION OF mRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS	5	0	5	-2,236067977	1.19E-4
REGULATION OF PTEN STABILITY AND ACTIVITY	5	0	5	-2,236067977	4.82E-5
REGULATION OF RAS BY GAPS	5	0	5	-2,236067977	4.62E-5

REGULATION OF RUNX2 EXPRESSION AND ACTIVITY	5	0	5	-2,236067977	6.12E-5
REGULATION OF RUNX3 EXPRESSION AND ACTIVITY	5	0	5	-2,236067977	2.01E-5
RESPONSE OF EIF2AK4 GCN2 TO AMINO ACID DEFICIENCY	7	0	7	-2,645751311	1.4E-6
RHO GTPASE EFFECTORS	11	0	11	-3,31662479	2.61E-7
RHO GTPASES ACTIVATE WASPS AND WAVES	7	0	7	-2,645751311	1.97E-9
RRNA PROCESSING	7	0	7	-2,645751311	6.37E-5
SARS COV 1 INFECTION	4	-	-	-	2.51E-4
SARS COV 2 INFECTION	5	-	-	-	4.56E-5
SARS COV INFECTIONS	6	-	-	-	1.02E-4
SCF SKP2 MEDIATED DEGRADATION OF P27 P21	5	0	5	-2,236067977	2.9E-5
SELENOAMINO ACID METABOLISM	7	0	7	-2,645751311	3.02E-6
SEPARATION OF SISTER CHROMATIDS	7	0	7	-2,645751311	4.56E-5
SIGNALING BY HEDGEHOG	7	0	7	-2,645751311	1.4E-5
SIGNALING BY NOTCH4	5	0	5	-2,236067977	9.67E-5
SIGNALING BY RHO GTPASES MIRO GTPASES AND RHOBTB3	14	2	12	-2,672612419	1.61E-6
SIGNALING BY ROBO RECEPTORS	14	0	14	-3,741657387	1.37E-12
SIGNALING BY WNT	8	0	8	-2,828427125	1.27E-4
SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	9	0	9	-3	7.69E-9
STABILIZATION OF P53	5	0	5	-2,236067977	2.32E-5
SWITCHING OF ORIGINS TO A POST REPLICATIVE STATE	5	0	5	-2,236067977	1.41E-4
THE ROLE OF GTSE1 IN G2 M PROGRESSION AFTER G2 CHECKPOINT	6	0	6	-2,449489743	5.87E-6
TNFR2 NON CANONICAL NF KB PATHWAY	6	0	6	-2,449489743	2.01E-5
TRANSCRIPTIONAL REGULATION BY RUNX2	6	0	6	-2,449489743	4.4E-5
TRANSCRIPTIONAL REGULATION BY RUNX3	5	0	5	-2,236067977	1.77E-4
TRANSLATION	12	0	12	-3,464101615	1.05E-8

TRANSLATION OF SARS COV 2 STRUCTURAL PROTEINS	4	-	-	-	1.56E-4
UCH PROTEINASES	6	0	6	-2,449489743	2.01E-5
VESICLE MEDIATED TRANSPORT	17	1	16	-3,638034376	7.69E-9

TABELA SUPLEMENTAR 5C – Triplo Negativo

Gene set (Via ou processo biológico)	DEPs em Overlap	DEPs super expressas nesta via/processo	DEPs sub expressas nesta via/processo	Z-score	Valor de P-FDR
ABC FAMILY PROTEINS MEDIATED TRANSPORT	16	14	2	3	2.11E-12
ACTIVATION OF THE mRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S	34	34	0	5,830951895	1.36E-48
ADAPTIVE IMMUNE SYSTEM	40	32	8	3,794733192	1.01E-12
ANTIGEN PROCESSING CROSS PRESENTATION	19	18	1	3,900067476	9.13E-16
APC C CDH1 MEDIATED DEGRADATION OF CDC20 AND OTHER APC C CDH1 TARGETED PROTEINS IN LATE MITOSIS EARLY G1	12	11	1	2,886751346	8.51E-10
APC C MEDIATED DEGRADATION OF CELL CYCLE PROTEINS	13	12	1	3,050851079	4.67E-10
APOPTOSIS	25	22	3	3,8	5.07E-18
ASPARAGINE N LINKED GLYCOSYLATION	21	16	5	2,400396793	1.29E-9
ASSEMBLY OF THE PRE REPLICATIVE COMPLEX	14	13	1	3,207134903	1.42E-12
ASYMMETRIC LOCALIZATION OF PCP PROTEINS	12	11	1	2,886751346	1.71E-10
AUF1 HNRNP D0 BINDS AND DESTABILIZES mRNA	13	11	2	2,496150883	1.66E-12
CELL CYCLE	38	32	6	4,217756949	1.21E-13
CELL CYCLE CHECKPOINTS	25	23	2	4,2	3.79E-13
CELL CYCLE MITOTIC	36	30	6	4	5.88E-15

CELLULAR RESPONSE TO CHEMICAL STRESS	25	21	4	3,4	3.14E-19
CELLULAR RESPONSE TO HYPOXIA	13	12	1	3,050851079	7.07E-11
CELLULAR RESPONSE TO STARVATION	67	66	1	7,941013883	7.1E-86
CELLULAR RESPONSES TO EXTERNAL STIMULI	108	98	10	8,467803948	1.51E-86
CLASS I MHC MEDIATED ANTIGEN PROCESSING PRESENTATION	26	24	2	4,314554973	1.16E-11
CLEC7A DECTIN 1 SIGNALING	14	13	1	3,207134903	1.91E-10
CROSS PRESENTATION OF SOLUBLE EXOGENOUS ANTIGENS ENDOSOMES	11	10	1	2,713602101	1.91E-10
CYCLIN A CDK2 ASSOCIATED EVENTS AT S PHASE ENTRY	15	13	2	2,840187787	1.86E-12
CYTOKINE SIGNALING IN IMMUNE SYSTEM	39	34	5	4,64371646	7.76E-14
CYTOPROTECTION BY HMOX1	17	14	3	2,667891875	2.78E-12
DECTIN 1 MEDIATED NONCANONICAL NF KB SIGNALING	14	13	1	3,207134903	3.92E-13
DEFECTIVE CFTR CAUSES CYSTIC FIBROSIS	12	11	1	2,886751346	9.73E-11
DEGRADATION OF AXIN	12	11	1	2,886751346	3.04E-11
DEGRADATION OF BETA CATENIN BY THE DESTRUCTION COMPLEX	13	12	1	3,050851079	3.1E-10
DEGRADATION OF DVL	12	11	1	2,886751346	4.53E-11
DEGRADATION OF GLI1 BY THE PROTEASOME	13	12	1	3,050851079	4.65E-12
DEVELOPMENTAL BIOLOGY	109	98	11	8,333088681	2.11E-65
DNA REPLICATION	16	15	1	3,5	4.53E-11
DNA REPLICATION PRE INITIATION	14	13	1	3,207134903	2.6E-11
DOWNSTREAM SIGNALING EVENTS OF B CELL RECEPTOR BCR	15	14	1	3,356585567	9.85E-13
EUKARYOTIC TRANSLATION ELONGATION	68	68	0	8,246211251	4.24E-109
EUKARYOTIC TRANSLATION INITIATION	70	70	0	8,366600265	2.72E-102
EXTRACELLULAR MATRIX ORGANIZATION	23	4	19	-3,127716211	2.98E-11

G1 S DNA DAMAGE CHECKPOINTS	12	11	1	2,886751346	3.17E-10
G2 M CHECKPOINTS	20	19	1	4,024922359	3.67E-13
GENE AND PROTEIN EXPRESSION BY JAK STAT SIGNALING AFTER INTERLEUKIN 12 STIMULATION	11	10	1	2,713602101	1,00E-11
HEDGEHOG LIGAND BIOGENESIS	12	11	1	2,886751346	1.95E-10
HEDGEHOG OFF STATE	14	12	2	2,672612419	8.87E-10
HIV INFECTION	26	-	-	-	1.93E-16
HOST INTERACTIONS OF HIV FACTORS	22	-	-	-	1.35E-17
INFECTIOUS DISEASE	122	-	-	-	1.22E-90
INFLUENZA INFECTION	69	-	-	-	1.41E-89
INNATE IMMUNE SYSTEM	65	53	12	5,085424118	3.91E-25
INTERLEUKIN 12 FAMILY SIGNALING	11	10	1	2,713602101	7.64E-10
INTERLEUKIN 12 SIGNALING	11	10	1	2,713602101	9.84E-11
IRE1ALPHA ACTIVATES CHAPERONES	11	9	2	2,110579412	1.91E-10
M PHASE	30	25	5	3,651483717	1.02E-13
MAPK6 MAPK4 SIGNALING	14	12	2	2,672612419	7.61E-11
METABOLISM OF AMINO ACIDS AND DERIVATIVES	86	79	7	7,763959671	3.96E-84
METABOLISM OF POLYAMINES	12	11	1	2,886751346	6.64E-11
METABOLISM OF RNA	105	102	3	9,661410722	4.86E-85
MITOTIC G1 PHASE AND G1 S TRANSITION	17	15	2	3,152963125	4.35E-11
MITOTIC G2 G2 M PHASES	19	15	4	2,523573073	5.47E-11
MITOTIC METAPHASE AND ANAPHASE	20	18	2	3,577708764	1.1E-10
mRNA SPLICING	20	20	0	4,472135955	2.54E-12
NEGATIVE REGULATION OF NOTCH4 SIGNALING	15	13	2	2,840187787	1.9E-15
NERVOUS SYSTEM DEVELOPMENT	96	-	-	-	7.56E-80
NEUTROPHIL DEGRANULATION	39	32	7	4,003203845	1.1E-19
NONSENSE MEDIATED DECAY NMD	66	66	0	8,124038405	1.64E-95
ORC1 REMOVAL FROM CHROMATIN	15	14	1	3,356585567	1.44E-13
PCP CE PATHWAY	15	13	2	2,840187787	5.08E-12
POST TRANSLATIONAL PROTEIN MODIFICATION	60	50	10	5,163977795	3.58E-16

PROCESSING OF CAPPED INTRON CONTAINING PRE mRNA	22	22	0	4,69041576	3.5E-12
PROGRAMMED CELL DEATH	25	22	3	3,8	1.82E-16
REGULATION OF EXPRESSION OF SLITS AND ROBOS	78	77	1	8,605305459	1.72E-102
REGULATION OF HMOX1 EXPRESSION AND ACTIVITY	13	12	1	3,050851079	1.27E-11
REGULATION OF mRNA STABILITY BY PROTEINS THAT BIND AU RICH ELEMENTS	16	13	3	2,5	1.71E-13
REGULATION OF PTEN STABILITY AND ACTIVITY	13	11	2	2,496150883	2.73E-11
REGULATION OF RAS BY GAPS	12	11	1	2,886751346	3.17E-10
REGULATION OF RUNX2 EXPRESSION AND ACTIVITY	13	12	1	3,050851079	5.2E-11
REGULATION OF RUNX3 EXPRESSION AND ACTIVITY	12	11	1	2,886751346	3.04E-11
RESPONSE OF EIF2AK4 GCN2 TO AMINO ACID DEFICIENCY	66	66	0	8,124038405	5.79E-101
RHO GTPASE EFFECTORS	25	23	2	4,2	3.16E-12
RNA POLYMERASE II TRANSCRIPTION	48	45	3	6,062177826	1.95E-10
RRNA PROCESSING	67	67	0	8,185352772	1.69E-76
RUNX1 REGULATES TRANSCRIPTION OF GENES INVOLVED IN DIFFERENTIATION OF HSCS	16	15	1	3,5	5.53E-11
S PHASE	18	16	2	3,299831646	1.68E-11
SARS COV INFECTIONS	16	-	-	-	2.74E-10
SCF SKP2 MEDIATED DEGRADATION OF P27 P21	13	12	1	3,050851079	4.65E-12
SELENOAMINO ACID METABOLISM	66	66	0	8,124038405	7.43E-95
SIGNALING BY INTERLEUKINS	31	26	5	3,771711343	2.2E-13
SIGNALING BY NOTCH	20	18	2	3,577708764	2.27E-10
SIGNALING BY NOTCH4	15	13	2	2,840187787	1.14E-12
SIGNALING BY RHO GTPASES MIRO GTPASES AND RHOBTB3	43	36	7	4,42245854	8.4E-17

SIGNALING BY ROBO RECEPTORS	83	81	2	8,671376539	1.41E-101
SIGNALING BY WNT	25	21	4	3,4	4.65E-12
SRP DEPENDENT COTRANSLATIONAL PROTEIN TARGETING TO MEMBRANE	74	74	0	8,602325267	1.94E-113
STABILIZATION OF P53	12	11	1	2,886751346	4.53E-11
SWITCHING OF ORIGINS TO A POST REPLICATIVE STATE	15	14	1	3,356585567	4.65E-12
TCF DEPENDENT SIGNALING IN RESPONSE TO WNT	19	16	3	2,98240454	6.54E-10
TNFR2 NON CANONICAL NF KB PATHWAY	14	13	1	3,207134903	2.12E-10
TRANSCRIPTIONAL REGULATION BY RUNX2	15	13	2	2,840187787	1.91E-10
TRANSLATION	90	89	1	9,27601447	2.43E-100
TRANSLATION OF SARS COV 2 STRUCTURAL PROTEINS	10	-	-	-	1.01E-9
UNFOLDED PROTEIN RESPONSE UPR	15	13	2	2,840187787	5.08E-12
VESICLE MEDIATED TRANSPORT	33	19	14	0,87038828	4.63E-10

TABELA SUPLEMENTAR 5D – HER2+ Enriquecido

Gene set (Via ou processo biológico)	DEPs em Overlap	DEPs super expressas nesta via/processo	DEPs sub expressas nesta via/processo	Z-score	Valor de P-FDR
REACTOME_APOPTOSIS	5	1	4	-1,341640786	3.2E-3
REACTOME_APOPTOTIC_CLEAVAGE_OF_CELLULAR_PROTEINS	4	1	3	-1	1.51E-4
REACTOME_APOPTOTIC_EXECUTION_PHASE	5	1	4	-1,341640786	1.23E-5
REACTOME_DEUBIQUITINATION	5	1	4	-1,341640786	2.3E-2
REACTOME_INITIATION_OF_NUCLEAR_ENVELOPE_REFORMATION	2	0	2	-1,414213562	4.23E-2
REACTOME_INNATE_IMMUNE_SYSTEM	13	2	11	-2,496150883	2.21E-5
REACTOME_METABOLISM_OF_RNA	17	1	16	-3,638034376	4.51E-12

REACTOME_METABOLISM_OF_VITAMINS_AND_COFACTORS	4	3	1	1	3.54E-2
REACTOME_mRNA_SPLICING	9	0	9	-3	3.04E-8
REACTOME_NEUTROPHIL_DEGRANULATION	10	2	8	-1,897366596	5.69E-6
REACTOME_NICOTINAMIDE_SALVAGING	2	1	1	0	4.23E-2
REACTOME_PEROXISOMAL_PROTEIN_IMPORT	3	1	2	-0,577350269	2.3E-2
REACTOME_PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_mRNA	10	0	10	-3,16227766	1.57E-8
REACTOME_PROGRAMMED_CELL_DEATH	5	1	4	-1,341640786	5.91E-3
REACTOME_PROTEIN_LOCALIZATION	4	2	2	0	2.34E-2
REACTOME_REGULATION_OF_mRNA_STABILITY_BY_PROTEINS_THAT_BIND_AU_RICH_ELEMENTS	3	0	3	-1,732050808	4.23E-2
REACTOME_RHO_GTPASE_CYCLE	6	1	5	-1,632993162	1.95E-2
REACTOME_SIGNALING_BY_RHO_GTPASES_MIRO_GTPASES_AND_RHOBTB3	7	1	6	-1,889822365	2.97E-2
REACTOME_TRNA_PROCESSING	4	0	4	-2	7.07E-3
REACTOME_TRNA_PROCESSING_IN_THE_NUCLEUS	4	0	4	-2	7.92E-4

TABELA SUPLEMENTAR 6 – RESULTADO DO RANQUEAMENTO ORIUNDO DA ANÁLISE DE MACHINE LEARNING DE CADA SUBTIPO.

Ranqueamento	Luminal A	Luminal B	Triplo Negativo	HER2+ Enriquecido
1 °	<i>TPP1</i>	<i>LGALS3</i>	<i>AOC3</i>	<i>FAM98B</i>
2 °	<i>NDUFS1</i>	<i>EPB41L2</i>	<i>ADH1B</i>	<i>NAMPT</i>
3 °	<i>HP1BP3</i>	<i>PDIA5</i>	<i>PTRF</i>	<i>HK1</i>
4 °	<i>DDX1</i>	<i>S100A10</i>	<i>FABP5</i>	<i>RTCB</i>
5 °	<i>IQGAP2</i>	<i>GSTP1</i>	<i>FSCN1</i>	<i>CCAR1</i>
6 °	<i>RTCB</i>	<i>NAMPT</i>	<i>EHD2</i>	<i>HMGCL</i>
7 °	<i>H2AFY</i>	<i>PGRMC1</i>	<i>ANP32E</i>	<i>RAB1A</i>
8 °	<i>MVP</i>	<i>DBI</i>	<i>EEF1A1</i>	<i>MYH2</i>
9 °	<i>HMGCL</i>	<i>UFC1</i>	<i>PDIA6</i>	<i>DBI</i>
10 °	<i>TROVE2</i>	<i>PSMD4</i>	<i>PIIB</i>	<i>TROVE2</i>
11 °	<i>ANXA7</i>	<i>DCPS</i>	<i>PFN1</i>	<i>DDX1</i>
12 °	<i>S100A9</i>	<i>SNX3</i>	<i>SNRPD1</i>	<i>ADH1B</i>
13 °	<i>CCAR1</i>	<i>TSN</i>	<i>FABP4</i>	<i>FABP4</i>
14 °	<i>LMNB2</i>	<i>EIF4H</i>	<i>RPS19</i>	<i>ANXA7</i>
15 °	<i>VCAN</i>	<i>ARPC5</i>	<i>SNX3</i>	<i>LMNB2</i>
16 °	<i>EEF1D</i>	<i>S100A11</i>	<i>HYOU1</i>	<i>AOC3</i>
17 °	<i>PRKAR2A</i>	<i>RPL22</i>	<i>CD59</i>	<i>ANP32E</i>
18 °	<i>MYOF</i>	<i>HYOU1</i>	<i>RAN</i>	<i>PSMD4</i>
19 °	<i>APEH</i>	<i>UGP2</i>	<i>RPL30</i>	<i>S100A10</i>
20 °	<i>FAM98B</i>	<i>EIF5A</i>	<i>RPL22</i>	<i>RNPEP</i>
21 °	<i>RNPEP</i>	<i>CD59</i>	<i>CAV1</i>	<i>FSCN1</i>
22 °	<i>ANXA11</i>	<i>IFI16</i>	<i>RPL23</i>	<i>CAV1</i>
23 °	<i>YBX1</i>	<i>RPL3</i>	<i>SSR4</i>	<i>ARMT1</i>
24 °	<i>ARMT1</i>	<i>ABCD3</i>	<i>TAGLN2</i>	<i>PGRMC1</i>
25 °	<i>AHNAK</i>	<i>MOGS</i>	<i>RPS28</i>	<i>TPP1</i>
26 °	<i>HK1</i>	<i>PFKP</i>	<i>RPS13</i>	<i>PFKP</i>
27 °	<i>DAK</i>	<i>ESYT1</i>	<i>LDHB</i>	<i>EPB41L2</i>
28 °	<i>SEC61A1</i>	<i>PHGDH</i>	<i>RBM8A</i>	<i>COLGALT1</i>
29 °	<i>IDH2</i>	<i>MYH2</i>	<i>PLEC</i>	<i>LRP1</i>
30 °	<i>RPL35</i>	<i>RPL13A</i>	<i>RPS12</i>	<i>EEF1A1</i>
31 °	<i>RPS20</i>	<i>PDIA4</i>	<i>ABCF1</i>	<i>NDUFS1</i>
32 °	<i>GSS</i>	<i>PFN1</i>	<i>FKBP1A</i>	<i>HP1BP3</i>
33 °	<i>FBP1</i>	<i>PLEC</i>	<i>FBP1</i>	<i>RPL35A</i>
34 °	<i>COLGALT1</i>	<i>RPL23A</i>	<i>TSN</i>	<i>SNRPD1</i>
35 °	<i>SRPRB</i>	<i>UBE2I</i>	<i>EIF4A1</i>	<i>PTRF</i>
36 °	<i>EHD2</i>	<i>TAGLN2</i>	<i>TPD52L2</i>	<i>SEC61A1</i>
37 °	<i>PIIB</i>	<i>DAD1</i>	<i>RPS21</i>	<i>LGALS3</i>
38 °	<i>RPL11</i>	<i>FABP5</i>	<i>RPL38</i>	<i>PDIA5</i>
39 °	<i>HSPE1</i>	<i>SNRPD1</i>	<i>GSS</i>	<i>RAN</i>
40 °	<i>LRP1</i>	<i>RPL35A</i>	<i>RPS15A</i>	<i>MYOF</i>
41 °	<i>PRDX4</i>	<i>ABCF1</i>	<i>RPL27</i>	<i>PRKAR2A</i>
42 °	<i>LDHB</i>	<i>CALR</i>	<i>ARFIP1</i>	<i>EHD2</i>
43 °	<i>PDIA6</i>	<i>RPS25</i>	<i>UBE2I</i>	<i>ARPC5</i>
44 °	<i>EEF1G</i>	<i>CSTB</i>	<i>RPS20</i>	<i>IDH2</i>
45 °	<i>RPS26</i>	<i>RPL27</i>	<i>LRP1</i>	<i>PRDX4</i>

46 °	<i>PTRF</i>	<i>RPS12</i>	<i>EIF5A</i>	<i>PFN1</i>
47 °	<i>RPS15A</i>	<i>RPL32</i>	<i>CALR</i>	<i>H2AFY</i>
48 °	<i>RPS27</i>	<i>ANP32E</i>	<i>RPL32</i>	<i>PHGDH</i>
49 °	<i>EIF4A1</i>	<i>PRDX4</i>	<i>ESYT1</i>	<i>PDIA4</i>
50 °	<i>CAV1</i>	<i>FKBP1A</i>	<i>EEF1G</i>	<i>DAD1</i>
51 °	<i>CALR</i>	<i>ARFIP1</i>	<i>RPS27</i>	<i>RPL3</i>
52 °	<i>S100A6</i>	<i>RPS13</i>	<i>IFI16</i>	<i>S100A9</i>
53 °	<i>RPL30</i>	<i>MANF</i>	<i>RPL13A</i>	<i>S100A11</i>
54 °	<i>PFKP</i>	<i>HEBP2</i>	<i>S100A11</i>	<i>DAK</i>
55 °	<i>EEF1A1</i>	<i>S100A6</i>	<i>RPL11</i>	<i>PDIA6</i>
56 °	<i>CSTB</i>	<i>SNX9</i>	<i>APEH</i>	<i>ARFIP1</i>
57 °	<i>SSR1</i>	<i>SSR4</i>	<i>SSR1</i>	<i>GSTP1</i>
58 °	<i>RPL35A</i>	<i>TPD52L2</i>	<i>CSTB</i>	<i>S100A6</i>
59 °	<i>RPS19</i>	<i>PPIB</i>	<i>HSPE1</i>	<i>ANXA11</i>
60 °	<i>RAN</i>	<i>RPS28</i>	<i>SNX9</i>	<i>TSN</i>
61 °	<i>HEBP2</i>	<i>SSR1</i>	<i>VCAN</i>	<i>RPL38</i>
62 °	<i>DAD1</i>	<i>COLGALT1</i>	<i>MOGS</i>	<i>RPS19</i>
63 °	<i>MANF</i>	<i>RPS15A</i>	<i>RAB1A</i>	<i>RPS21</i>
64 °	<i>PHGDH</i>	<i>HSPE1</i>	<i>RPL35</i>	<i>ABCF1</i>
65 °	<i>RPS25</i>	<i>RPL38</i>	<i>SRPRB</i>	<i>IQGAP2</i>
66 °	<i>RPL32</i>	<i>RPS21</i>	<i>UFC1</i>	<i>RBM8A</i>
67 °	<i>RPL23</i>	<i>RPL23</i>	<i>PRDX4</i>	<i>EIF5A</i>
68 °	<i>RPS12</i>	<i>RPS27</i>	<i>MANF</i>	<i>SSR4</i>
69 °	<i>SNX9</i>	<i>ARMT1</i>	<i>RPS26</i>	<i>RPL13A</i>
70 °	<i>AOC3</i>	<i>RBM8A</i>	<i>UGP2</i>	<i>IFI16</i>
71 °	<i>PDIA4</i>	<i>LDHB</i>	<i>PHGDH</i>	<i>RPS13</i>
72 °	<i>FKBP1A</i>	<i>HK1</i>	<i>DAK</i>	<i>HSPE1</i>
73 °	<i>RBM8A</i>	<i>FSCN1</i>	<i>MYOF</i>	<i>UGP2</i>
74 °	<i>FABP5</i>	<i>RPS26</i>	<i>RPL23A</i>	<i>RPS25</i>
75 °	<i>RPL27</i>	<i>RPS19</i>	<i>ABCD3</i>	<i>RPL30</i>
76 °	<i>RPS28</i>	<i>RPL11</i>	<i>EIF4H</i>	<i>RPL23</i>
77 °	<i>S100A11</i>	<i>EEF1G</i>	<i>AHNAK</i>	<i>DCPS</i>
78 °	<i>TPD52L2</i>	<i>RPL30</i>	<i>RPS25</i>	<i>CALR</i>
79 °	<i>MYH2</i>	<i>PRKAR2A</i>	<i>YBX1</i>	<i>RPS28</i>
80 °	<i>HYOU1</i>	<i>SRPRB</i>	<i>ARPC5</i>	<i>PLEC</i>
81 °	<i>SSR4</i>	<i>EIF4A1</i>	<i>MVP</i>	<i>EIF4H</i>
82 °	<i>SNX3</i>	<i>RNPEP</i>	<i>HEBP2</i>	<i>FABP5</i>
83 °	<i>RPL22</i>	<i>AOC3</i>	<i>DAD1</i>	<i>EEF1D</i>
84 °	<i>RPS13</i>	<i>FBP1</i>	<i>S100A6</i>	<i>CSTB</i>
85 °	<i>RPL23A</i>	<i>DAK</i>	<i>DCPS</i>	<i>HEBP2</i>
86 °	<i>DBI</i>	<i>RPS20</i>	<i>PDIA4</i>	<i>AHNAK</i>
87 °	<i>UGP2</i>	<i>IDH2</i>	<i>ARMT1</i>	<i>ESYT1</i>
88 °	<i>CD59</i>	<i>MYOF</i>	<i>IDH2</i>	<i>YBX1</i>
89 °	<i>RPS21</i>	<i>SEC61A1</i>	<i>NDUFS1</i>	<i>ABCD3</i>
90 °	<i>TAGLN2</i>	<i>LMNB2</i>	<i>RPL3</i>	<i>TAGLN2</i>
91 °	<i>RPL38</i>	<i>ADH1B</i>	<i>S100A9</i>	<i>TPD52L2</i>
92 °	<i>RPL3</i>	<i>PDIA6</i>	<i>ANXA11</i>	<i>EIF4A1</i>
93 °	<i>PSMD4</i>	<i>TROVE2</i>	<i>PFKP</i>	<i>MVP</i>
94 °	<i>ABCD3</i>	<i>RAN</i>	<i>EEF1D</i>	<i>SRPRB</i>

95 °	<i>UBE2I</i>	<i>FABP4</i>	<i>SEC61A1</i>	<i>MANF</i>
96 °	<i>FSCN1</i>	<i>GSS</i>	<i>COLGALT1</i>	<i>RPL23A</i>
97 °	<i>EPB41L2</i>	<i>RPL35</i>	<i>RNPEP</i>	<i>RPS26</i>
98 °	<i>PLEC</i>	<i>PTRF</i>	<i>RPL35A</i>	<i>SNX3</i>
99 °	<i>NAMPT</i>	<i>IQGAP2</i>	<i>PSMD4</i>	<i>RPL11</i>
100 °	<i>ABCF1</i>	<i>DDX1</i>	<i>PDIA5</i>	<i>RPL32</i>
101 °	<i>ADH1B</i>	<i>ANXA7</i>	<i>GSTP1</i>	<i>APEH</i>
102 °	<i>MOGS</i>	<i>EHD2</i>	<i>H2AFY</i>	<i>RPS20</i>
103 °	<i>ARFIP1</i>	<i>EEF1A1</i>	<i>LMNB2</i>	<i>VCAN</i>
104 °	<i>PFN1</i>	<i>HP1BP3</i>	<i>LGALS3</i>	<i>RPL35</i>
105 °	<i>FABP4</i>	<i>YBX1</i>	<i>PRKAR2A</i>	<i>RPL22</i>
106 °	<i>ESYT1</i>	<i>HMGCL</i>	<i>PGRMC1</i>	<i>RPS15A</i>
107 °	<i>PGRMC1</i>	<i>ANXA11</i>	<i>ANXA7</i>	<i>UBE2I</i>
108 °	<i>SNRPD1</i>	<i>APEH</i>	<i>TPP1</i>	<i>RPS27</i>
109 °	<i>RPL13A</i>	<i>RAB1A</i>	<i>DBI</i>	<i>SSR1</i>
110 °	<i>GSTP1</i>	<i>S100A9</i>	<i>S100A10</i>	<i>FKBP1A</i>
111 °	<i>IFI16</i>	<i>CAV1</i>	<i>IQGAP2</i>	<i>HYOU1</i>
112 °	<i>UFC1</i>	<i>EEF1D</i>	<i>HK1</i>	<i>SNX9</i>
113 °	<i>EIF5A</i>	<i>MVP</i>	<i>HP1BP3</i>	<i>GSS</i>
114 °	<i>ANP32E</i>	<i>TPP1</i>	<i>HMGCL</i>	<i>CD59</i>
115 °	<i>TSN</i>	<i>NDUFS1</i>	<i>MYH2</i>	<i>PPIB</i>
116 °	<i>RAB1A</i>	<i>CCAR1</i>	<i>NAMPT</i>	<i>EEF1G</i>
117 °	<i>EIF4H</i>	<i>RTCB</i>	<i>CCAR1</i>	<i>LDHB</i>
118 °	<i>PDIA5</i>	<i>VCAN</i>	<i>EPB41L2</i>	<i>FBP1</i>
119 °	<i>DCPS</i>	<i>LRP1</i>	<i>RTCB</i>	<i>RPS12</i>
120 °	<i>S100A10</i>	<i>AHNAK</i>	<i>DDX1</i>	<i>RPL27</i>
121 °	<i>LGALS3</i>	<i>FAM98B</i>	<i>FAM98B</i>	<i>UFC1</i>
122 °	<i>ARPC5</i>	<i>H2AFY</i>	<i>TROVE2</i>	<i>MOGS</i>

11 REFERÊNCIAS

- AKTER, K.; MANSOUR, M, A.; HYODO, T.; SENGA, T. FAM98A associates with DDX1-C14orf166-FAM98B in a novel complex involved in colorectal cancer progression. **Biochemistry & cell biology**, 2017.
- AL-WAJEE, A, S.; SALHIMI, S, M.; AL-MANSOUB, A, A.; KHALID, I, A.; HARVEY, T, M.; LATIFF, A.; ISMAIL, M, A. Comparative proteomic analysis of different stages of breast cancer tissues using ultra high performance liquid chromatography tandem mass spectrometer. **Plos One**, 2020.
- AN, X.; LIN, X.; YANG, A.; JIANG, Q.; GENG, B.; HUANG, M.; LU, J.; XIANG, Z.; YUAN, Z.; WANG, S.; SHI, Y.; ZHU, J. Cavin3 Suppresses Breast Cancer Metastasis via Inhibiting AKT Pathway. **Frontiers in pharmacology**, 11, 2020.
- ANAND, S.; SAMUEL, M.; ANG, C, S.; KEERTHIKUMAR, S.; MATHIVANAN, S. Label-based and label-free strategies for protein quantitation. **Proteome Bioinformatics, Methods in Molecular Biology**, 1549, 2017.
- ANDRADE, L, J, O.; OLIVEIRA, G, C, M.; BITTENCOURT, A, M, V.; MELO, P, R, S. Protein molecular modeling of genetic markers for thyroid cancer. **J Bras Patol Med Lab**, 52(4), 324-337, 2016.
- AZIZI, E.; CARR, A, J.; PLITAS, G.; MAZUTIS, L.; RUDENSKY, A, Y.; PE'ER, S. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. **Cell** 174, 1293–1308, 2018.
- BAI, L.; DENG, X.; LI, Q.; WANG, M.; AN, W.; A, D.; GAO, Z.; XIE, Y.; DAI, Y.; CONG, Y, S. Down-Regulation of the Cavin Family Proteins in Breast Cancer. **Journal of Cellular Biochemistry**, 113:322–328, 2012.
- BALIC, M.; THOMSEN, C.; WURSTLEIN, R.; GNENT, M.; HARBECK, N. St. Gallen/Vienna 2019: A Brief Summary of the Consensus Discussion on the Optimal Primary Breast Cancer Treatment. **Breast Care**, 14:103–110, 2019.
- BARH, D. Omics Approaches in Breast Cancer. Springer, 2014.
- BOUCHAL, P.; SCHUBERT, O, T.; FAKTOR, J.; BUDINSKA, E.; NENUTIL, R.; AEBERSOLD, R. Breast Cancer Classification Based on Proteotypes Obtained by SWATH Mass Spectrometry. **Cell Reports**, 28, 832–843, 2019.
- BRADFORD, M, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical biochemistry**, 72, 248-254, 1976.
- BRUNORO, G, V, F.; CARVALHO, P, C.; BARBOSA, V, C.; PAGNONCELLI, D.; GALLO, G, V, D, M.; PERALES, J.; ZAHEDI, R, P.; VALENTE, R, H.; NEVES-FERREIRA, A, G, C. Differential proteomic comparison of breast cancer secretome using a quantitative paired analysis workflow. **BMC Cancer**, 19:365, 2019.

CAO, B.; FANG, Z.; LIAO, P.; ZHOU, X.; XIONG, J.; XENG, S.; LU, H. Cancer-mutated ribosome protein L22 (RPL22/eL22) suppresses cancer cell survival by blocking p53-MDM2 circuit. **Oncotarget**, 8(53): 90651-90661, 2017.

CAPPELLETTI, V.; IORIO, E.; MIODINI, P.; SILVESTRI, M.; DUGO, M.; DAIDONE, M, G. Metabolic footprints and molecular subtypes in breast cancer. **Disease markers**, 10, 2017.

CARLSSON, H.; PETERSSON, S.; ENERBACK, C. Cluster analysis of S100 gene expression and genes correlating to psoriasin (S100A7) expression at different stages of breast cancer development. **International journal of oncology** 27: 1473-1481, 2005.

MUTHU, M.; CHERIYAN, V, T.; RISHI, A, K. CARP-1 / CCAR1: A biphasic regulator of cancer cell growth and apoptosis. **Oncotarget**, 6(9), 6490-6510, 2015.

CARVALHO, A, S.; BAETA, H.; SILVA, B, C.; MORAES, M, C, S.; BODO, C.; BECK, H, C.; RODRIGUEZ, M, S.; SARASWAT, M.; PANDEY, A.; MATTHIESEN, R. Extra-cellular vesicles carry proteome of cancer hallmarks. *Frontiers in Bioscience*, **Landmark**, 25, 398-436, 2020.

CASTRO, R, X. Adesão das usuárias das unidades básicas de saúde do município de Rubelita – Minas Gerais, aos métodos de detecção precoce de câncer de mama. Trabalho de conclusão de curso, Araçuí, 2011.

CGHFBC. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. **Lancet Oncol**, 13: 1141–51, 2012.

CHAKAFANA, G.; SHONHAI, A. The Role of Non-Canonical Hsp70s (Hsp110/Grp170) in Cancer. **Cells**. 28;10(2):254, 2021.

CHANG, H, Y.; JUAN, H, F. Proteomic Data Analysis: Functional Enrichment. In: *A Practical Guide to Cancer Systems Biology*, 2017.

CHEN, C.; HOU, J.; TANNER, J, J.; CHENG, J. Bioinformatics methods for mass spectrometry-based proteomics data analysis. **Int. J. Mol. Sci.**, 21, 2873, 2020.

CHUNG, L.; SHIBLI, S.; MOORE, K.; ELDER, E, E.; BOYLE, F, M.; MARSH, D, J.; BAXTER, R, C. Tissue biomarkers of breast cancer and their association with conventional pathologic features. *British Journal of Cancer*, 108, 351–360, 2013.

CONGLONG, X.; ZHANG, Z.; YU, H.; YU, M.; YANG, T.; MIAO, M.; SHI, H. Up-regulation of CNDP2 facilitates the proliferation of colon cancer. **BMC Gastroenterology**, 14:96, 2014.

- CORRÊA, S.; PANIS, C.; BINATO, R.; HERRERA, A, C.; PIZZATTI, L.; ABDELHAY, E. Identifying potential markers in Breast Cancer subtypes using plasma label-free proteomics. **Journal of Proteomics**, 151, 33–42, 2017.
- DA COSTA, G, G.; GOMIG, T, H, B.; KAVISKI, R.; SOUSA, K, S.; KUKOLJ, C.; LIMA, R, S.; URBAN, C, A.; CAVALLI, I, J.; RIBEIRO, E, M, S, F. Comparative proteomics of tumor and paired normal breast tissue highlights potential biomarkers in breast cancer. **Cancer genomics & proteomics**, 12: 251-262, 2015.
- DAI, X.; LI, T.; BAI, Z.; YANG, Y.; LIU, X.; ZHANG, J.; SHI, B. Breast cancer intrinsic subtype classification, clinical use and future trends. **Am J Cancer Res**; 5(10):2929-2943, 2015.
- DALAL, N.; JALANDRA, R.; SHARMA, M.; PRAKASH, H.; MAKHARIA, G, K.; SOLANKI, P, R.; SINGH, R.; KUMAR, A. Omics technologies for improved diagnosis and treatment of colorectal cancer: Technical advancement and major perspectives. **Biomedicine & Pharmacotherapy**, 131, 110648, 2020.
- DIAZ, M, I.; DÍAZ, P.; BENNETT, J, C.; URRÁ, H.; ORTIZ, R.; ORELLANA, P, C.; HETZ, .; QUEST, A, F, G. Caveolin-1 suppresses tumor formation through the inhibition of the unfolded protein response. **Cell Death and Disease**, 11:648, 2020.
- DIREITO, I.; MONTEIRO, L.; MELO, T.; FIGUEIRA, D.; LOBO, J.; ENES, V.; MOURA, G.; HENRIQUE, R.; SANTOS, M, A, S.; JERÓNIMO, C.; AMADO, F.; FARSILHA, M.; HELGUERO, L, A. Protein aggregation patterns inform about breast cancer response to antiestrogens and reveal the RNA ligase RTCB as mediator of acquired tamoxifen resistance. **Cancers**, 13, 3195, 2021.
- ELLIS, H.; MAHADEVAN, V. Anatomy and physiology of the breast. **Surgery (Oxford)**, 31(1), 11-14, 2013.
- EL-SHARKAWY, A, M. Principles and Practice of Cancer Prevention and Control. **OMICS Group eBooks**, 2014.
- ERCAN, C.; VAN DIEST, P, J.; VOOIJS, M. Mammary development and breast cancer: The Role of Stem Cells. **Curr Mol Med**; 11(4): 270–285, 2011.
- FANG, Y, F.; LEE, Y, C.; WANG, Y, Y.; WANG, C, H.; HOU, M, F.; YUAN, S, S.; YWHAE promotes proliferation, metastasis, and chemoresistance in breast cancer cells. **J Med Sci**.;35:408–416, 2019.
- FRAGOMENI, S, M.; SCIALIIS, A.; JERUSS, J, S. Molecular subtypes and local-regional control of breast cancer. **Surg Oncol Clin N Am**; 27(1): 95–120. 2019.
- FU, W.; SUN, H.; ZHAO, Y.; CHEN, M.; YANG, X.; LIU, Y.; JIN, W. BCAP31 drives TNBC development by modulating ligand-independent EGFR trafficking

and spontaneous EGFR phosphorylation. **Theranostics**, 9(22), 6468-6484, 2019.

FUNG, S, M.; WONG, X, Y.; LEE, S, X.; MIAO, H.; HARTMAN, M.; WEE, H, M. Performance of single-nucleotide polymorphisms in breast cancer risk prediction models: A systematic review and meta-analysis. **Cancer Epidemiol Biomarkers Prev**; 28(3), 2019.

GAM, L, H. Breast cancer and protein biomarkers. **World J Exp Med**, 20(2): 86-91, 2012.

GAN, Y.; YE, F.; HE, X, X.; The role of YWHAZ in cancer: A maze of opportunities and challenges. **Journal of Cancer**, 11, 2020.

GERMAIN, D, R.; GRAHAM, K.; GLUBRECHT, S, D.; HUGH, J, C; MACKEY, J, R.; GODBOUT R. DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. **Breast Cancer Res Treat**, 127:53–63, 2011.

GOLDHIRSCH, A.; WINER, E, P.; COATES, A, S. GELBER, R, D.; PICCART-GEBHART, M.; THURLIMANN, B.; SEN, H, J. Personalizing the treatment of women with early breast cancer: Highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. **Annals of Oncology**, 24: 2206–2223, 2013.

GOMIG, T, H, B.; CAVALLI, I, J.; SOUZA, L, R, R.; VIEIRA, E.; LUCENA, A, C, R.; BATISTA, M.; MACHADO, K, C.; MARCHINI, F, K.; MARCHI, F, A.; LIMA, R, S.; URBAN, C, A.; CAVALLI, L, R.; RIBEIRO, E, M, S, F. Quantitative label-free mass spectrometry using contralateral and adjacent breast tissues reveal differentially expressed proteins and their predicted impacts on pathways and cellular functions in breast cancer. **Journal of Proteomics**, 199, 1–14, 2019.

GRAVES, P, S.; HAYSTEAD, T, A, J.; Molecular Biologist's Guide to Proteomics. **Microbiology and molecular biology reviews**, 39–63, 2002.

GROSSET, A, A.; LABRIE, M.; VLADOIU, M, C.; YOUSEF, E, M.; GABOURY, L.; PIERRE, Y, S. Galectin signatures contribute to the heterogeneity of breast cancer and provide new prognostic information and therapeutic targets. **Oncotarget**, 7(14), 2016.

HADDAD, C, F. Hereditary breast cancer: review and current approach. **Mastology**;30:e20200042, 2020.

HANAHAAN, D.; WEINBERG, R, A. Hallmarks of Cancer: The Next Generation. **Cell**, 144, 4, 2011.

HAUKAAS, T H.; EUCEDA, L,R.; GISKEØDEGÅRD, G,F.; LAMICHHANE, S.; KROHN, M.; JERNSTRÖM, S.; AURE, M,R.; LINGJÆRDE, O,C.; SCHLICHTING, E.; GARRED, Ø.; DUE, E, U.; MILLS, G, B.; SAHLBERG, K, K.; BØRRESEN-DALE, A, L.; BATHEN, T, F. Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. **Cancer & Metabolism**, 4:12, 2016.

HEBERLE, H.; MEIRELLES, G, V.; SILVA, F, R, da.; TELLES, G, P.; MINGHIM, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics**, 16:169, 2015.

HESKE, C, M. Beyond energy metabolism: Exploiting the additional roles of NAMPT for cancer therapy. **Frontiers in oncology**, 9(1514), 2020.

HUANG, D, W.; SHERMAN, B, T.; LEMPICKI, R, A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature protocols**, 4(1), 44-57, 2009.

INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA. Estimativa 2020: Incidência de câncer no Brasil. Rio de Janeiro, **INCA**, 2019.

IQBAL, M, J.; JAVED, Z.; SADIA, H.; QURESHI, I, A.; IRSHAD, A.; AHMED, R.; MALIK, K.; ABBAS, A.; PEZZANI, R.; SHARIFI-RAD, J. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future. **Cancer Cell Int**, 21:270, 2021.

JAIN, K, K.; Role of proteomics in diagnosis of cancer. **Technology in Cancer Research & Treatment**, 1(4), 2002.

JERBY, L.; WOLF, L.; DENKERT, G.; STEIN, G, Y.; HILVO, M.; ORESIC, M.; GEIGER, T.; RUPPIN, R. Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. **Cancer Res**; 72(22), 2012.

JUNAID, M, A.; CLARK, G, M.; PULLARKAT, R, K. A lysosomal pepstatin-insensitive proteinase as a novel biomarker for breast carcinoma. **The International Journal of Biological Markers**, 15(2), 129-134, 2000.

KAMINKÁS, M.; CISZEWKI, T.; LOPACKA-SZATAN, K.; MIOTLA, P.; STAROSLAWSKA, E. Breast cancer risk factors. **Prz Menopauzalny**; 14(3): 196-202, 2015.

KANG, G, J.; PARK, M, K.; BYUN, H, J.; KIM, H, J.; KIM, E, J.; YU, L.; KIM, B.; SHIM, J, G.; LEE, H.; LEE, C, H. SARNP, a participant in mRNA splicing and export, negatively regulates E-cadherin expression via interaction with pinin. **J Cell Physiol**. 1–13. 2019.

KANHAIYA, K.; CZEIZLER, E.; GRATIE, C.; PETRE, I. Controlling Directed Protein Interaction Networks in Cancer. **Scientific Reports**, 7: 10327, 2017.
KIM, J.; GAO, L.; TAN, K. Multi-Analyte Network Markers for Tumor Prognosis. **Plos ONE**, 7(12): e52973, 2012.

KOOMEN, J, M.; HAURA, E, B.; BEPLER, G.; SUTPHEN, R.; REMILY-WOOD, E, R.; BENSON, K.; HUSSEIN, M.; HAZLEHURST, L, A.; YEATMAN, T, J.; HILDRETH, L, T.; SELLERS, T, A.; JACOBSEN, P, B.; FENSTERMACHER, D, A.; DALTON, W, S. Proteomic Contributions to Personalized Cancer Care. **Molecular & Cellular Proteomics**, 7.10, 2008.

KOUROU, K.; EXARCHOS, T, P.; EXARCHOS, K, P.; KARAMOUZIS, M, V.; FOTIADIS, D, I. Machine learning applications in cancer prognosis and prediction. **Computational and Structural Biotechnology Journal**, 13, 8–17, 2015.

KOVACS, I, A.; LUCK, K.; SPIROHN, K.; WANG, Y.; POLLIS, C.; SCHLABACH, S.; BIAN, W.; KIM, D, K.; KISHORE, N.; HAO, T.; CALDERWOOD, M, A.; VIDAL, M.; BARABÁSI, A, L. Network-based prediction of protein interactions. **Nature communications**, 10:1240, 2019.

KRUG, K.; JAEHNIG, E, J.; SATPATHY, S.; ELLIS, M, J.; GILLETTE, M, A.; Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. **Cell**, 183, 1436–1456, 2020.

LANG, G, T.; JIANG, Y, Z.; SHI, J, X.; YANG, F.; LI, X, G.; PEI, Y, C.; ZHANG, C, H.; MA, D.; XIAO, Y.; HU, P, C.; WANG, H.; YANG, Y, S.; GUO, L, W.; LU, X, X.; XUE, M, Z.; WANG, P.; CAO, A, W.; LING, W.; WNAG, Z, W.; YU, K, D.; DI, G, H.; LI, D., Q.; WANG, Y, J.; YU, Y.; SHI, L, M.; HU, X.; HUANG, W.; SHAO, Z, M. Characterization of the genomic landscape and actionable mutations in Chinese breast cancers by clinical sequencing. **Nature communications**, 11:5679, 2020.

LEE, S, B.; BOSE, S.; AHN, S, H.; SON, B, H.; KO, B, S.; KIM, H, J.; CHUNG, I, J.; KIM, J.; LEE, W.; KO, M, S.; LEE, K.; CHANG, S.; PARK, H, S.; LEE, J, W.; KIM, D, C. Breast cancer diagnosis by analysis of serum N-glycans using MALDI-TOF mass spectroscopy. **Plos One**, 2021.

LI, H.; ZHANG, H.; HUANG, G.; BING, Z.; LUO, H.; XU, D.; LIU, J.; AN, X. Loss of RPS27a expression regulates cell cycle, apoptosis, and proliferation via the RPS27a-RPL11-MDM2-p53 pathway in lung adenocarcinoma cells. **Research Square**, 2020.

LI, J.; GUAN, X.; FAN, Z.; CHING, L, M.; LI, Y.; WANG, Z.; CAO, W, M.; LIU, D, X. Non-invasive biomarkers for early detection of breast cancer. **Cancers**, 12, 2767, 2020.

LI, Y.; HUANG, J.; SUN, J.; XIANG, S.; YANG, D.; YING, X.; LI, H.; REN, G. The transcription levels and prognostic values of seven proteasome alpha subunits in human cancers. **Oncotarget**, 8(3), 4501-4519, 2017.

LIBERZON, A.; SUBRAMANIAN, A.; PINCHBACK, R.; PABLO, T.; MESIROV, J, P. Molecular signatures database (MSigDB) 3.0. **Bioinformatics**, 27(12), 1739–1740, 2011.

LIN, Z.; PENG, R.; SUN, Y.; ZHANG, L.; ZHANG, Z. Identification of ribosomal protein family in triple-negative breast cancer by bioinformatics analysis. **Bioscience Reports**, 41, 2021.

LIU, R, A.; GRAHAM, K.; GLUBRECHT, D, D.; GERMAIN, D, R.; MACHEY, J, R.; GODBOUT, R. Association of FABP5 expression with poor survival in triple-negative breast cancer. **AJP**, 178(3), 2011.

LOBO, M, D, P.; MORENO, F, B, M, B.; SOUZA, G, H, M, F, S.; VERDE, S, M, M, L.; MOREIRA, R, A.; MONTEIRO-MOREIRA, A, C, O. Label-Free Proteome analysis of Plasma from Patients with Breast cancer: stage-specific Protein expression. **Front. Oncol.** 7:14, 2017.

LÓPEZ-CORTÉS, A.; CABRERA-ANDRADE, A.; VÁZQUEZ-NAYA, J, M.; PAXOS, A.; GONZÁLES-DÍAZ, H.; PAZ-Y-MIÑO, C.; GUERRERO, S.; PÉREZ-CASTILLO, Y.; TEJERA, E.; MUNTEANU, C, R. Prediction of breast cancer proteins involved in immunotherapy, metastasis, and RnA-binding using molecular descriptors and artificial neural networks. **Scientific Reports**,10:8515, 2020.

LUALDI, M.; FASANO, M. Statistical analysis of proteomics data: A review on feature selection. **Journal of Proteomics**;198(30), 18-26, 2019.

LUCK, K.; SHEYNKMAN, G, M.; ZHANG, I.; VIDAL, M. Proteome-scale human interactomics. **Trends Biochem Sci**; 42(5): 342–354, 2017.

MACKLIN, A.; KHAN, S.; KISLINGER, T. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. **Clin Proteom**, 17:17, 2020.

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

MANNELL, A. The role of Ki-67 in breast cancer. **SAJS**, 54(2), 2016.

MCVEIGH, T, P.; KERIN, M, J. Clinical use of the Oncotype DX genomic test to guide treatment decisions for patients with invasive breast cancer. **Breast Cancer - Targets and Therapy**:9 393–400, 2017.

MERWE, D, E, V, der.; OIKONOMOPOULOU, K.; MARSHALL, J.; DIAMANDIS, E, P. Mass spectrometry: uncovering the cancer proteome for diagnostics. **Advances in cancer research**, 2007.

MILIOLI, H, H.; SOUSA, K, S.; KAVISKI, R.; OLIVEIRA, N, C, D, S.; URBAN, C, A.; LIMA, R, S.; CAVALLI, I, J.; RIBEIRO, E, M, S, F. Comparative Proteomics of Primary Breast Carcinomas and Lymph Node Metastases Outlining Markers of Tumor Invasion. **Cancer genomics & proteomics**, 12: 89-102 (2015).

MIYAKE, T.; NAKAYAMA, T.; KAGARA, N.; YAMAMOTO, N.; NAKAMURA, Y.; OTANI, Y.; UJI, K.; NAOI, Y.; SHIMODA, M.; MARUYAMA, N.; SHIMOMURA, A.; SHIMAZU, K.; KIM, S, J.; NOGUSHI, S. Association of GSTP1 methylation with aggressive phenotype in ER-positive breast cancer. **Anticancer research**, 33: 5617-5624, 2013.

MIYAKE, T.; NAKAYAMA, T.; NAOI, Y.; TAMAMOTO, N.; OTANI, Y.; KIM, S, J.; SHIMAZU, K.; SHIMOMURA, A.; MARUYAMA, N.; TAMAKI, Y.; NOGUCHI, S. GSTP1 expression predicts poor pathological complete response to neoadjuvant chemotherapy in ER-negative breast cancer. **Cancer science**, 103(5), 913-920, 2012.

MOLINIE, N.; GAUTREAU, A. The Arp2/3 regulatory system and its deregulation in cancer. **Physiol Rev**, 98: 215–238, 2018.

MOMENIMOVAHED, Z.; SALEHINIA, H. Epidemiological characteristics of and risk factors for breast cancer in the world. **Breast Cancer - Targets and Therapy**:11 151–164, 2019.

NASCIMENTO, R, G.; OTONI, K, N. Histological and molecular classification of breast cancer: what do we know? **Mastology**;30: e20200024, 2020.

NEVE, R,M.; CHIN ,K.; FRIDLAND, J.; YEH, J.; BAEHNER, F, L.; FEVR, T.; CLARK, L.; BAYANI, N.; COPPE, J, P.; TONG, F.; SPEED, T.; SPELLMAN, P, T.; DEVRIES, S.; LAPUK, A.; WANG, N,J.; KUO, W, L.; STILWELL, J, L.; PINKEL, D.; ALBERTSON, D, G.; WALDMAN, F, M.; MCCORMICK, F.; DICKSON, R, B.; JOHNSON, M,D.; LIPPMAN, M.; ETHIER, S.; GAZDAR, A.; GRAY, J, W.; A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. **Cancer cell** 10, 515–527, 2006.

NISHIMURA, Y.; KOMATSU, S.; ICHIKAWA, D.; NAGATA, H.; HIRAJIMA, S.; TAKESHIDA, H.; KAWAGUCHI, T.; ARITA, T.; KONISHI, H.; KASHIMOTO, K.; SHIOZAKI, A.; FUJIWARA, H.; OKAMOTO, K.; TSUDA, H.; OTSUJI, E. Overexpression of YWHAZ relates to tumor cell proliferation and malignant outcome of gastric carcinoma. **British Journal of Cancer**,108, 1324–1331, 2013.

OSZ, A.; LANCZKY, A.; GYORFFY, B. Survival analysis in breast cancer using proteomic data from four independent datasets. PREPRINT, 2020.
OTTEWELL, P, D.; O'DONNELL, L.; HOLEN, I. Molecular alterations that drive breast cancer metastasis to bone. **BoneKEY Reports** 4(643), 2015.

PALAMIDESSI, A.; FRITTOLI, E.; GARRÉ, M.; FARETTA, M.; MIONE, M.; TESTA, I.; DIASPRO, A.; LANZETTI, L.; SCITA, G.; FIORE, P, D, D.; Endocytic trafficking of RAC is required for the spatial restriction of signaling in cell migration. **Cell** 134, 135–147, 2008.

PANIS, C.; PIZZATTI, L.; HERRERA, A, C.; CORRÊA, S.; BINATO, E.; ABDELHAY, E. Label-Free Proteomic Analysis of Breast Cancer Molecular Subtypes. **J. Proteome Res.**, 13, 4752–4772, 2014.

PAPALAZAROU, V.; MACHESKY, L, M. The cell pushes back: The Arp2/3 complex is a key orchestrator of cellular responses to environmental forces. **Current Opinion in Cell Biology**, 68:37–44, 2021.

PARISE, C, A.; CAGGIANO, V. Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers. **Journal of Cancer Epidemiology**; 2014, 469251, 2014.

PARSONS, J.; FRANCAVILLA, C. 'Omics approaches to explore the breast cancer landscape. **Cell Dev. Biol.** 7:395, 2020.

PEROU, C, M.; SORLIE, T.; EISEN, M, B.; de RIJN, M, V.; JEFFREY, S, S.; REES, C, A.; POLLACK, J, R.; ROSS, D, T.; JOHNSEN, H.; AKSLEN, L, A.; FLUGE, O.; PERGAMENSHIKOV, A.; WILLIAMS, C.; ZHU, S, X.; LONNING, P, E.; BORRENSEN-DALE, A, L.; BROWN, P, O.; BOTSTEIN, D. Molecular portraits of human breast tumors. **Nature**, 406, 2000.

PLAYDON, M, C.; ZIEGLER, R, G.; SOMPSON, J, N.; STOLZENBERG-SOLOMON, R.; THOMPSON, H, J.; IRWIN, M, L.; MAYNE, S, T.; HOOVER, R, N.; MOORE, S, C. Nutritional metabolomics and breast cancer risk in a prospective study. **Am J Clin Nutr**;106: 637–49, 2017.

POLYAK, A. Breast cancer: origins and evolution. **J. Clin. Invest.** 117:3155–3163, 2007.

POWELL, C, A.; NASSER, M, W.; ZHAO, H.; WOCHNA, J, C.; ZHANG, X.; SHAPIRO, S.; SHILO, K.; GANJU, R, G. Fatty acid binding protein 5 promotes metastatic potential of triple negative breast cancer cells through enhancing epidermal growth factor receptor stability. **Oncotarget**, 6(8), 2015.

PRIETO, D, A. Mass spectrometry in cancer biomarker research: a case for immunodepletion of abundant blood-derived proteins from clinical tissue specimens. **Biomark Med**; 8(2): 269–286, 2014.

RAMDAS, P.; RADHAKRISHNAN, A, K.; SANI, A, A, A.; KUMARI, M.; RAO, J, S, A.; RAHMAN, P, S, A. Advancing the role of gamma-tocotrienol as proteasomes inhibitor: a quantitative proteomic analysis of MDA-MB-231 human breast cancer cells. **Biomolecules**, 10, 19; 2020.

RAO, S.; LEE, S, Y.; GUTIERREZ, A.; PERRIGOUE, J.; THAPA, R, J.; TU, Z.; JEFFERS, J, R.; RHODES, M.; ANDERSON, S.; ORAVECZ, T.; HUNGER, S, P.; TIMAKHOV, R, A.; ZHANG, R.; BALACHANDRAN, S.; ZAMBETTI, G, P.; TESTA, J, R.; LOOK, A, T.; WIEST, D, L. Inactivation of ribosomal protein L22 promotes transformation by induction of the stemness factor, Lin28B. **Blood**, 120: 3764-3773, 2013.

RICHARD, A, L.; ECKHARDT, M.; KROGAN, N, J. Mass spectrometry-based protein–protein interaction networks for the study of human diseases. **Mol Syst Biol.**17: e8792, 2021.

RISHA, Y.; MIMIC, Z.; GHOBADLOO, S, M.; BEREZOVSKI, M, V. The proteomic analysis of breast cell line exosomes reveals disease patterns and potential biomarkers. **Scientific Reports**, 10:13572, 2020.

SAHARA, S.; AOTO, M.; EGUCHI, Y.; IMAMOTO, N.; YONEDA, Y.; TSUJIMOTO, Y. Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation. **Nature**, 401, 168-173, 1999.

SCHMID, S, L. Reciprocal regulation of signaling and endocytosis: Implications for the evolving cancer cell. **J. Cell Biol.**, 216(9), 2623–2632, 2016.

SERRANO-CARBAJAL, E, A.; ESPINAL-ENRIQUEZ, J.; HERNÁNDEZ-LEMUS, E. Targeting metabolic deregulation landscapes in breast cancer subtypes. **Front. Oncol.** 10:97, 2020.

SIBIO, A.; ABRIATA, G.; BUFFA, S.; VINIEGRA, M.; FORMAN, D.; SIERRA, M, S. Etiology of breast cancer (c50) in central and south america. In: cancer in central and south America. Lyon: International agency for research on cancer. Disponível em: http://www-dep.iarc.fr/CSU_resources.htm, 2016.

SIMPSON, P, T.; REIS-FILHO, J, S.; GALE, T.; LAKHANI, S, R. Molecular evolution of breast cancer. **J Pathol**; 205: 248–254, 2005.

SOLIMAN, H.; SHAH, V.; SRKALOVIC, G.; MAHTANI, R.; LEVINE, E.; MAVROMATIS, B.; SRINIVASIAH, J.; KASSAR, M.; GABORDI, R.; QAMAR, R.; UNTCH, S.; KLING, H, M.; TREECE, T.; AUDEN, W. MammaPrint guides treatment decisions in breast Cancer: Results of the IMPACt trial. **BMC Cancer**; 20:81, 2020.

SOYSAL, S.; PTP1B expression is an independent positive prognostic factor in human breast cancer. **Breast Cancer Res Treat.** 137(2), 2013.

STINGL, J. Estrogen and Progesterone in normal mammary gland development and in cancer. **Horm canc**, 2:85–90, 2010.

STRICKER, T, P.; BROWN, C, D.; BANDLAMUDI, C.; MCNERNEY, M.; KITTLER, R.; MONTOYA, V.; PETERSON, A.; GROSSMAN, R.; WHITE, K, P; Robust stratification of breast cancer subtypes using differential patterns of transcript isoform expression. **Plos Genet**, 13(3), 2017.

SUÁREZ-ARROYO, I, J.; FELIZ-MOSQUEA, Y, R.; PÉREZ-LASPIUR, J.; ARJU, R.; GIASHUDDIN, S.; MALDONADO,-MARTÍNEZ, G.; CUBANO, L, A.; SCHNEIDER, R, J.; MARTÍNEZ-MONTEMAYOR, M, M, M. The proteome signature of the inflammatory breast cancer plasma membrane identifies novel molecular markers of disease. **Am J Cancer Res**, 6(8):1720-1740, 2016.

SUZUKI, R.; ORSINI, N.; SAJI, S.; KEY, T, J.; WOLK, A.; Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status—A meta-analysis. **Int. J. Cancer**: 124, 698–712, 2009.

TELKERSEN, T.; PERNEMALM, M.; GROMOV, P.; BORRENSSEN-DALE, A, L; KROGH, A.; HAAKENSEN, V, D.; LETHIO, J.; PAPALEO, E.; GROMOVA, I. High-throughput proteomics of breast cancer interstitial fluid: identification of

tumor subtype-specific serologically relevant biomarkers. **Molecular Oncology** 15, 429–461, 2021.

TENGA, M, J.; LAZAR, L, M.; A Proteomic Snapshot of Breast Cancer Cell Cycle: The G1/S Transition Point. **Proteomics**. 13(1): 48–60, 2013.

TWIMAN, R, M. Proteomics. **Elsevier**, 2021.

TYANOVA, S.; ALBRECHTSEN, R.; KRONKVIST, P.; COX, J.; MANN, M.; GEIGER, T. Proteomic maps of breast cancer subtypes. **Nature communication**, 7:10259, 2016.

VASAIKAR, S, V.; STRAUB, P.; WANG, J.; ZHANG, B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. **Nucleic Acids Research**, Vol. 46, 2018.

WALLDEN, B.; STORHOFF, J.; NIELSEN, T.; DOWIDAR, N.; SCHAPER, C.; FERREE, S.; LIU, S.; LEUNG, S.; GEISS, G.; SNIDER, J.; VICKERY, T.; DAVIES, S, R.; MARDIS, E, R.; GNANT, M.; SESTAK, I.; ELLIS, M, J.; PEROU, C, M.; BERNARD, P, S.; PARKER, J, S.; Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. **BMC Medical Genomics**; 8:54, 2015.

WALTER, W.; SÁNCHEZ-CABO, R.; RICOTE, M. GOplot: an R package for visually combining expression data with functional analysis. **Bioinformatics**, 31(17), 2912–2914, 2015.

WANG, C, Q, LI, Y.; HUANG, B, F.; ZHAO, Y, M.; YUAN, H.; GOU, D.; SU, C, M, U, G, N, WNAG, Q.; LONG, T.; WNAG, Y.; TANG, C, H.; LI, X. EGFR conjunct FSCN1 as a novel therapeutic strategy in triple negative breast cancer. **Scientific Reports**, 15654, 2017.

WANG, C, Q.; TANG, C, H.; CHANG, H, T.; LI, X, N.; ZHAO, Y, M.; SU, N, X, X.; ZENG, Y.; DU, Z.; WANG, Y.; HUANG, B, F. Fascin -1 as a novel diagnostic marker of triple-negative breast cancer. **Cancer Medicine**; 5(8):1983–1988, 2015.

WANG, H.; XIE, B.; KONG, Y.; TAO, Y.; YANG, G.; GAO, M.; XU, H.; ZHAN, F.; SHI, J.; ZHANG, Y.; WU, X. Overexpression of RPS27a contributes to enhanced chemoresistance of CML cells to imatinib by the transactivated STAT3. **Oncotarget**, 7(14), 2015.

WAZIR, U.; AHMED, M, H.; BRIDGER, J, M.; HARVEY, A.; KIANG, W, G.; SHARMA, A, K.; MOKBEL, K. The Clinicopathological Significance of Lamin A/C, Lamin B1 and Lamin B receptor mRNA Expression in Human Breast Cancer. **Cellular & Molecular Biology Letters**, 2013.

WISNIEWSKI, J, R.; GOUGAZ, F, Z. Fast and sensitive total protein and peptide assays for proteomic analysis. **Anal. Chem**, 87, 4110–4116, 2015.

WU, J.; CHEN, J.; ZHANG, Y.; DAI, W.; LIN, X.; YANG, X.; CHEN, M. ADH2 is downregulated by methylation and acts as a novel biomarker for breast carcinoma prognosis. **Annals of Clinical & Laboratory Science**, 51(1), 2021.

WU, X.; HASAN, M, A.; CHEN, J, Y. Pathway and network analysis in proteomics. **Journal of Theoretical Biology**, 362, 44–52, 2014.

YANOVICH, G.; AGMON, H.; HAREL, M.; SONNENBLICK, A.; PERETZ, T.; GEIGER, T. Clinical Proteomics of Breast Cancer Reveals a Novel Layer of Breast Cancer Classification. **Cancer Res**, 78(20), 6001-6010, 2018.

YERSAL, O.; BARUTCA, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. **World J Clin Oncol**, 10; 5(3): 412-424, 2014.

ZAMANIAN-AZODI, M.; RAZAEI-TAVIRANI, R.; RAHMAT-RAD, S.; HASANZADEH, H.; TAVINANI, M, R.; SEYYDI, S, S. Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer. **Gastroenterol Hepatol**; 8(3):215-224, 2015.

ZHANG, L, Q.; YANG, H, Y.; YANG, S, Q.; WANG, Y.; CHEN, X, J.; LU, H, S.; ZHAO, L, P. CNDP2 acts as an activator for human ovarian cancer growth and metastasis via the PI3K/AKT pathway. **Technology in Cancer Research & Treatment**, 18: 1-10, 2019.

ZHANG, M.; LEE, A, V.; ROSEN, J, M. The Cellular Origin and Evolution of Breast Cancer. **Cold Spring Harb Perspect Med**;7:a027128, 2017.

ZHAO, Y.; YANG, X.; XU, X.; ZHANG, L.; YU, H.; MIAO, Z.; LI, D.; WANG, S. Deubiquitinase PSMD7 regulates cell fate and is associated with disease progression in breast cancer. **Am J Transl Res**, 12(9):5433-5448, 2020.

ZHENG, L.; DAI, H.; ZHOU, N.; LI, M.; SINGH, P.; QIU, J.; TSARK, W.; HUANG, Q.; KERNSTINE, K.; ZHANG, X.; LIN, D.; SHEN, B. Fen1 mutations result in autoimmunity, chronic inflammation and cancers. **Nature medicine**, 13(7), 821-819, 2007.

ZHU, Y.; GUO, M.; ZHANG, L.; XU, T.; WANG, L, XU, G. Biomarker triplet NAMPT/VEGF/HER2 as a de novo detection panel for the diagnosis and prognosis of human breast cancer. **Oncology reports** 35: 454-462, 2016.

ZIANG, Y, Z.; MA, D.; SUO, C.; SHI, L.; HUANG, W.; SHAO, Z, M. Genomic and transcriptomic landscape of triple negative breast cancers: Subtypes and treatment strategies. **Cancer Cell** 35, 428–440, 2019.

12 CAPÍTULO 2

“Comprehensive analysis of the large and small ribosomal proteins in breast cancer: Insights on proteomic and transcriptomic expression patterns, regulation, mutational landscape and prognostic significance”.

ABSTRACT

The *RPL* and *RPS* genes codify the small and large ribosomal proteins (RPs), the major components of the ribosomal structure. Several evidences demonstrated the involvement of the RPs in several malignancies and its pathogenesis, however, various aspects related to the function and clinical relevance of the RPs in breast cancer remains unclear. In our study, we integrated several online databases and bioinformatic approaches to perform a comprehensive analysis of the RPs in breast cancer, providing insights on the proteomic and transcriptomic expression of the RPs and evaluating its clinical significance and prognostic value, besides its clinicopathological associations, mutational and functional landscape and regulation. Through our expression analysis, we demonstrated that 33 RPs present differential expression at protein and mRNA level, and that the majority of the RPs, but not all, tend to be overexpressed in tumor tissue regarding non-tumor tissue, in addition to being markedly overexpressed in triple negative tumors. Our RP gene regulatory network indicates the transcription factors *MYC*, *ETS1* and *SPI1*, and the miRNAs has-let-7c-5p, has-mir-20b-5p and has-mir-4668-3p as main regulators of the RPs expression in breast cancer. The survival analysis showed that 26 RPs have potential as prognostic markers in a subtype-dependent way, and also that mutations in RP genes are frequent in breast tumors and related to overall survival and relapse-free status. The gene functional enrichment analysis indicated that the RPs are also potentially involved in breast tumorigenesis through ribosomal-independent mechanisms, like metabolism regulation and regulation of *SLITS* and *ROBOS* expression. Finally, we report genes that share genomic neighborhood with the RPs genes, including *EIF3S6*, *NPM1*, *FBL* and *HDAC1*, which are potentially functionally-related to the RPs and compose a list of candidates to future studies in the context of investigation of the ribosomal involvement in cancer. Together, our results suggest that the RPs have potential as biomarkers of diagnostic and prognostic breast cancer, and provide an in-depth view of the RPs significance in breast cancer.

12.1 INTRODUCTION

Along with rRNAs, the ribosomal proteins (RPs) are the elements that compound the ribosomes. At protein level, each eukaryotic ribosome is composed by the combination of a 40S small ribosomal subunit (RPS), responsible to interact with the tRNA anticodon-containing end to translate the information of the mRNA in an amino acid sequence, and a 60S large ribosomal subunit (RPL), that link the amino acids in polypeptide chains through its peptidyl transferase activity (KANG et al., 2021). According to the HUGO gene nomenclature committee (HGNC), the human genome contains a total of 35 RPS genes and 54 RPL genes (TWEEDIE et al., 2021).

The ribosomes are the complexes responsible to convert the genetic information into proteins, in a process that involves the assembly of the ribosome subunits, rRNA processing, RNA folding and interactions with auxiliary factors involved in the mRNA translation (SULIMA et al., 2018). However, beyond this structural and translational role, the RPs also have ribosome-independent functions, including regulation of proliferation, cell growth and differentiation, DNA repair and apoptosis (WANG et al., 2014). All the RPs activities occur under strong regulation, thus its dysregulation tends to presents a great impact in cell homeostasis, with RPs alterations being related to several diseases, including cancer (WARNER & MCINTOSH, 2009; WANG et al., 2014).

There is a wide variety of evidence that links RPs dysregulations with cancer, whether through down- or overexpression of RPs by mutation-associated defects, or by oncogenic hyperactivation of RPs biogenesis. In general, the RPs presents tumorigenic and prognostic roles associated with several neoplasias (SAHIN et al., 2005; HUANG et al., 2006; BEE et al., 2011; WU et al., 2011; DUTTON-REGESTER et al., 2014; RAO et al., 2016; HOFMAN et al., 2017). For example, in colon cancer, *RPL15* is overexpressed and involved in carcinogenesis via cell cycle and apoptosis regulation (DONG et al., 2019); in non-small cell lung cancer, overexpression of *RPL34* is related to malignant proliferation stimulation (YANG et al., 2016); *RPS7* is appointed to suppress tumorigenesis and metastasis through *PI3K/AKT* and *MAPK* signaling pathways in ovarian cancer (WANG et al., 2013). This also include breast cancer: *RPL15*

and *RPL27A* are related to invasion and metastasis (EBRIGHT et al., 2020; ZHAO et al., 2021), while *RPS9*, *RPS14*, *RPS27*, *RPL11* and *RPL14* were found downexpressed and correlated with poor prognosis (LIN et al., 2021). However, to date, our understanding of the involvement of each individual RP in breast cancer is still incomplete and fragmented.

Breast cancer is a major global health issue, being one of the most diagnosed and lethal cancers among women (SUNG, 2021). Breast tumors can be classified in molecular subtypes by gene expression (Sørli et al., 2003) and in immunohistochemical subtypes (BALIC et al., 2019) by the immunodetection of the Estrogen Receptor (ER) and Progesterone Receptor (PR), Human Epidermal Growth Factor Receptor 2 (*HER2*) and the marker of proliferation Ki-67 (*MKI67*). The use of the classical markers ER, PR, HER2 and Ki-67, together with clinicopathological parameters as tumor grade and stage, age and lymph node impairment, is highly relevant in clinics, yet insufficient to accurately define prognosis or classify subtypes (PARISE & CAGGIANO, 2014; FRAGOMENI et al., 2018). Therefore, the identification of new biomarkers of diagnostic and prognostic prediction can lead to relevant implications in the field of breast cancer treatment and patient's management (GAMBLE et al., 2021).

The present study aims to contribute to the understanding of the role of the RPs in breast tumorigenesis and its clinical implications in the field of biomarker discovery and outcome prediction. To this end, we comprehensively investigated the proteomic and transcriptomic expression patterns, potential functions, regulatory and mutational landscape, and prognostic values of the ribosomal proteins in breast cancer using mass spectrometry and transcriptomic data, bioinformatics approaches and online tools and databases.

12.2 MATERIALS AND METHODS

12.2.1 Patients and sample preparation

This study was approved by the National Commission of Ethics in Research (CAAE 19870319.3.0000.0102). All tissue samples were collected during surgery procedure at the Hospital Nossa Senhora das Graças in Curitiba/Paraná, Brazil, under patients signed informed consent. In total, nineteen

tumor samples and eighteen paired adjacent non-tumor samples from different patients, all diagnosed with Invasive Ductal Carcinoma (IDC), were used in this research. Among these, six were classified immunohistochemically (GOLDHIRSCH et al., 2013; BALIC et al., 2019) as triple negative breast cancer (TNBC), four as HER2+ enriched breast cancer (HER2+), five as luminal A breast cancer (LABC) and four as luminal B breast cancer (LBBC). All non-breast tissues were removed from the samples by macrodissection, and the remaining tissue was stored in RNeasy lysis buffer at -80°C until protein extraction.

The protein extracts were obtained through chemical and mechanical lysis, according to standardized protocol adapted from Gomig and coworkers (2019). Each sample was lysed in 4% SDS; 0.1M Tris-HCl, pH 7.5; 0.1M DTT (100µ to every 10mg of tissue) and then homogenized in the TissueLyser II sample disruptor (Qiagen) in 3-minutes cycles at 30 Hz, with each cycle interspersed with heating at 95°C. After, all samples were subjected to sonication and centrifugation to remove cell debris. The final peptide concentration of each extract was determined by the tryptophan quantification method (WISNIEWSKI & GOUGAZ, 2015). All subsequent sample preparation steps for mass spectrometry were performed as described by Gomig and coworkers (2019).

12.2.2 Protein identification and quantification by LC-MS/MS

The peptides were separated by liquid chromatography (3000 RSLCnano chromatograph, Thermo Scientific) and analyzed by mass spectrometry (Orbitrap Fusion Lumos mass spectrometer, Thermo Scientific) in the mass spectrometry facility (RPT02H) of Carlos Chagas Institute - Fiocruz, Paraná. All chromatography and mass spectrometry run parameters are described in **Supplementary Table 1**. All runs were performed in duplicate.

The resulting mass spectra were analyzed in the MaxQuant software (v. 1.6.17), and the protein identification was performed in the human UniProt protein database (7577 entries). Carbamidomethylation of cysteine was set as fixed modification, and oxidation of methionine and acetylation of protein N-terminal were set as variable modifications. An FDR of 1% was independently applied for

both peptide and protein identification, with at least seven amino acids required for peptide identification.

12.2.3 Proteomic data analysis and statistical approach

The raw proteomic data was processed in Perseus software (v. 1.6.15.0), with protein expression established based on the normalized spectral label-free protein intensity (LFQ intensity). Potential contaminants, reverse peptides and proteins only identified based on modified peptides were removed, and only proteins identified in at least 70% of the samples were maintained for analysis. LFQ intensities were then logarithmized ($\log(x)$) and normalized by Z-score, with missing values replaced by values from normal distribution (width = 0.3; down-shift = 1.8), as recommended by Perseus developers (TYANOVA & COX, 2018).

The ribosomal proteins differentially expressed proteins (DEPs) were obtained from two distinct approaches: the first was t-tests (P -value < 0.05; $\text{Log}^2\text{FC} \pm 0.58$) comparisons of tumor samples of each breast cancer subtype against its paired non-tumor counterparts (T vs. NT). The second approach was an ANOVA (FDR/ P -value < 0.05) comparison between tumor samples of each subtype (T vs. T), followed by Tukey's post-hoc test for identification of significant pairs. Correlations between protein levels of each RP were determined through Spearman's correlation test in the Hmisc R package (HARREL, 2021).

12.2.4 Breast Cancer Gene-Expression Miner analysis

The transcriptomic expression levels of the ribosomal proteins in breast cancer were obtained through the Breast Cancer Gene-Expression Miner database (bc-GenExMiner, v. 4.7) (JÉZÉQUEL et al., 2021). Data from The Cancer Genome Atlas (TCGA) (WEINSTEIN et al., 2013) and The Genotype-Tissue Expression (GTEx) (CARITHERS et al., 2021) compendium were retrieved by the database; the Dunnett-Tukey-Kramer's tests (P -value < 0.05) were applied to obtain RPs differentially expressed genes (DEGs), comparing tumor samples of each breast cancer subtype against non-tumor breast samples (T vs. NT) and tumor samples of each molecular breast cancer subtype (T vs. T). PAM50 classification method (WALLDEN et al., 2015) was used to determine the

molecular subtypes of the breast cancer samples. In this study, molecular and immunohistochemical subtypes were considered correspondents (luminal A = ER+, PR+, HER2-, Ki67 < 14%; luminal B = ER+, PR-, HER-, Ki67 > 14% or ER+, PR+/-, HER2+, any Ki67; HER2+ enriched = ER-, PR-, HER2+; Basal-like = ER-, PR-, HER2-). Correlations between the mRNA levels of each RP were performed according to the database algorithm.

The prognostic value of each RP was determined by means of univariate Cox proportional hazards model considering all RNA-microarray data available, with median as splitting criteria and P-value < 0.05. The Hazard ratios were retrieved as well. The association between the mRNA levels of the RPs were performed to the clinical outcomes of overall survival (OS), distant metastasis-free survival (DMFS) and disease-free survival (DFS), with and without sample subgrouping by subtypes (to both Sorlie's and PAM50 classification systems). Moreover, the association between mRNA levels and the clinicopathological parameters ER, PR and HER2 status (pos/neg), TP53 status (wild type/mutated), node involvement (pos/neg), patients age (< 51/ ≥ 51), tumor stage, Nottingham Prognostic Index (NPI) and Scarff-Bloom-Richardson grade (SBR) (P-value < 0.05) were also investigated in bc-GenExMiner (v. 4.7).

12.2.5 Genomic alteration analysis by cBioPortal database

In the cBioPortal (v. 3.7.21) database (CERAMI et al., 2012), the differentially expressed RPs were investigated in terms of genomic alterations and prognostic significance of somatic changes. Data from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) (PEREIRA et al., 2016) and TCGA (WEINSTEIN et al., 2013) were considered for analysis, using the search parameters 'mutations' and 'copy-number alterations'.

Moreover, through separating the samples into 'altered' and 'unaltered' groups, we determined the prognostic significance of the somatic alterations in RPs genes concerning overall survival and relapse-free survival (q-value < 0.05).

12.2.6 Gene enrichment and functional analysis of GSEA/MSigDB

The GSEA Molecular Signatures Databases (MSigDB, v. 7.4) (SUBRAMANIAN et al., 2005) was accessed to perform the enrichment analysis and functional predictions aiming to investigate the biological relevance of the differentially expressed RPs and its neighbor genes in breast cancer. GeneOntology terms (Molecular function, cellular component and biological processes) and REACTOME pathways (FDR/P-value < 0.05) were the datasets used to compute the overlaps, along with the compendia Expression Profiles of Global Cancer Map (Broad Institute). Also, genes that share 'genomic neighborhood' with the RPs genes were determined by computing overlaps between our gene set (the differentially expressed RP gene list) and the 'cancer gene neighborhoods' dataset of MSigDB (FDR/P-value < 0.05).

12.2.7 Prediction of regulation of RPs expression by miRNA and transcription factor

The identification of potential regulators of RP expression was performed through identification of differentially expressed miRNAs and transcription factors (TFs) in breast cancer, and subsequent intersection of several in silico and experimentally validated interactions databases, aiming to identify significant interactions between the regulators and RPs. The differentially expressed miRNAs and TFs were identified following the analysis pipeline of the GDCRNATools package described by Li and coworkers (2018), using the TCGA breast cancer dataset (P-value < 0.05; $\text{Log}^2\text{FC} \pm 0.58$).

The MIRgate (ANDRÉS-LEÓN et al., 2015) platform was used to determine potential miRNAs involved in RP regulation: OncomiRDB, Tarbase, Mirecords, Microtar, Targetscan and Miranda were the datasets resorted to analysis, with only interactions with agreement score (AS) higher than 1 ($\text{AS} > 1$) considered significant. To the TF-gene interactions, the HARMONIZE (ROUILLARD et al., 2016) platform was consulted to access the CHEA, ENCODE, JASPAR and TRANSFACT datasets, with only interactions simultaneously predicted in at least two datasets selected for subsequent study. Lastly, the main regulators-RPs pairs (outdegree > 5) were visualized in Cytoscape (v. 3.8.2) (SHANNON et al., 2003).

12.3 RESULTS

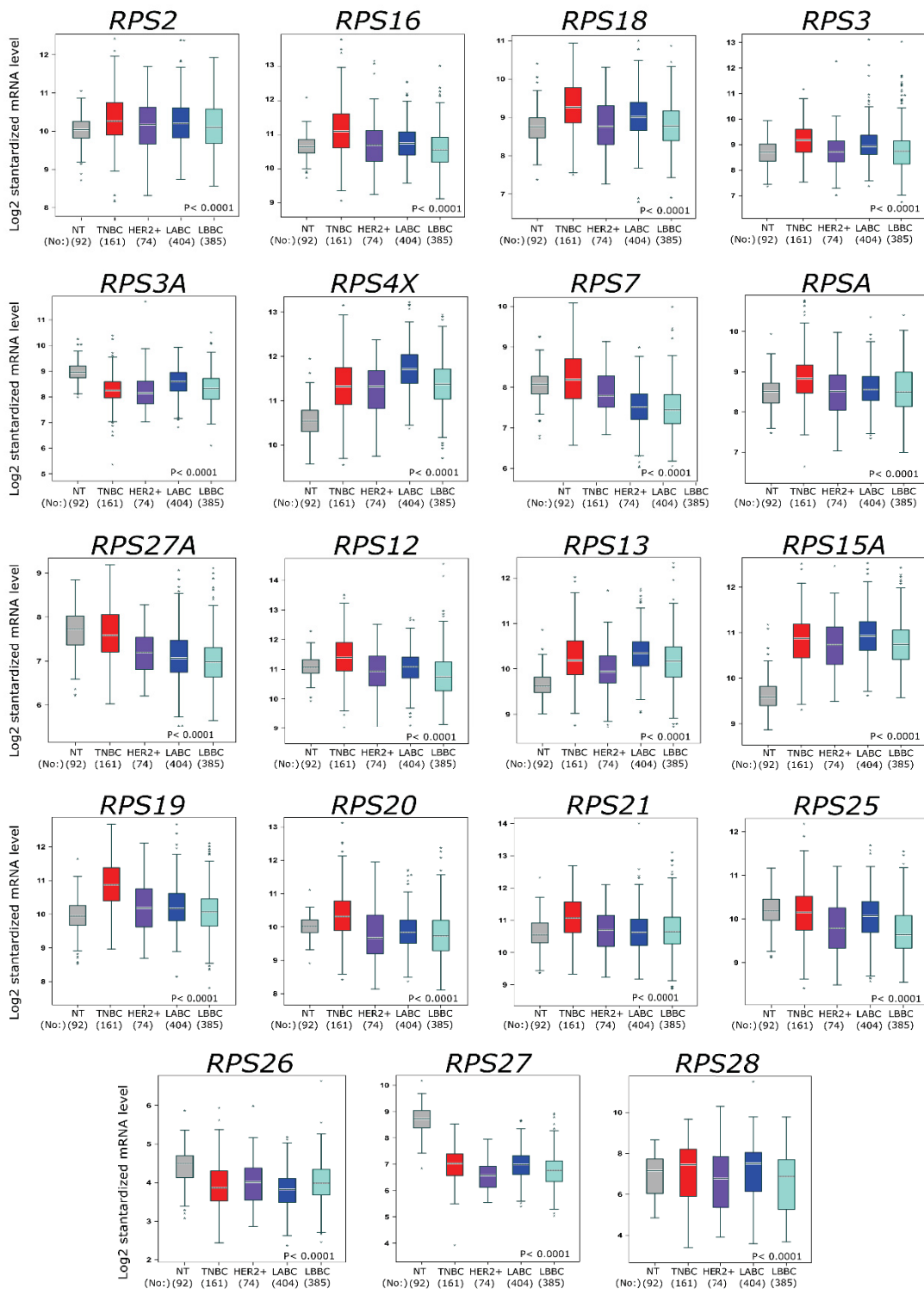
3.1 Transcriptomic and proteomic expression levels of large and small ribosome proteins in cancer

The differential mRNA and protein expression of large and small ribosomal proteins were evaluated in the bc-GenExMiner (v. 4.7) database and our mass-spectrometry data, respectively, comparing breast cancer tumors of each subtype and non-tumor tissue (T vs. NT) as well as the different breast cancer subtypes tumors (T vs. T). Only RPs differentially expressed at both mRNA and protein levels were maintained to subsequent analysis and discussion. The total number of differentially expressed RPs covered in this study can be verified in the **Figure 1A**, and detailed data referring to the expression analysis are detailed in the **Supplementary table 2**.

As shown in the **Figure 1B**, 13 RPs were overexpressed at proteomic level in the T vs. NT comparison (P -value < 0.05 ; $\text{Log}_2\text{FC} \pm 0.58$). Among these, 12 were differentially expressed in all T vs. NT comparisons of all subtypes. In the T vs. T comparison, 22 RPs were found with significant differential expression (P -value < 0.05) (**Figure 1C**), with a recurrent overexpression of RPs in the TNBC subtype regarding luminal subtypes, indicating that the overexpression of RPs may be specially related to ER/PR negative breast tumors. Only *RPS12* and *RPS13* presented differential expression in both T vs. NT and T vs. T comparisons at proteomic level.

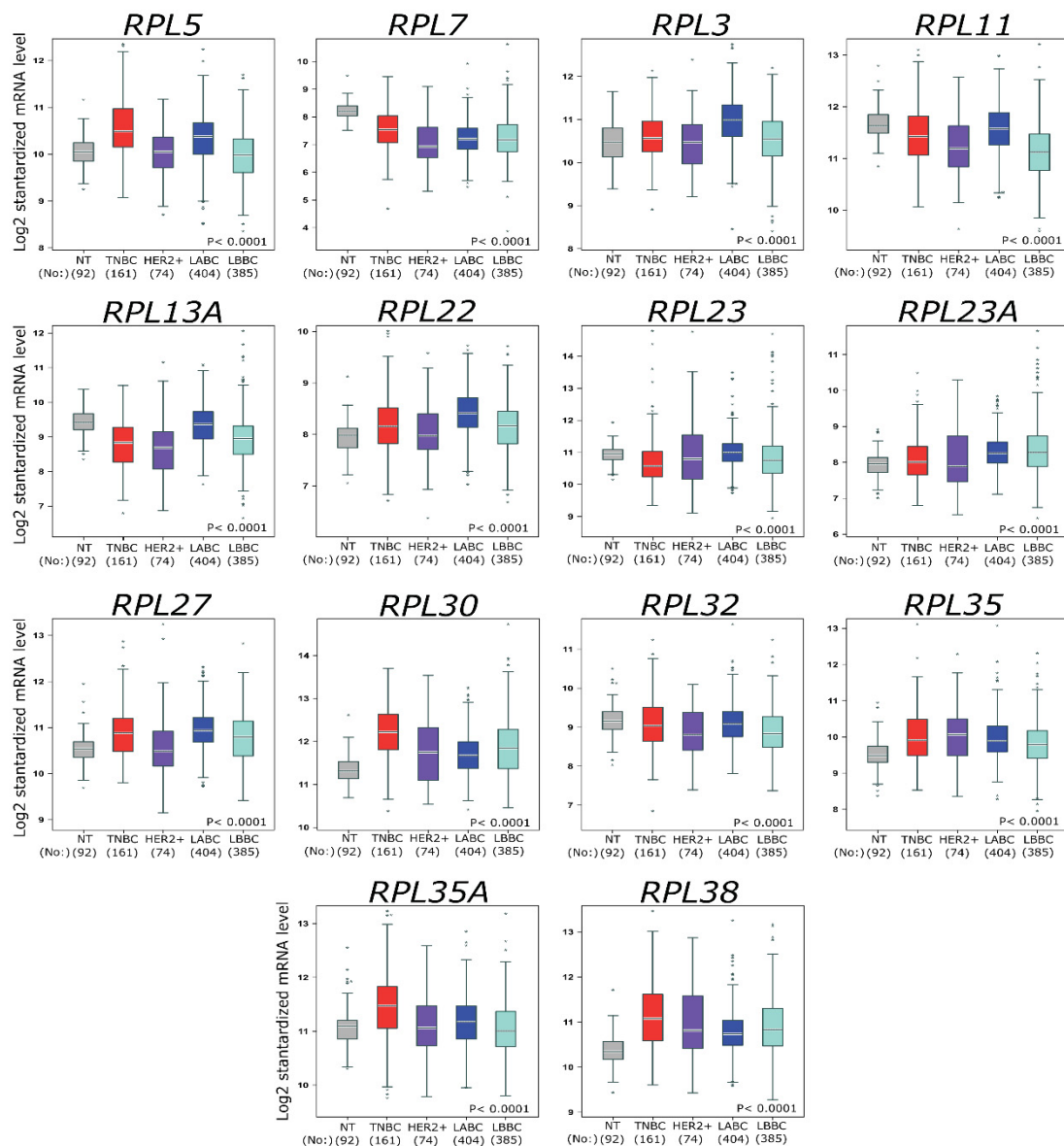
FIGURE 1 - DIFFERENTIALLY EXPRESSED RIBOSOMAL PROTEINS (RPs) AND PROTEOMIC EXPRESSION LEVELS OF SPECIFIC RPs.

FIGURE 2 - DIFFERENTIALLY EXPRESSED SMALL RIBOSOMAL PROTEINS AT mRNA LEVELS.



NT = Non-tumor breast samples. TNBC = Triple negative breast cancer. HER2 = HER2+ enriched breast cancer. LABC = Luminal A breast cancer. LBBC = Luminal B breast cancer

FIGURE 3 - DIFFERENTIALLY EXPRESSED LARGE RIBOSOMAL PROTEINS AT mRNA LEVELS.



NT = Non-tumor breast samples. TNBC = Triple negative breast cancer. HER2 = HER2+ enriched breast cancer. LABC = Luminal A breast cancer. LBBC = Luminal B breast cancer.

A total of 33 RPs showed differences in the T vs NT and/or T vs T comparisons in at least one subtype comparison (P-value < 0.05), with *RPS2* being only differentially expressed in the T vs NT comparison, and *RPS28* only in the T vs T comparison.

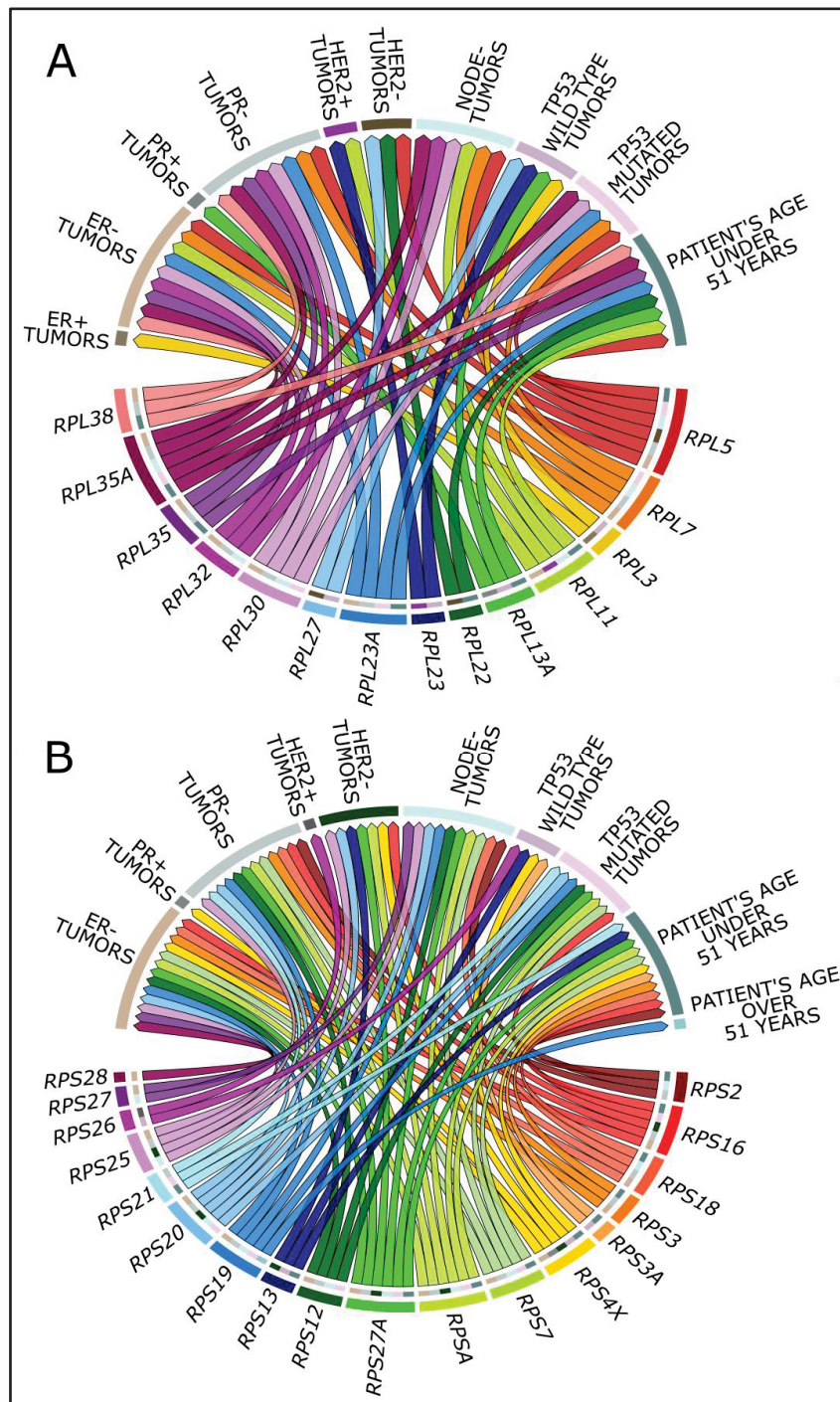
While the T vs NT comparisons of proteomic data presented only overexpressed RPs, the mRNA data included downexpressed RPs, like *RPS27A*, *RPL13A* and *RPL7*. Furthermore, while all RPs presented higher expression in the TNBC subtype regarding luminal subtypes at protein level, the mRNA data

included downexpressed RPs in the TNBC subtype. *RPL3*, *RPS3A* and *RPL13A*, for example, showed higher expression in LABC compared to TNBC tumors. Overall, a fair correspondence between protein and mRNA expression levels were observed, however, there were also inconsistencies between mRNA and protein levels, suggesting a post-transcriptional regulation of RPs expression in breast cancer.

3.2 Association between RPs expression and clinicopathological parameters of breast cancer

Next, we used bc-GenExMiner (v. 4.7) to investigate the relationship between the RPs mRNA levels and classical clinicopathological parameters of breast cancer - ER, PR and *HER2* status (pos/neg), *TP53* status (wild type/mutated), node involvement (pos/neg) and patients age (< 51/ ≥ 51) - besides its association with tumor stage, Nottingham Prognostic Index (NPI) and Scarff-Bloom-Richardson grade (SBR). (**Figures 4A** and **4B**)

FIGURE 4 - ASSOCIATIONS BETWEEN RPs mRNA EXPRESSION LEVELS AND CLINICOPATHOLOGICAL PARAMETERS OF BREAST CANCER.



(A) Significant positive associations (P-value < 0.05) between high mRNA expression of each small ribosomal protein and ER/PR/HER2 status (pos/neg), node commitment (pos/neg), *TP53* mutational status (Wild-type/mutated) and patients age (< 51/ ≥ 51). (B) Significant positive associations (P-value < 0.05) between high mRNA expression of each large ribosomal protein and ER/PR/HER2 status, node commitment (pos/neg), *TP53* mutational status (Wild-type/mutated) and patients age (under 51 years/ over 51 years).

In this analysis, 66.7% of the RPs had its high expression significantly associated with ER negative tumors (P-value < 0.05); 60.61% of them had its high expression significantly associated with PR negative tumors (P-value < 0.05), and 30.3% of the RPs were associated with *HER2* negative status (P-value < 0.05), corroborating the strong association of RPs expression with TNBC tumors. Among all RPs, only *RPL38* was significantly associated with tumor stage (lower expression in advanced tumor stage; P-value = 0.0440).

Regarding age, node involvement and *TP53* status, most of the RPs presented a positive association between its high expression and patients under 51 years (57.6%), node negative tumors (48.48%) and *TP53* mutated tumors (39.4%) (**Figure 4A** and **4B**).

The Nottingham Prognostic Index (NPI) comprehends a well-established classification system (based in tumor size, number of involved lymph nodes and tumor grade) that is commonly used to determine patient's prognosis following surgery of breast tumors (GALEA et al., 1992; FONG et al., 2015). Among the RPs, 17 were significantly associated with NPI, being 13 highly expressed in NPI-1 tumors, and four with high expression associated with NPI-3 tumors (**Table 1**). The Scarff-Bloom-Richardson grade (SBR) is a histologic grading system that evaluates tubule formation, nuclear characteristics of pleomorphism and mitotic index, classically used to predict response to neoadjuvant treatment (AMAT et al., 2002). Twenty-eight RPs were significantly (P-value < 0.05) associated with the SBR grade, with 17 RPs highly expressed in SBR-1 tumors, and 11 RPs highly expressed in SBR-3 tumors (**Table 1**).

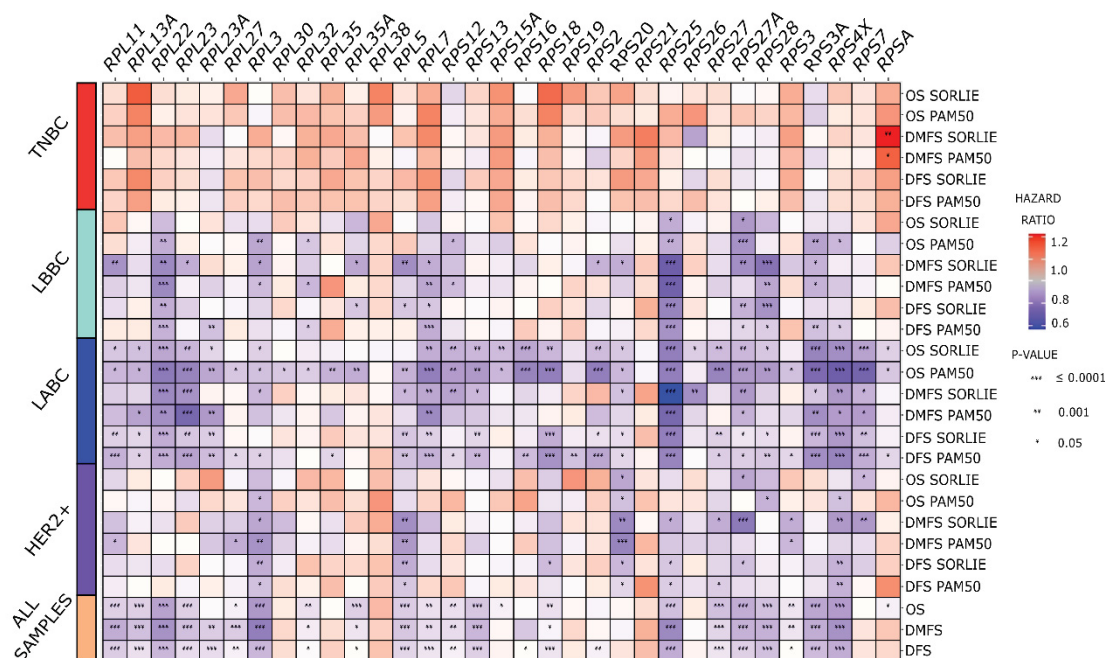
TABLE 1 – ASSOCIATION BETWEEN RPs EXPRESSION AND NOTTINGHAM PROGNOSTIC INDEX AND SCARFF-BLOOM-RICHARDSON GRADES

Ribosomal Proteins	Expression in Scarff-Bloom-Richardson Grades	Expression in Nottingham Prognostic Index	Ribosomal Proteins	Expression in Scarff-Bloom-Richardson Grades	Expression in Nottingham Prognostic Index
<i>RPL3</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2>NPI3 (P < 0.0001)	<i>RPS4X</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2>NPI3 (P = 0.0001)
<i>RPL5</i>	SBR1>SBR2=SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0121)	<i>RPS7</i>	SBR3>SBR2=SBR1 (P = 0.0001)	NPI3=NPI2>NPI1, NPI3>NPI1 (P < 0.0001)
<i>RPL7</i>	SBR3>SBR2=SBR1 (P = 0.0455)	NPI1=NPI2=NPI3 (P = 0.5600)	<i>RPS12</i>	SBR3>SBR2>SBR1, SBR3=SBR1 (P = 0.0072)	NPI1=NPI2=NPI3 (P = 0.7765)
<i>RPL11</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0024)	<i>RPS13</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0001)
<i>RPL13A</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0120)	<i>RPS15A</i>	SBR1>SBR2=SBR3 (P = 0.0103)	NPI1=NPI2=NPI3 (P = 0.4397)
<i>RPL22</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2>NPI3 (P < 0.0001)	<i>RPS16</i>	SBR3>SBR2=BR1 (P = 0.0012)	NPI1=NPI2=NPI3 (P = 0.0968)
<i>RPL23</i>	SBR1=SBR2=SBR3 (P = 0.3577)	NPI1=NPI2=NPI3 (P = 0.1452)	<i>RPS18</i>	SBR1>SBR2=SBR3 (P = 0.0076)	NPI1=NPI2=NPI3 (P = 0.1117)
<i>RPL23A</i>	SBR1=SBR2=SBR3 (P = 0.3423)	NPI1=NPI2=NPI3 (P = 0.0976)	<i>RPS19</i>	SBR3>SBR2>BR1 (P < 0.0001)	NPI3=NPI2>NPI1 (P < 0.0001)
<i>RPL27</i>	SBR1=SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0003)	<i>RPS20</i>	SBR3>SBR2=BR1 (P = 0.0004)	NPI1=NPI2=NPI3 (P = 0.1188)
<i>RPL30</i>	SBR3>SBR2>SBR1 (P = 0.0001)	NPI3>NPI2>NPI1 (P < 0.0001)	<i>RPS21</i>	SBR3>SBR2>BR1 (P < 0.0001)	NPI3>NPI2>NPI1 (P < 0.0001)
<i>RPL32</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0189)	<i>RPS25</i>	SBR1>SBR2=SBR3 (P = 0.0013)	NPI1>NPI2=NPI3, NPI1=NPI3 (P = 0.0129)
<i>RPL35</i>	SBR1=SBR2=SBR3 (P = 0.0524)	NPI1=NPI2=NPI3 (P = 0.1233)	<i>RPS26</i>	SBR1=SBR2=SBR3 (P = 0.7110)	NPI1=NPI2=NPI3 (P = 0.5845)
<i>RPL35A</i>	SBR1=SBR2=SBR3, SBR1>SBR3 (P = 0,0383)	NPI1=NPI2>NPI3 (P = 0.0124)	<i>RPS27</i>	SBR1=SBR2>SBR3 (P = 0.0001)	NPI1=NPI2=NPI3 (P = 0.0682)
<i>RPL38</i>	SBR3>SBR2=SBR1 (P = 0.0017)	NPI1=NPI2=NPI3 (P = 0.8026)	<i>RPS27A</i>	SBR3>SBR2=SBR1, SBR3=SBR1 (P = 0.0230)	NPI1=NPI2=NPI3 (P = 0.1623)
<i>RPS2</i>	SBR1=SBR2=SBR3, SBR1>SBR3 (P = 0,0494)	NPI1=NPI2=NPI3 (P = 0.0743)	<i>RPS28</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P < 0.0001)
<i>RPS3</i>	SBR1=SBR2=SBR3 (P = 0.0574)	NPI1=NPI2=NPI3 (P = 0.2390)	<i>RPSA</i>	SBR3>SBR2=SBR1 (P = 0.0230)	NPI1=NPI2=NPI3 (P = 0.1102)
<i>RPS3A</i>	SBR1>SBR2>SBR3 (P < 0.0001)	NPI1>NPI2=NPI3 (P = 0.0083)			

3.3 Prognostic value of ribosomal proteins in breast cancer

To determine the prognostic value of each differentially expressed RP, we used the univariate Cox proportional Hazard model to correlate the mRNA expression of the RPs with overall survival (OS), distant metastasis-free survival (DMFS) and disease-free survival (DFS). All analyses were performed with and without sample subsection by subtypes, according to Sorlie's and PAM50 subtype classification separately. The results are summarized in **Figure 5** and detailed in the **Supplementary Table 3**.

FIGURE 5 - PROGNOSTIC SIGNIFICANCE OF EACH RP REGARDING OS, DMFS AND DFS IN BREAST CANCER PATIENTS.



'ALL SAMPLES' refers to results obtained without subgrouping samples by subtype. OS = Overall survival. DMFS = Distant metastasis-free survival. DFS = Disease-free survival.

The prognosis analysis revealed that at least 23 out of the total 33 RPs have predictive value to OS, DMFS or DFS, with low expression associated with poor prognosis, while *RPS21*, *RPSA* and *RPL38* had their high expression associated with poor OS, DMFS or DFS, despite not all associations being significant (P -value < 0.05; HR > 1).

When the samples were stratified by subtypes, a slightly different pattern of associations was observed. In non-TNBC tumors, low expression of almost all RPs could be associated with poor prognosis, but in TNBC tumors, poor prognosis was associated with high RPs expression, pointing that each RP present different prognostic values depending on the subtype.

In summary, *RPS27A*, *RPL7*, *RPS27*, *RPS20* and *RPS25* demonstrated potential as candidates to predict OS, DMFS and DFS in LABC, LBBC and HER2+ tumors, (P-value < 0.05; HR < 1), with low expression associated with poor prognosis, but not in TNBC tumors (P-value > 0.05). In the TNBC tumors, only *RPSA* was significantly associated with poor DMFS (P-value < 0.05; HR > 1). *RPS3*, *RPS15A* and *RPS18* also were associated with poor OS, DMFS and DFS in TNBC tumors, although the associations found were not significant (P-value > 0.05; HR > 1).

3.4 Genomic alterations in RPs coding genes associated with breast cancer samples

The aberrant expression can result from gene amplification, deletion or mutations and others. To investigate this context, we used the cBioPortal (v. 3.7.21) database to search for the frequency of genomic alterations in RP genes in both TCGA and METABRIC datasets. As shown in **Figure 6**, the most frequently altered genes were *RPL30* (16%), *RPS27* (16%), *RPL7* (13%), *RPS20* (10%) and *RPL23* (8%), with a predominance of amplification-type and deletion-type alterations, which occur slightly more frequently in ER/PR negative tumors. Moreover, structural variants (*RPL38*, *RPL37*, *RPL23*, *RPL13A*, *RPL11*, *RPS19*, *RPS15A*, *RPS27A*, *RPS4X*, *RPS3*, *RPL5*), inframe mutations (*RPS3*), missense mutations (*RPL35*, *RPL32*, *RPL27*, *RPL23A*, *RPL23*, *RPL13A*, *RPS25*, *RPS20*, *RPS15A*, *RPS12*, *RPS7*, *RPS4X*, *RPS18*, *RPS16*, *RPL7*, *RPS2*, *RPL5*), splice mutations (*RPL11*) and truncating mutations (*RPL22*, *RPS25*, *RPL5*) were reported in this analysis.

FIGURE 6 - cBIOPORTAL ANALYSIS OF MUTATIONAL LANDSCAPE OF THE RPs IN THE TCGA AND METABRIC BREAST TUMOR SAMPLES.

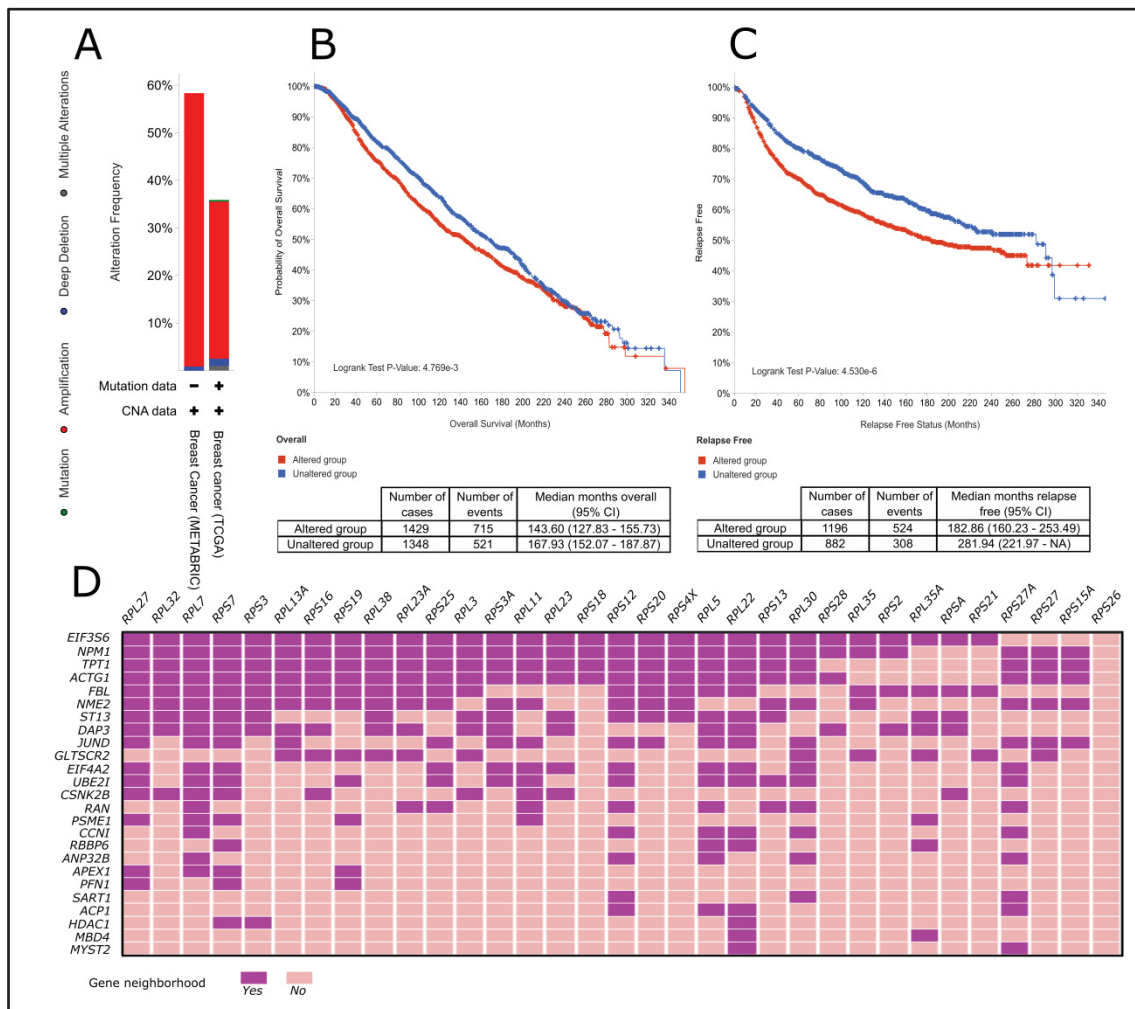


The oncoprint represents the distribution and proportion of samples that carry different types of alterations in RPs genes. Each sample (column) is colored accordingly with the alteration carried for each gene (lines).

Among the METABRIC breast cancer data, almost 60% carry genomic alterations in at least one RP gene, with a smaller percentage in the TCGA samples (~35%) (**Figure 7A**). When subgrouping the tumors of both datasets in “altered” and “unaltered” groups, based in the occurrence of genomic alterations in RPs genes, the Kaplan-Meier analysis demonstrates that patients with “altered” tumors tends to have poor overall survival (**Figure 7B**) and poor relapse-free survival (**Figure 7C**), highlighting the relevance of genomic alterations in RPs genes to breast cancer malignancy and prognostic. Furthermore, we used the “cancer gene neighborhoods” collection of MSigDB to determine genes that may share genomic neighborhood with the RPs genes, and found *EIF3S6*, *NPM1*, *TPT1*, *ACTG1*, *FBL* and other genes related to the RPs genomic regions (**Figure**

7D). As they share genomic neighborhood, these genes may have their expression and function affected by mutations and alterations in the RPs genes.

FIGURE 7 – ALTERATION FREQUENCIES AND CLINICAL RELEVANCE OF MUTATIONS IN THE RPs GENES AND ITS GENOMIC NEIGHBORS.



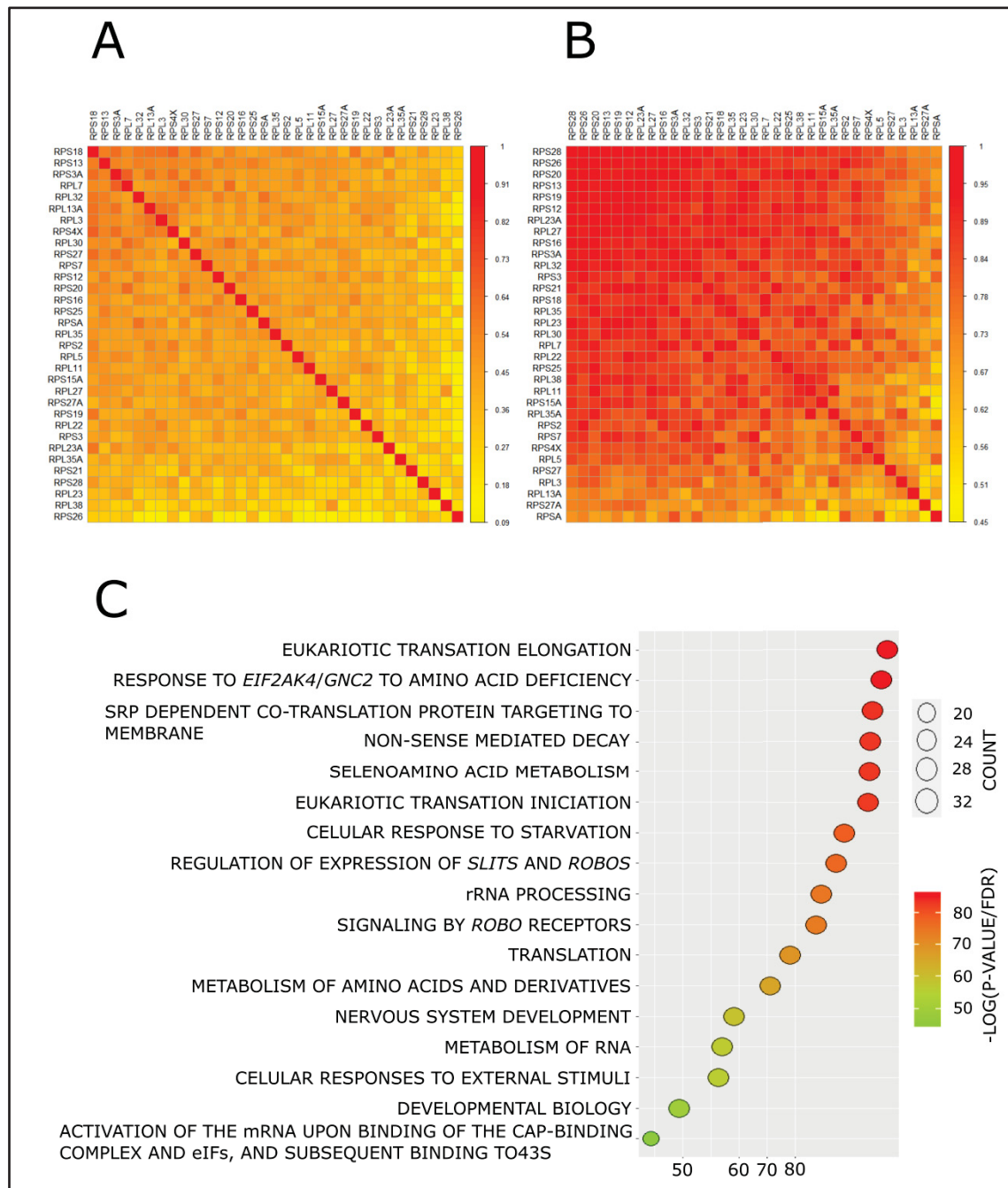
(A) Percentage of samples from METABRIC and TCGA databases that carry at least one RP mutated gene. (B) Kaplan-Meier analysis of overall survival and (C) Relapse-free survival. Patients were divided in 'altered' and 'unaltered' groups based in occurrence of RPs gene mutations, with the 'unaltered' group composed by patients with tumors that do not carry any RP mutation. (D) Heatmap showing the RP genes and its neighbor genes, determined in MSigDB (FDR/P-value < 0.05).

3.5 Correlation between mRNA and protein levels of the RPs and enrichment analysis

Aiming to determine the potential functional mechanisms and pathways in which the RPs are involved, we investigated the correlation between each RP at transcriptomic and proteomic level, and performed an enrichment analysis in the MSigDB platform, searching for the REACTOME pathways and GeneOntology (GO) terms enriched in the set of RPs.

The correlation analysis shows that at mRNA, but mainly at protein levels, the RPs tend to be positively co-expressed in breast tumors, with correlations higher than 0.9 being observed at proteomic level (**Figure 8A** and **8B**). The functional analysis of REACTOME pathways (**Figure 8C**) presented enrichment of classic ribosome related mechanisms, such as translation initiation and elongation, rRNA processing and non-sense mediated decay, but also mechanisms such as metabolism of seleno amino acid, amino acid and RNAs, response to starvation and to external stimuli and regulation of expression of SLITS and ROBOS.

FIGURE 8 - CO-EXPRESSION AND FUNCTIONAL ENRICHMENT ANALYSIS OF THE RPs IN BREAST CANCER



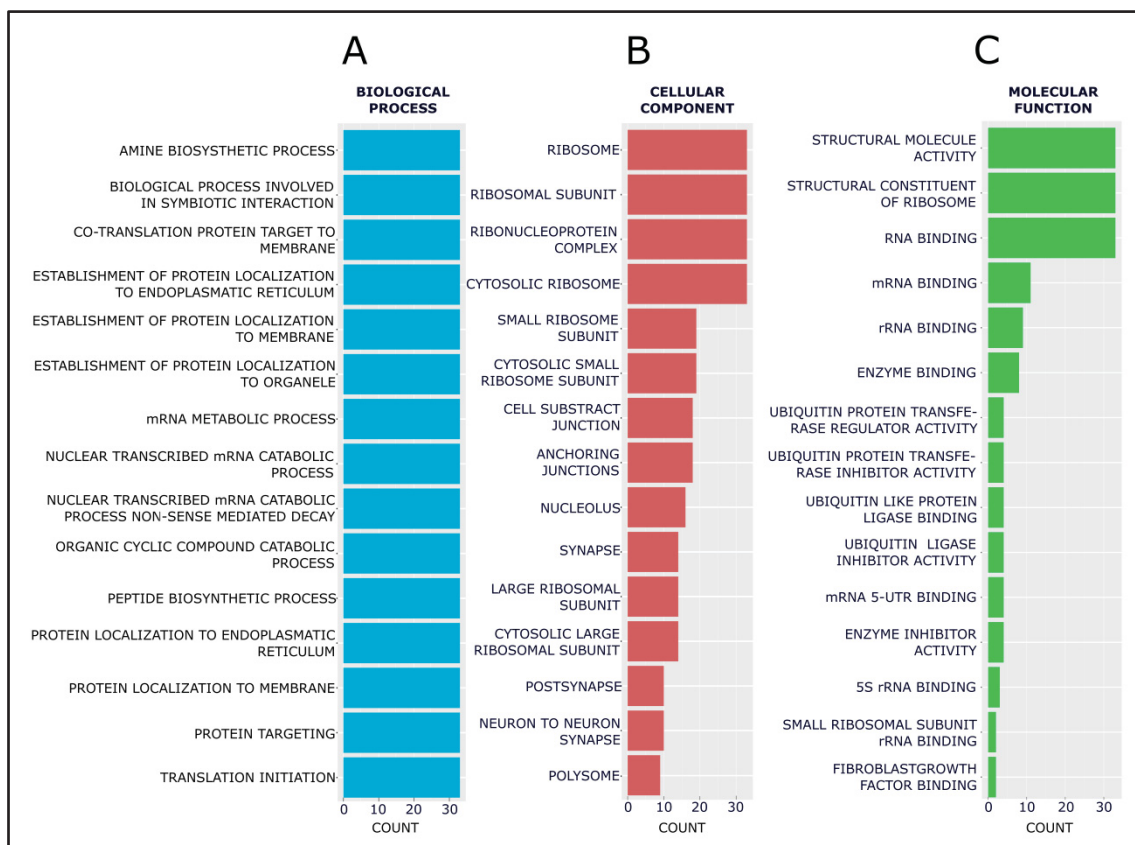
(A) Correlation between RPs expression at mRNA level and (B) protein level. (C) Significantly enriched REACTOME pathways (FDR/P-value < 0.05) in the total set of RPs, according to The Molecular Signatures Database (MSigDB). FDR = False Discovery Rate.

The GO terms analysis predicted the biological role of the investigated RP genes according to three categories: Biological processes, cellular component and molecular function (**Figure 9A-C**). The biological processes significantly enriched to our gene set included: establishment of protein localization and targeting to organelles and membrane, metabolic/catabolic processes such as

nuclear transcribed mRNA catabolic process non-sense mediated decay, amine and peptide biosynthetic process and organic cyclic compound catabolic process.

Regarding cellular components, the location of RBs included, besides the structure of ribosomal complexes and polysomes, cell-substrate junctions, anchoring junctions and in synapse and post-synapse localizations. The molecular functions of the RPs in turn was enriched in its structural and binding functions, activity as inhibitors and regulators of ubiquitin protein transferases and fibroblast growth-factor binding.

FIGURE 9 - GENE ONTOLOGY FUNCTIONAL ENRICHMENT ANALYSIS



(A) Functional enrichment analysis based in Gene ontology terms to (A) Biological processes (B) Cellular component and (C) Molecular function. COUNT = Number of RPs enriched to each GO term.

The REACTOME pathway analysis was also applied to the set of genes that share genomic neighborhood with the RP genes: By sharing genomic neighborhoods, these genes potentially encode interacting or functionally related proteins that may act together with the RP in cancer related pathways or mechanisms. The results are compiled in the **Table 2**.

The main pathways enriched were related to the cell-cycle related pathways, *TCF/WNT* signaling pathways, metabolism of RNA, regulation of *PTEN* and regulation of *RUNX1* and *AP2/TFAP2* transcription factors. In general, we highlight *PSME1*, *RAN*, *UBE2I*, *HDAC1* and *CSNK2B* genes as potential molecular partners of the ribosomal proteins.

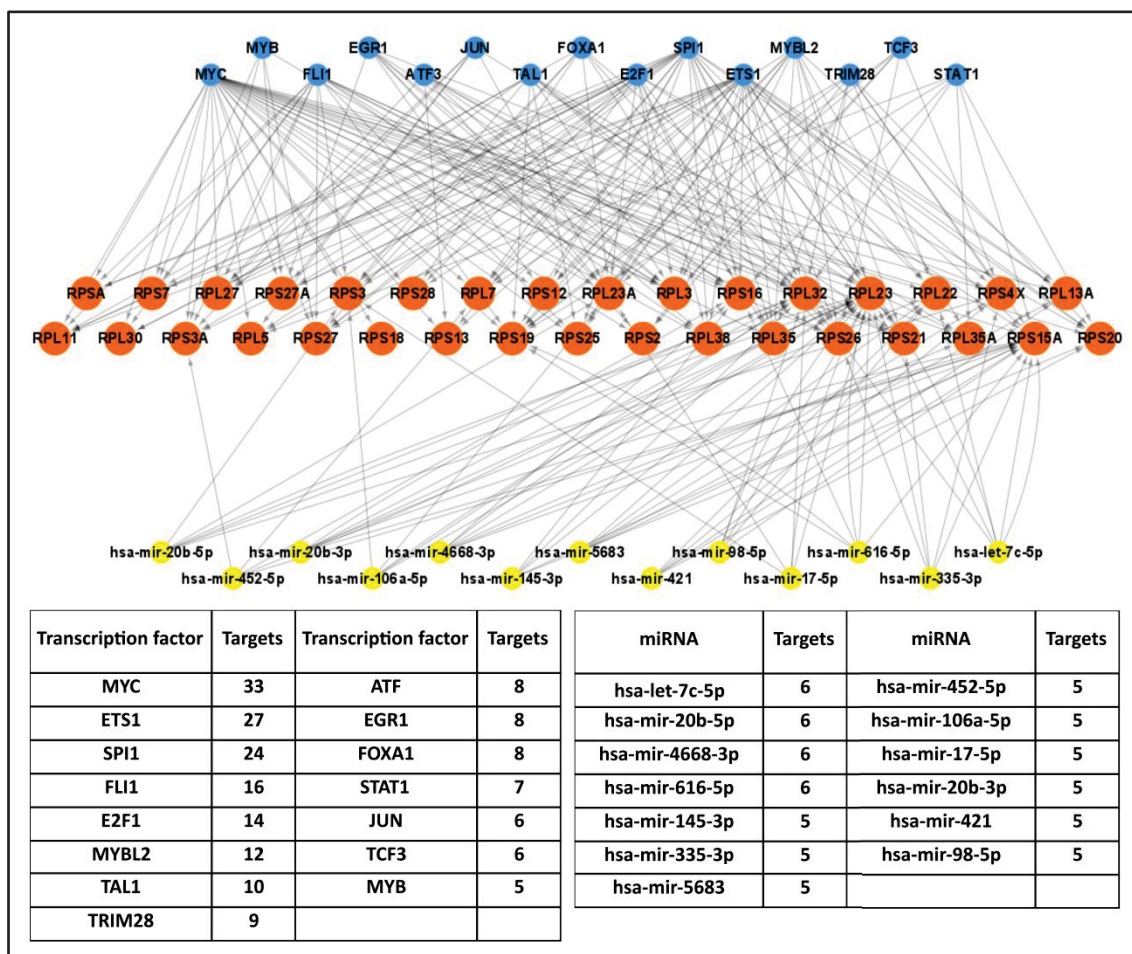
TABLE 2 – REACTOME ENRICHMENT ANALYSIS PERFORMED TO THE SET OF GENES THAT SHARE GENOMIC NEIGHBORHOOD WITH THE RPS GENES

Reactome pathway	Overlapping genes	FDR-value	Reactome pathway	Overlapping genes	FDR-value
CELL CYCLE	PSME1, HDAC1, RAN, UBE2I, NPM1, CSNK2B	5.43E-3	RHOBTB GTPASE CYCLE	ACTG1, RBBP6	2.57E-2
DISPLACEMENT OF DNA GLYCOSYLASE BY APEX1	MBD4, APEX1	8.68E-3	RESOLUTION OF ABASIC SITES/AP SITES	MBD4, APEX1	2.57E-2
CELL CYCLE MITOTIC	PSME1, HDAC1, RAN, UBE2I, CSNK2B	8.68E-3	TRANSCRIPTIONAL REGULATION BY THE AP2/TFAP2 FAMILY OF TRANSCRIPTION FACTORS	UBE2I, NPM1	2.57E-2
SIGNALING BY WNT	PSME1, HDAC1, CSNK2B, PPFN1	1.37E-2	SUMOYLATION OF TRANSCRIPTION COFACTORS	UBE2I, NPM1	3.1E-2
METABOLISM OF RNA	PSME1, RAN, EIF2A2, SART1, FBL	1.37E-2	TCF DEPENDENT SIGNALING IN RESPONSE TO WNT	PSME1, HDAC1, CSNK2B	3.13E-2
PTEN REGULATION	PSME1, HDAC1, CSNK2B	1.75E-2	MITOTIC METAPHASE AND ANAPHASE	PSME1, RAN, UBE2I	3.13E-2
M PHASE	PSME1, RAN, UBE2I, CSNK2B	2.06E-2	NERVOUS SYSTEM DEVELOPMENT	PSME1, CSNK2B, ACTG1, PPFN1	3.13E-2
POSTMITOTIC NUCLEAR PORE COMPLEX REFORMATION	RAN, UBE2I	2.06E-2	TRANSCRIPTIONAL REGULATION BY RUNX1	PSME1, HDAC1, CSNK2B	3.13E-2
NUCLEAR IMPORT OF REV PROTEIN	RAN, NPM1	2.57E-2	ACTIVATION OF THE mRNA UPON BINDING OF THE CAP BINDING COMPLEX AND EIFS AND SUBSEQUENT BINDING TO 43S	EIF2A2, EIF3S6	4.46E-2
SUMOYLATION	HDAC1, UBE2I, NPM1	2.57E-2			

3.6 Prediction of main regulators of RP expression

Lastly, aiming to provide a better understanding of the molecular mechanisms and regulatory circuits in which the RPs are involved, we obtained from the TCGA database a list of the miRNAs and transcription factors (TFs) differentially expressed in breast cancer (data not shown), and performed analysis in the miRGate (ANDRÉS-LEÓN et al., 2015) and HARMONIZE (ROUILLARD et al., 2016) platforms to access validated and predicted miRNA and TF interactions and construct a regulatory network of RPs. In total, 185 miRNAs and 34 TFs were found to regulate the expression of, at least, one RP. The main regulators found (**Figure 10**) include the proto oncogenes *MYC* (33 targets), *ETS1* (27 targets) and *SPI1* (24 targets), and the miRNAs mir-let-7c-5p (6 targets), hsa-mir-20b-5p (6 targets), hsa-mir-4668-3p (6 targets) and hsa-mir-616-5p (6 targets). A more detailed list of regulators-RP pairs is presented in the **Supplementary Table 4**.

FIGURE 10 - GENE REGULATORY NETWORK OF RIBOSOMAL PROTEIN EXPRESSION BY miRNAS AND TRANSCRIPTION FACTORS.



(Legend: The predicted network is composed by RP genes (orange circles), transcription factors (blue circles), miRNAs (yellow circles) and edges that determine the interaction relationship between the TF-RP and miRNA- RP pairs.)

12.4 DISCUSSION

The small and large ribosomal proteins (RPs) are the main components of the ribosomal structure (KANG et al., 2021). It's well known that dysregulation of RPs expression and function, either by oncogenic activation, or changes related to mutations in RP genes, are associated with several cancer types (SAHIN et al., 2005; HUANG et al., 2006; BEE et al., 2011; WU et al., 2011; DUTTON-REGESTER et al., 2014; RAO et al., 2016; HOFMAN et al., 2017). Classically, high RP biogenesis is related to sustained cell proliferation and tumor growth through induction of proteomic expression changes that are beneficial to tumors,

but emerging evidence also suggest that the RPs can exert roles in cancer by ribosomal-independent ways (VENTICINQUE et al., 2011; SULIMA et al., 2018): also, it has been demonstrated that some drugs have the potential to treat cancer patients by reducing the biogenesis of ribosomes (DRYGIN et al., 2009; BRUNO et al., 2017), underscoring the relevance of the RPs in cancer. Nevertheless, the true relevance of each individual RP in breast cancer tumorigenesis is under-documented and poor explored.

Here, multiple bioinformatic approaches and online tools and databases were used to investigate the small and large ribosomal proteins in the breast cancer. As far as we know, this study is the first to comprehensively and systematically explore the expression of the RPs in breast cancer at both proteomic and transcriptomic level, and investigate its potential as biomarkers of diagnostic and prognostic, including also the mutational, functional and regulatory landscape of the RPs in the breast cancer.

Comparing non-tumor and tumor breast cancer tissues of different subtypes (TNBC, HER2+, LABC and LBBC), we found a total of 33 RPs that presented differential expression in at least one comparison at both mRNA and protein levels. In the T vs. NT comparison, 13 RPs were differentially expressed, with *RPS2*, *RPL5*, *RPS16*, *RPS18*, *RPS3*, *RPS4X* and *RPSA* standing out as concordantly overexpressed in tumors at both mRNA and protein levels. At protein level, these differences were significant unregarding tumor subtypes, but at mRNA level, only *RPSA* was overexpressed in tumors of all subtypes. In the T vs. T subtype comparisons, 22 RPs were differentially expressed in at least one subtype-comparison pair. At protein level, all RPs were overexpressed in TNBC tumors regarding LBBC and LABC subtypes, of which five (*RPL23*, *RPL30*, *RPS13*, *RPS19* and *RPS28*) were overexpressed in TNBC regarding all non-TNBC subtypes. The transcriptomic data revealed that almost all RPs are overexpressed in TNBC at mRNA level as well, with *RPS19* and *RPL30* significantly overexpressed in TNBC regarding all non-TNBC tumors. Nevertheless, the mRNA data also included RPs with high expression in luminal tumors (*RPL13A*, *RPL22*, *RPL23*, *RPL23A*, *RPL3*, *RPS3A* and *RPS4X*). Together, our results expand the understanding of individual RP expression in breast cancer and indicate that the RPs have good potential as biomarkers to

differentiate tumor and non-tumor breast tissues and also potential to complement the discrimination of TNBC from non-TNBC tumors.

To achieve a more in-depth understanding of the RP expression in breast cancer, we obtained lists of differentially expressed miRNAs and TFs and predicted the main regulators of RPs expression. Among the TFs, the proto-oncogenes *MYC*, *ETS1* and *SPI1*, together with *FLI1*, *E2F1* and *MYBL2*, presented the most RPs targets. Among the miRNAs, hsa-let-7c-5p, hsa-mir-20b-5p, hsa-mir-4668-3p and hsa-mir-616-5p bounce as the miRNAs with the most RP regulated. Altogether, these results point to new potential regulators of the ribosomal expression in breast cancer, and corroborate previous studies that support the involvement of *MYC*, *FLI1* and *SPI1* in ribosome biogenesis in various cancer types (JUBAN et al., 2009; POPAY et al., 2021) and also the association between hsa-let-7c-5p (NOTTROT et al., 2006) and hsa-mir-20b-5p (CHENG et al., 2021) with ribosomal regulation.

According the clinicopathological features, almost all RPs were associated with ER/PR negative tumors, and to a lesser extent, with HER2 negative tumors, which is consistent with our expression analysis, also corroborating findings that exploits the positive relation between ribosome abundance and ER/PR negative breast tumors (STEGNER et al., 1986; HE et al., 2018). Regarding lymph node commitment, *RPL7*, *RPL11*, *RPL32*, *RPS7*, *RPS25*, *RPS27* and *RPS27A* stand out: all were downexpressed in tumors. Its low expression was related to node positivity, implying a potential role of these RPs as markers of risk of breast cancer metastasis.

To further clarify the potential of the RPs in breast cancer prognostic prediction, we examined the expression of the RPs regarding OS, DMFS, DFS and the NPI and SBR indexes. Either by down- or overexpression, some RPs have been associated with prognostic prediction in cancer. For example, overexpression of *RPL19* in prostate cancer (BEE et al., 2006); overexpression of *RPS11* and *RPS20* in glioblastoma (YONG et al., 2015) or downexpression of *RPL15* in pancreatic carcinoma (YAN et al., 2015) were pointed to predict prognosis in these cancer types. In our report, we demonstrated that several RPs can be associated with prognostic prediction in breast cancer. Interestingly, all 23

significant associations were determined with low mRNA expression leading to poor prognosis (to all samples, without subtype subgrouping). When considering the samples according to the subtypes, these associations of RP low expression and poor prognosis remain almost the same to HER2+, LABC and LBBC tumors, but not to TNBC tumor. In TNBC, high expression of many RPs was associated with poor prognosis, although the association of *RPSA* and poor DMFS being the only significant.

In general, each subtype presented a set of most fit RPs in terms of subtype-specific prognostic prediction, accordingly with the values of Hazard ratio and significant associations with OS, DMFS and DFS. Here, we highlight *RPS20*, *RPL3* and *RPL5* in the HER2+ tumors; *RPS25*, *RPS4X*, *RPS7* and *RPS3A* in the LABC tumors; *RPS27A*, *RPS25* and *RPL22* in the LBBC tumors and *RPSA*, *RPS3A* and *RPL38* in the TNBC tumors, although not all TNBC associations were significant. Regarding the NPI and SBR indexes, the expression at least 50% of all RPs could be associated significantly with both grading systems, suggesting its potential as prognostic markers.

Mutations in RP genes are related to several ribosomopathies, like Dyskeratosis congenita, Diamond-Blackfan anemia, Shwachman-Diamond syndrome (NARLA et al., 2010) and with elevated risk of developing cancer (LAWRENCE et al., 2014; GOUDARZI & LINDSTROM, 2016). In fact, sundry mutation patterns in different RP genes are related to cancer development (DAVE et al., 2014; NIEMINEN et al., 2014; ORSOLIC et al., 2020). Also, several evidences link the expression of RPs, as well mutations in RPs genes, to *TP53* mutational status in cancer (AJORE et al., 2017; ZONNEVILLE et al., 2020). Our analysis corroborates the *TP53*-RPs relationship in breast cancer, pointing that 22 out of 33 RPs had its expression related to the mutational status of *TP53*. Besides, through cBioPortal analysis, we identified that tumors with mutations in at least one RP genes are frequent in breast cancer (METABRIC, 60%; TCGA, ~35%), with high frequency of amplification and deletion-type mutations, but also missense, truncation, splicing and variant mutations included. Interestingly, the most frequently mutated genes (*RPL30*, *RPS27*, *RPL7*, *RPS20* and *RPL23*) were also related to prognostic prediction and node commitment. To confirm the relevancy of RPs gene mutations in breast cancer, we performed Kaplan-Meier

analysis to verify if patients with mutation-carrying tumors presented poor prognosis. Our results acknowledge that mutations in RPs genes tend to correlate with poor overall survival and poor relapse-free survival in breast cancer patients.

Through genomic neighborhood analysis, we report that *EIF3S6*, *NPM1*, *UBE2I*, *PSME1* and *FBL* genes, among others, are candidates to exert oncogenic functions together with the RPs via ribosomal dependent and independent mechanisms. In chronic myelogenous leukemia, for example, nucleophosmin 1 (*NPM1*) is related to be involved in ribosome biogenesis (CHAN et al., 2015) and overexpressed in several cancers, being implicated with mitotic spindle, centromeres, cytoskeleton binding and, consequently, to the cell cycle and proliferation (CELA et al., 2010). Here, *NPM1* was found sharing genomic neighborhood with 29 RPs, being related with cell cycle regulation and sumoylation in our enrichment analysis. Fibrillarin (*FBL*) is related to modulation of *TP53* expression and ribosome biogenesis in breast cancer (SU et al., 2013). We found *FBL* sharing genomic neighborhood with 22 RPs, and involved with RNA metabolism. Also, the Histone Deacetylase 1 gene (*HDAC1*), related to proliferation and invasion in breast cancer cells (TANG et al., 2017), was found to be a genomic neighbor of three RPs genes, and related, via our analysis, to several tumorigenic processes. Thus, through enrichment and functional predictions, we suggest that these neighbor genes are related to ribosomal functions in breast cancer in cell cycle contexts, sumoylation, regulation of *PTEN*, *AP2/TFAP2* and *RUNX1*, and *WNT* signaling pathway. These genes compose a list of further research targets in the context of understanding the ribosomal involvement in cancer. Also, our results highlight that the ribosomal proteins have a role in breast tumorigenesis participating in regulation and signaling by *SLITS* and *ROBOS*, metabolic pathways, regulation of ubiquitination and responses to external and stress stimuli.

12.4 CONCLUSIONS

In this study, bioinformatic analytic approaches were applied to comprehensively analyze proteomic and transcriptomic data of breast cancer tumor and non-tumor tissue samples, and to provide a better understanding of multiple aspects related to the involvement of the small and large ribosomal proteins in breast tumorigenesis. Collectively, our results contribute to expanding the knowledge about RPs expression in breast cancer, and demonstrate that RPs expression patterns may be useful as diagnostics and predictive biomarkers in patients with breast cancer. In addition, we present insights about functional mechanisms, regulatory systems, mutational landscape and gene neighbors of the RPs in breast cancer. Our analysis demonstrated how the RPs appears to have a special relation with TNBC tumors, not only at expression levels, but also at prognostic impact: Additional investigations are needed to clarify why the ribosome proteins present a different “behavior” in triple negative breast cancer. Nevertheless, clinical analyzes and further validation experiments are needed to confirm the clinical relevance of RPs.

12.5 REFERENCES

- AJORE, R.; RAISER, D.; MCCONKEY, M.; JOUD, M.; BOIDOL, B.; MAR, B.; SAKSENA, G.; WEINSTOCK, D, M.; ARMSTRONG, S.; ELLIS, S, R.; ELBERT, B, L.; NILSSON, B. Deletion of ribosomal protein genes is a common vulnerability in human cancer, especially in concert with TP53 mutations. **EMBO Molecular Medicine**, 9(4):498-507, 2017.
- AMAT, S.; PENAULT-LLORCA, F.; CURE, H.; BOUEDEC, G.; ACHARD, J, L.; PRAGH, I, V.; FEILLEL, V.; MUORET-REYNIER, M, A.; DAUPLAT, J.; CHOLLET, P. Scarff-Bloom-Richardson (SBR) grading: a pleiotropic marker of chemosensitivity in invasive ductal breast carcinomas treated by neoadjuvant chemotherapy. **International journal of oncology**, 20:791-796, 2002.
- ANDRÉS-LEÓN, E.; PEÑA, D, G.; GÓMEZ-LÓPEZ, G.; PISANO, D, G. miRGate: a curated database of human, mouse and rat miRNA–mRNA targets. **Database**, 2015.
- BALIC, M.; THOMSEN, C.; WURSTLEIN, R.; GNANT, M.; HARBECK, N. St. Gallen/Vienna 2019: A Brief Summary of the Consensus Discussion on the Optimal Primary Breast Cancer Treatment. **Breast care**, 14:103-110, 2019.
- BEE, A.; BREWER, D.; BEESLEY, C.; DODSON, A.; FOROOTAN, S.; DICKISON, T.; GERARD, P.; LANE, B.; YAO, S.; COOPER, C, S.; DJAMGOZ, M, B, A.; GOSDEN, C, M.; KE, Y.; FOSTER, C, S. siRNA Knockdown of Ribosomal Protein Gene RPL19 Abrogates the Aggressive Phenotype of Human Prostate Cancer. **PlosOne**, 6(7), 2011.
- BEE, A.; KE, Y.; FOROOTAN, S.; LIN, K.; BEESLEY, C.; FORREST, S, E.; FOSTER, C, S. Ribosomal Protein L19 Is a Prognostic Marker for Human Prostate Cancer. **Clinical cancer research**, 12(7), 2006.
- BRUNO, P.; LIU, Y.; PARK, G, Y.; MURAI, J.; KOCH, C, E.; EISEN, T, J.; PRITCHARD, J, R.; POMMIER, Y.; LIPPARD, S, J.; HEMANN, M, T. A subset of platinum-containing chemotherapeutic agents kills cells by inducing ribosome biogenesis stress. **Nature medicine**, 2018.
- CARITHERS, L, J.; ARDLIE, K.; BARCUS, M.; BRANTON, P, A.; BRITTON, A.; BUIA, S, A.; COMPTON, C, C.; DELUCA, D, S.; PETER-DEMCHOK, J.; GELFAND, E, T.; GUAN, P.; KORZENIEWSKI, G, E.; LOCKHART, N, C.; RABINER, C, A.; RAO, A, K.; ROBINSON, K, L.; ROCHE, N, V.; SAWYER, S, J.; SEGRÈ, A, V.; SHIVE, C, E.; SMITH, A, M.; SOBIN, L, H.; UNDALE, A, H.; VALENTINO, K, M.; VAUGHT, J.; YOUNG, T, R.; MOORE, H, M. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. **BIOPRESERVATION AND BIOBANKING**, 13(5), 2015.

CELA, I.; MATTEO, A, D.; FEDERICI, L.; Nucleophosmin in Its Interaction with Ligands. **International Journal of Molecular Science**, 21(4885), 2020.

CERAMI, E.; GAO, J.; DOGRUSOZ, U.; GROSS, B, E.; SUMER, S, O.; AKSOY, B, A.; JACOBSEN, A.; BYRNE, C, J.; HEUER, M, L.; LARSSON, E.; ANTIPIN, Y.; REVA, B.; GOLDBERG, A, P.; SANDER, C.; SCHULTZ, N. The cBio Cancer Genomics Portal: An Open Platform for exploring Multidimensional Cancer Genomics Data. **American association of cancer research**, 2(5):401-404, 2012.

CHAN, L, W, C.; LIN, X.; YUNG, G.; LUI, T.; CHIU, Y, M.; WANG, F.; TSUI, N, B, Y.; CHO, W, C, S.; YIP, S, P.; SIU, P, M.; WONG, S, C, C.; TUNG, B. Novel structural co-expression analysis linking the NPM1- associated ribosomal biogenesis network to chronic myelogenous leukemia. **Scientific reports**, 5(10973), 2015.

CHENG, Z.; DAI, Y.; HUANG, W.; ZHONG Q.; ZHU, P.; ZHANG, W.; WU, Z.; LIN, Q.; ZHU, H.; CUI, L.; QIAN, T.; DENG, C.; FU, L.; LIU, Y.; ZENG, T. Prognostic Value of MicroRNA-20b in Acute Myeloid Leukemia. **Frontiers in oncology**, 10, 2021.

DAVE, B.; GRANADOS-PRINCIPAL, S.; ZHU, R.; BENZ, S.; RABIZADEH, S.; SOON-SHION, P.; YU, K, D.; SHAO, Z.; LI, X.; GILCREASE, M.; LAI, Z.; CHEN, Y.; HUANG, T, H, M.; SHEN, H.; LIU, X.; FERRARI, M.; ZHAN, M.; WONG, S, T, C.; KUMARASWAMI, M.; MITTAL, V.; CHEN, X.; GROSS, S, S.; CHANG, J, C. Targeting RPL39 and MLF2 reduces tumor initiation and metastasis in breast cancer by inhibiting nitric oxide synthase signaling. **PNAS**, 111(24), 8838-8843, 2014.

DONG, Z.; JIANG, H.; LIANG, S.; WANG, Y.; JIANG, W.; ZHU, C. Ribosomal Protein L15 is involved in Colon Carcinogenesis. **International Journal of Medical Sciences**, 16(8): 1132-1141, 2019.

DRYGIN, D.; SIDDIQUI-JAIN, A.; O'BRIEN, S. Anticancer Activity of CX-3543: A Direct Inhibitor of rRNA biogenesis. **Cancer research**, 69:7653-7661, 2009.

DUTTON-REGISTER, K.; GARTNER, J, J.; EMMANUEL, R.; QUTOB, N.; DAVIES, M, A.; GERSHENWALD, J, E.; ROBINSON, W.; ROBINSON, S.; ROSENBERG, S, A.; SCOLYER, R, A.; MANN, G, J.; THOMPSON, J, F.; HAYWARD, N, K.; SAMUELS, Y. A highly recurrent RPS27 5'UTR mutation in melanoma. **Oncotarget**, 5(10):2912-2917, 2014.

EBRIGHT, R, Y.; LEE, S.; WITTNER, B, S. NIEDERHOFFER, K, L.; NICHOLSON, B, T.; BARDIA, A.; TRUESDELL, S.; WILEY, D, F.; WESLEY, B.; LI, S.; MAI, A.; ACETO, N.; VINCENT-JORDAN, N.; SZABOLCS, A.; CHIRN, B.; KREUZER, J.; COMAILLS, V.; KALINNICH, M.; HAAS, W.; TING, D, T.; TONER, M.; VASUDEVAN, S.; HABER, D, A.; MASHESWARAN, S.;

MICALIZZI, D, S. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. **Science**, 367(6485): 1468-1473, 2020.

FONG, Y.; EVANS, J.; BROOK, D.; KENKRE, J.; JARVIS, P.; GOWER-THOMAS, K. The Nottingham Prognostic Index: five- and ten-year data for all-cause survival within a screened population. **Ann R Coll Surg Engl**, 97:137-139, 2015.

FRAGOMENI, S, M.; SCIALLI, A.; JERUSS, J, S. Molecular subtypes and local-regional control of breast cancer. **Surg Oncol Clin N Am**; 27(1): 95–120. 2019.

GALEA, M, H.; BLAMEY, R, W.; ELSON, C, E.; ELLIS, I, O. The Nottingham Prognostic Index in primary breast cancer. **Breast cancer research and treatment**, 22:207-219, 1992.

GAMPLE, P.; JANOENSRI, R.; WANG, H.; TAN, F.; MORAN, M.; BROWN, T.; FLAMENT-AUVIGNE, I.; RAKHA, E, A.; TOSS, M.; DABBS, D, J.; REGITNIG, P.; OLSON, N.; WREN, J, H.; ROBINSOM ,C.; CORRADO, G, S.; PENG, L, H.; LIU, Y.; MERMEL, G, H.; STEINER, D, F.; CHEN, P, H, C. Determining breast cancer biomarker status and associated morphological features using deep learning. **COMMUNICATIONS MEDICINE**, 1(14), 2021.

GOLDHIRSCH, A.; WINER, E, P.; COATES, A, S.; GELBER, R, D.; PICCART-GEBHART, M.; THURLIMANN, B.; SENN, H, J. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. **Annals of Oncology**, 2:2206-2223, 2013.

GOMIG, T, H, B.; CAVALLI, I, J.; SOUZA, L, R, R.; VIEIRA, E.; LUCENA, A, C, R.; BATISTA, M.; MACHADO, K, C.; MARCHINI, F, K.; MARCHI, F, A.; LIMA, R, S.; URBAN, C, A.; CAVALLI, L, R.; RIBEIRO, E, M, S, F. Quantitative label-free mass spectrometry using contralateral and adjacent breast tissues reveal differentially expressed proteins and their predicted impacts on pathways and cellular functions in breast cancer. **Journal of Proteomics**, 199, 1–14, 2019.

GOUDARZI, K, M.; LINDSTROM, M, S. Role of ribosomal protein mutations in tumor development (Review). **International journal of oncology**, 48(1313-1324), 2016.

HARRELL, F, J. With contributions from Dupont C and many others. Hmisc: Harrell Miscellaneous. R package version 3.17–1. CRAN.R-project.org/package=Hmisc. **Accessed 31 dez, 2021.**

HE, Z.; XU, Q.; WANG, X.; WANG, J.; MU, X.; CAI, Y.; QIAN, Y.; SHAO, W.; SHAO, Z. RPLP1 promotes tumor metastasis and is associated with a poor prognosis in triple-negative breast cancer patients. **Cancer cell int.**, 18, 2018.

HOFMAN, I, J, F.; PATCHETT, S.; DUIN, M, V., GEERDENS, E.; VERBEECK, J.; MICHAUX, L.; DELFORGE, M.; SONNEVELD, P.; JOHANSON, A, W.; KEERSMAECKER, K. Low frequency mutations in ribosomal proteins RPL10 and RPL5 in multiple myeloma. **HAEMATOLOGICA**, 102, 2017.

HONG, W, H.; SHABIHKHANI, M.; TELESKA, D.; YANG, S.; TSO, J, L.; MENKIVAR, J, C.; WEI, B.; LUCEY, G, M.; MARENINOV, S.; CHEN, Z.; LIAU, L, M.; LAI, A.; NELSON, S, F.; CLOUGHESY, T, F.; TSO, C, L. Ribosomal Proteins RPS11 and RPS20, Two Stress-Response Markers of Glioblastoma Stem Cells, Are Novel Predictors of Poor Prognosis in Glioblastoma Patients. **PlosOne**, 2015.

HUANG, X, P.; ZHAO, C, X.; LI, Q, J.; CAI, Y.; LIU, F, X.; H, H., XU, X.; HAN, Y, L.; WU, M.; ZHAN, Q, M.; WANG, M, R. Alteration of RPL14 in squamous cell carcinomas and preneoplastic lesions of the esophagus. **Gene**, 366:161-168, 2006.

JÉZÉQUEL, P.; GOURAUD, W.; AZZOUZ, F, B.; GUÉRIN-CHARBONNEL, C.; JUIN, P, P.; LASLA, H.; COMPONE, M. bc-GenExMiner 4.5: new mining module computes breast cancer differential gene expression analyses. **DATABASE**, 00: 1-10, 2021.

JUBAN, G.; GIRAUD, G.; GUYOT, B.; BELIN, S.; DIAZ, J, J.; STARCK, J.; GUILLOUF, C.; MOREAU-GACHELIN, F.; MORLÉ, F. Spi-1 and Fli-1 Directly Activate Common Target Genes Involved in Ribosome Biogenesis in Friend Erythroleukemic Cells. **Molecular and cellular biology**, 29(10):2852-2864, 2009.

KANG, J.; BRAJANOVSKI, N.; CHAN, K, T.; XUAN, J.; PEARSON, R, B.; SANJI, E. Ribosomal proteins and human diseases: molecular mechanisms and targeted therapy. **Signal Transduction and Targeted Therapy**. 6(323), 2021.

LAWRENCE, M, S.; STOJANOV, P.; MERMEL, C, H.; ROBINSON, J, T.; GARRAWAY, L, A.; GOLUB, T, R.; MEYERSON, M.; GABRIEL, S, B.; LANDER, E, S.; GETZ, G. Discovery and saturation analysis of cancer genes across 21 tumour types. **Nature**, 505(495), 2014.

LI, R.; QU, H.; WANG, S.; WEI, J.; ZHANG, L.; MA, R.; LU, J.; ZHU, J.; DHONG, W, D.; KIA, Z. GDCRNATools: an R/Bioconductor package for integrative analysis of lncRNA, miRNA, and mRNA data in GDC. **Bioinformatics**, 2018.

LIN, Z.; PENG, R.; SUN, Y.; ZHANG, L.; ZHANG, Z. Identification of ribosomal protein family in triple-negative breast cancer by bioinformatics analysis. **Bioscience Reports**, 41, 2021.

NARLA, A.; EBERT, B, L. Ribosomopathies: human disorders of ribosome dysfunction. **Blood**, 15(16), 2010.

NIEMINEN, T, T.; O'DONOHUE, M, F.; WU, Y.; LOHI, H.; SCHERER, S, W.; PATERSON, A, D.; ELLONEN, P.; ABDEL-RAHMAN, W, M.; CALO, S.; MECKLIN, J, P.; JARVINEN, H, J.; GLEIZES, P, E.; PELTOMAKI, P. Germline Mutation of RPS20, Encoding a Ribosomal Protein, Causes Predisposition to Hereditary Nonpolyposis Colorectal Carcinoma Without DNA Mismatch Repair Deficiency. **Gastroenterology**, 147:595-598, 2014.

NOTTROT, S.; SIMARD, M, J.; RICHTER, J, D. Human let-7a miRNA blocks protein production on actively translating polyribosomes. **Nature structural & molecular biology**, 13(12), 2006.

ORSOLIC, I.; BURSAC, S.; JURADA, D.; HOFMAN, I, D.; DEMBIC, Z.; BARTEK, J.; MIHALEK, I.; VOLAREVIC, S. Cancer-associated mutations in the ribosomal protein L5 gene dysregulate the HDM2/p53-mediated ribosome biogenesis checkpoint. **Oncogene**, 2020.

PARISE, C, A.; CAGGIANO, V. Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers. **Journal of Cancer Epidemiology**, 2014.

PEREIRA, B.; CHIN, S, F.; RUEDA, O, M.; VOLLAN, H, K, M.; PROVENZANO, E.; BARDWELL, H, A.; PUGH, M.; JONES, L.; RUSSELL, R.; SAMMUT, S, J.; TSUI, D, W, Y.; LIU, B.; DAWSON, S, J.; ABRAHAM, J.; NORTHEN, H.; PEDEN, J, F.; MUKHERJEE, A.; TURASHBILI, G.; GREEM A, R.; MCKINNEY, D, R.; OLOUMI, A.; SHAH, S.; ROSENFELD, N.; MURPHY, L.; BENTLEY, D, R.; ELLIS, I, O.; PURUSHOTHAM, A.; PINDER, S, E.; BORRESEN-DALE, A, L.; EARL, H, M.; PHAROAH, P, D.; ROSS, M, T.; APARICIO, S.; CALDAS, C. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. **Nature communications**, 7, 2016.

POPAY, T, M.; WANG, J.; ADAMS, C, M.; HOWARD, G, C.; CODREANU, S, G.; SHERROD, S, D.; MCLEAN, J, A.; THOMAS, L, R.; LOREY, S, .; MACHIDA, Y, J.; WEISSMILLER, A, M.; EISCHEN, C, M.; LIU, Q.; TANSEY, W, P. MYC regulates ribosome biogenesis and mitochondrial gene expression programs through its interaction with host cell factor-1. **eLife**, 10, 2019.

RAO, S.; CAI, K, Q.; STADANLICK, J, E.; GREENBERG-KUSHNIR, N.; SOLANKI-PATEL, N.; LEE, S, Y.; FAHL, S, P.; TESTA, J, R.; WIEST, D, L. Ribosomal Protein Rpl22 Controls the Dissemination of T-cell Lymphoma. **Cancer research**, 76(11), 2016.

ROUILLARD, A.; GUNDERSEN, G, W.; FERNANDEZ, N.; WANG, Z.; MONTEIRO, C, D.; MCDERMOTT, M, G.; MA1YAN, A. The harmonizome: a

collection of processed datasets gathered to serve and mine knowledge about genes and proteins. **Database**, 2016.

SAHIN, F.; QIU, W.; WILENTZ, R, E.; IACOBUZIO-DONAHUE, C, A.; GROSMARK, A.; SU, G, H. RPL38, FOSL1, and UPP1 Are Predominantly Expressed in the Pancreatic Ductal Epithelium. **Pancreas**, 30(2):158-167, 2005.

SHANNON, P.; MARKIEL, A.; OZIER, O.; BALIGA, N, S.; WANG, J, T.; RAMAGE, D.; AMIN, N, SCHWIKOWSKI, B.; IDEKER, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. **Genome research**, 12:2498-2504, 2003.

SØRLIE, T.; TIBSHIRANI, R.; PARKER, J.; HASTIE, T.; MARRON, J, S.; NOBEL, S.; DENG, S.; JOHNSEN, H.; PESICH, R.; GEISLER, S.; DEMETER, J.; PEROU, C, M.; LØNNING, P, E.; BROWN, P, O.; BØRRESEN-DALE, A, L.; BOTSTEIN, D. Repeated observation of breast tumor subtypes in independent gene expression data sets. **PNAS**, 100(14), 8418-8423, 2003.

STEGNER, H, R.; BRORKEN, W.; BAHNSEN, J. Relation Between Receptor Status and Ultrastructure in Breast Cancer: E+P+ versus E-P-. **Path. Res. Pract.**, 181:377-381, 1986.

SU, H.; XU, T.; GANAPATHY, S.; SHADAFAN, M.; LONG, M.; HUANG, T, H, M.; THOMPSON, I.; YUAN, Z, M. Elevated snoRNA biogenesis is essential in breast cancer. **Oncogene**, 1-11, 2013.

SUBRAMANIAN, A.; TAMAYO, P.; MOOTHA, V, K.; MUKHERJEE, S.; ELBERT, B, J.; GILLETTE, M, A.; PAULOVICH, A.; POMEROY, S., GOLUB, T, R.; LANDER E, S.; MERISOV, J, P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. **PNAS**, 102 (43):15545-155505, 2005.

SULIMA, S, O.; HOFMAN, I, J, H.; KEERSMAECKER, K.; DINMAN, J, D. How ribosomes translate cancer. **Cancer Discovery**, 7(10):1069-1087, 2018.

SUNG, H.; FERLAY, J.; SIEGEL, R, L.; LAVERSANNE, M.; SOERJOMATARAM, I.; JEMAL, A.; BRAY, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. **CA Cancer**, 71:209-249, 2021.

TANG, Z.; DING, S.; HUANG, H.; LUO, P.; QING, B.; ZHANG, S.; TANG, R. HDAC1 triggers the proliferation and migration of breast cancer cells via upregulation of interleukin-8. **Biological Chemistry** 'Just Accepted' paper, 2018.

TWEEDIE, S.; BRASCHI, B.; GRAY, K.; JONES, T, E, M.; SEAL, R, L.; YATES, B.; BRUFORD, E, A. Genenames.org: the HGNC and VGNC resources in 2021. **Nucleic Acids Research**, 49:939-946, 2021.

TYANOVA, S.; COX, J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. **Cancer Systems Biology: Methods and Protocols, Methods in Molecular Biology**, 1711, 2018.

VENTICINQUE, L.; JAMISON, K, .; MERUELO, D. Interactions Between Laminin Receptor and the Cytoskeleton During Translation and Cell Motility. **PlosOne**, 6(1), 2011.

WALLDEN, B.; STORHOFF, J.; NIELSEN, T.; DOWIDAR, N.; SCHAPER, C.; FERREE, S.; LIU, S.; LEUNG, S.; GEISS, G.; SNIDER, J.; VICKERY, T.; DAVIES, S, R.; MARDIS, E, R.; GNANT, M.; SESTAK, I.; ELLIS, M, J.; PEROU,C, M.; BERNARD, P, S.; PARKER, J, S.; Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. **BMC Medical Genomics**; 8:54, 2015.

WANG, W.; NAG, S.; ZHANG, X.; WANG, M, H.; WANG, H.; ZHOU, J.; ZHANG, R. Ribosomal Proteins and Human Diseases: Pathogenesis, Molecular Mechanisms, and Therapeutic Implications. **Medicinal Research Reviews**, 35(2):225-285, 2015.

WANG, Z.; HOU, L.; QI, Z.; SUN, J.; GAO, W.; MENG, J.; WANG, Y.; SUN, H.; GU, H.; XIN, Y.; GUO, Z.; YANG, G. Small Ribosomal Protein Subunit S7 Suppresses Ovarian Tumorigenesis through Regulation of the PI3K/AKT and MAPK Pathways. **PlosOne**, 8(11), 2013.

WARNER, J, R.; MCLNTOSH, K, B. How Common Are Extraribosomal Functions of Ribosomal Proteins? **Molecular cell**, 34, 2009.

WEINSTEIN, J, N.; COLLISSON, E, A.; MILLS, G, B.; SHAW, K, M.; OZENBERGER, B, A.; ELLROTT, K.; SHMULEVICH, I.; SANDER, C.; STUART, J, M.; CANCER GENOME ATLAS RESEARCH NETWORK. The Cancer Genome Atlas Pan-Cancer Analysis Project. **Nature genetics**, 45(10):1113-1120, 2013.

WISNIEWSKI, J, R.; GOUGAZ, F, Z. Fast and sensitive total protein and peptide assays for proteomic analysis. **Anal. Chem**, 87, 4110–4116, 2015.

WU, Q.; GOU, Y.; WANG, Q.; JIN, H.; CUI, L.; ZHANG, Y.; HE, L.; WANG, J.; NIE, Y.; SHI, Y.; FAN, D. Downregulation of RPL6 by siRNA Inhibits Proliferation and Cell Cycle Progression of Human Gastric Cancer Cell Lines. **PlosOne**, 6(10), 2011.

YAN, T, T.; FU, X, L.; LI, J.; BIAN, Y, N.; LIU, D, J.; HUA, R.; REN, L, L.; LI, C, T.; SUN, Y, W.; CHEN, H, Y.; FANG, J, Y.; HONG, J. Downregulation of RPL15

may predict poor survival and associate with tumor progression in pancreatic ductal adenocarcinoma. **Oncotarget**, 6(35):37028-37042, 2015.

YANG, S.; CUI, J.; YANG, Y.; LIU, Z.; YAN, H.; TANG, C.; WANG, H.; QIN, H.; LI, X.; WANG, W.; HUANG, Y.; GAO, H. Over-expressed RPL34 promotes malignant proliferation of non-small cell lung cancer cells. **Gene**, 576:421-428, 2016.

ZHAO, W.; LI, X.; NIAN, W.; WANG, J.; WANG, X.; SUN, L.; ZHU, Y.; TONG, Z. Ribosome Proteins Represented by RPL27A Mark the Development and Metastasis of Triple-Negative Breast Cancer in Mouse and Human. **Frontiers in Cell and Developmental Biology**, 9, 2021.

ZONNEVILLE, J.; WONG, V.; LIMOGÉ, M.; NIKIFOROV, M.; BAKIN, A, V. TAK1 signaling regulates p53 through a mechanism involving ribosomal stress. **Scientific reports**, 10(2517), 2020.

12.6 SUPPLEMENTARY MATERIAL

SUPPLEMENTARY TABLE 1 – PARAMETERS USED TO PERFORM THE LC-MS/MS RUNS

Ion Source
Positive Ion (V): 2300
Ion Transfer Tube Temp (°C): 175
Master Scan (MS1)
Detector Type: Orbitrap
Orbitrap Resolution: 120000
Use Quadrupole Isolation: True
Scan Range (m/z): 300-1500
RF Lens (%): 30
AGC Target: standard
Maximum Injection Time (ms): 50
Microscans: 1
Data Type: Profile
Polarity: Positive
Use EASY-IC™ (Internal calibration): True
Advanced Peak Determination: True
Intensity
Filter Type: Intensity Threshold
Intensity Threshold for MS2: 2.0e4
Charge State
Include charge state(s): 2-7
Include undetermined charge states: False
Include charge states 25 and higher: False
Dynamic Exclusion
Exclude after n times: 1
Exclusion duration (s): 60
Mass Tolerance: ppm
Low: 10
High: 10
Exclude Isotopes: True
Data Dependent
Data Dependent Mode: Cycle Time
Time between Master Scans (sec): 2
ddMS² OT HCD (MS2)
Isolation Mode: Quadrupole
Isolation Window (m/z): 1.6
Isolation Offset: Off
Activation Type: HCD
HCD Collision Energy (%): 30
Detector Type: Orbitrap
Scan Range Mode: Auto: m/z Normal

Orbitrap Resolution: 15000
AGC Target: standard
Maximum Injection Time (ms): 22
Microscans: 1
Data Type: Centroid

SUPPLEMENTARY TABLE 2 - RIBOSOMAL PROTEINS DIFFERENTIALLY EXPRESSED. DIFFERENTIALLY EXPRESSED (A) LARGE RIBOSOMAL PROTEINS AND (B) SMALL RIBOSOMAL PROTEINS AT PROTEOMIC LEVEL IN THE TUMOR VS NON-TUMOR (T VS NT) COMPARISON (P-VALUE < 0.05; LOG2FC \pm 0.58) AND T VS T COMPARISON (P-VALUE < 0.05).

SUPPLEMENTARY TABLE 2A

Ribosomal protein	Significant pairs in T vs NT comparison	Significant pairs in T vs T comparison
<i>RPL11</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL13A</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL22</i>	-	HER2+ > LBBC; ; TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPL23</i>	-	TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPL23A</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL27</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL3</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL30</i>	-	TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPL32</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL35</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL35A</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL38</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPL5</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPL7</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS12</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	TNBC > LABC; TNBC > LBBC
<i>RPS13</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPS15A</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPS16</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS18</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS19</i>	-	TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPS2</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS20</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPS21</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPS25</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPS26</i>	-	TNBC > LABC; TNBC > LBBC
<i>RPS27</i>	-	TNBC > LABC; TNBC > LBBC

<i>RPS27A</i>	TNBC > NT; HER2+ > NT; LABC > NT	-
<i>RPS28</i>	-	TNBC > HER2+; TNBC > LABC; TNBC > LBBC
<i>RPS3</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS3A</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS4X</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPS7</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-
<i>RPSA</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	-

SUPPLEMENTARY TABLE 2B

Ribosomal protein	Significant pairs in T vs NT comparison	Significant pairs in T vs T comparison
<i>RPL11</i>	NT > HER2+; NT > LBBC; NT > TNBC; NT > LABC	TNBC > LBBC; LABC > LBBC; LABC > HER2+; TNBC > HER2+
<i>RPL13A</i>	NT > TNBC; NT > HER2+; NT > LBBC	LABC > TNBC; LABC > HER2+; LABC > LBBC
<i>RPL22</i>	LABC > NT; LBBC > NT; TNBC > NT	LABC > HER2+; LABC > LBBC; LABC > TNBC
<i>RPL23</i>	NT > TNBC	LABC > TNBC; LABC > LBBC
<i>RPL23A</i>	LABC > NT; LBBC > NT	LBBC > TNBC; LABC > TNBC
<i>RPL27</i>	TNBC > NT; LABC > NT; LBBC > NT	LABC > LBBC; LABC > HER2+; TNBC > HER2+
<i>RPL3</i>	LABC > NT	LABC > TNBC; LABC > HER2+; LABC > LBBC
<i>RPL30</i>	TNBC > NT; LABC > NT; LBBC > NT; HER2+ > NT	TNBC > HER2+; TNBC > LABC; TNBC > LBBC; LBBC > LABC
<i>RPL32</i>	NT > LBBC; NT > HER2+	LABC > LBBC; LABC > HER2+; TNBC > LBBC
<i>RPL35</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	LABC > LBBC
<i>RPL35A</i>	TNBC > NT	TNBC > LABC; TNBC > LBBC; TNBC > HER2+; LABC > LBBC

<i>RPL38</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	TNBC > LABC; TNBC > LBBC
<i>RPL5</i>	TNBC > NT; LABC > NT	TNBC > HER2+; TNBC > LBBC; LABC > LBBC; LABC > HER2+; TNBC > LA
<i>RPL7</i>	NT > TNBC; NT > HER2+; NT > LABC; NT > LBBC	TNBC > LABC; TNBC > HER2+; TNBC > LBBC
<i>RPS12</i>	TNBC > NT; NT > LBBC	TNBC > HER2+; TNBC > LABC; TNBC > LBBC; LABC > LBBC
<i>RPS13</i>	TNBC > NT; LABC > NT; LBBC > NT	LABC > HER2+; LABC > LBBC; TNBC > HER2+; LBBC > HER2+
<i>RPS15A</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	LABC > LBBC
<i>RPS16</i>	TNBC > NT	TNBC > LABC; TNBC > LBBC; TNBC > HER2+; LABC > LBBC
<i>RPS18</i>	TNBC > NT; LABC > NT	TNBC > HER2+; TNBC > LABC; TNBC > LBBC; LABC > LBBC; LABC > HER2+;
<i>RPS19</i>	TNBC > NT; LABC > NT	TNBC > HER2+; TNBC > LABC; TNBC > LBBC; LABC > LBBC
<i>RPS2</i>	TNBC > NT; LABC > NT	-
<i>RPS20</i>	NT > LBBC; NT > LABC; TNBC > NT	TNBC > LABC; TNBC > LBBC; TNBC > HER2+
<i>RPS21</i>	TNBC > NT	TNBC > LABC; TNBC > LBBC; TNBC > HER2+
<i>RPS25</i>	NT > HER2+; NT > LBBC; NT > LABC	TNBC > LBBC; LABC > LBBC; TNBC > HER2+; LABC > HER2+
<i>RPS26</i>	NT > TNBC; NT > LABC ; NT > LBBC; NT > HER2+	LBBC > LABC; HER2+ > LABC
<i>RPS27</i>	NT > TNBC; NT > HER2+; NT > LABC; NT > LBBC	TNBC > HER2+; LABC > HER2+; LABC > LBBC; TNBC > LBBC; LBBC > HER2+
<i>RPS27A</i>	NT > HER2+; NT > LABC; NT > LBBC;	TNBC > HER2+; TNBC > LABC; TNBC > LBBC; HER2+ > LBBC; LBC > LBBC
<i>RPS28</i>	-	LABC > LBBC; TNBC > LBBC
<i>RPS3</i>	TNBC > NT; LABC > NT	TNBC > LBBC; LABC > LBBC; TNBC > HER2+;

<i>RPS3A</i>	NT > TNBC; NT > HER2+; NT > LABC; NT > LBBC	LABC > TNBC; LABC > LBBC; LABC > HER2+
<i>RPS4X</i>	TNBC > NT; HER2+ > NT; LABC > NT; LBBC > NT	LABC > TNBC; LABC > HER2+; LABC > LBBC
<i>RPS7</i>	NT > LABC; NT > LBBC	TNBC > LABC; HER2+ > LABC; TNBC > LBBC; HER2+ > LBBC; TNBC > HER2+
<i>RPSA</i>	TNBC > NT	TNBC > LABC; TNBC > LBBC; TNBC > HER2+

SUPPLEMENTARY TABLE 3 – DETAILED RESULTS OF THE KAPLAN-MEIER ANALYSIS OF OVERALL-SURVIVAL, DISTANT METASTASIS-FREE SURVIVAL AND DISEASE-FREE SURVIVAL. (A) SURVIVAL ANALYSIS WITHOUT SAMPLE SUBGROUPING BY SUBTYPES. (B) SURVIVAL ANALYSIS WITH SAMPLE SUBGROUPING BY SORLIE'S SUBTYPE CLASSIFICATION. (C) SURVIVAL ANALYSIS WITH SAMPLE SUBGROUPING BY PAM50 SUBTYPE CLASSIFICATION

SUPPLEMENTARY TABLE 3A

RS gene	Subtype/Survival parameter	P-value	Hazard Ratio	RS gene	Subtype/Survival parameter	P-value	Hazard Ratio
RPL11	ALL-DFS	0,0001	0,83	RPS15A	ALL-OS	0,02	0,9
RPL11	ALL-DMFS	0,0001	0,77	RPS16	ALL-DFS	0,0338	0,93
RPL11	ALL-OS	0,0001	0,84	RPS16	ALL-DMFS	0,0919	0,82
RPL13A	ALL-DFS	0,0001	0,87	RPS16	ALL-OS	0,0558	0,92
RPL13A	ALL-DMFS	0,0001	0,81	RPS18	ALL-DFS	0,0003	0,88
RPL13A	ALL-OS	0,0002	0,84	RPS18	ALL-DMFS	0,0426	0,91
RPL22	ALL-DFS	0,0001	0,79	RPS18	ALL-OS	0,0019	0,87
RPL22	ALL-DMFS	0,0001	0,72	RPS19	ALL-DFS	0,3785	0,97
RPL22	ALL-OS	0,0001	0,75	RPS19	ALL-DMFS	0,853	1,01
RPL23	ALL-DFS	0,0001	0,86	RPS19	ALL-OS	0,2102	0,94
RPL23	ALL-DMFS	0,0001	0,81	RPS2	ALL-DFS	0,0041	0,9
RPL23	ALL-OS	0,0001	0,84	RPS2	ALL-DMFS	0,1859	0,94
RPL23A	ALL-DFS	0,0001	0,86	RPS2	ALL-OS	0,0557	0,91
RPL23A	ALL-DMFS	0,002	0,83	RPS20	ALL-DFS	0,6887	0,98
RPL23A	ALL-OS	0,088	0,92	RPS20	ALL-DMFS	0,0623	0,91
RPL27	ALL-DFS	0,0027	0,9	RPS20	ALL-OS	0,1902	0,94
RPL27	ALL-DMFS	0,0001	0,83	RPS21	ALL-DFS	0,0638	1,07
RPL27	ALL-OS	0,0112	0,9	RPS21	ALL-DMFS	0,0531	1,1
RPL3	ALL-DFS	0,0001	0,79	RPS21	ALL-OS	0,9851	1
RPL3	ALL-DMFS	0,0001	0,68	RPS25	ALL-DFS	0,0001	0,77
RPL3	ALL-OS	0,0001	0,73	RPS25	ALL-DMFS	0,0001	0,7

RPL30	ALL-DFS	0,9697	1	RPS25	ALL-OS	0,0001	0,81
RPL30	ALL-DMFS	0,9153	1,01	RPS26	ALL-DFS	0,2888	0,95
RPL30	ALL-OS	0,577	0,97	RPS26	ALL-DMFS	0,2066	0,92
RPL32	ALL-DFS	0,0192	0,92	RPS26	ALL-OS	0,5063	0,97
RPL32	ALL-DMFS	0,017	0,89	RPS27	ALL-DFS	0,0003	0,88
RPL32	ALL-OS	0,0019	0,87	RPS27	ALL-DMFS	0,0007	0,85
RPL35	ALL-DFS	0,3264	1,03	RPS27	ALL-OS	0,0001	0,81
RPL35	ALL-DMFS	0,9438	1	RPS27A	ALL-DFS	0,0001	0,84
RPL35	ALL-OS	0,127	0,93	RPS27A	ALL-DMFS	0,0001	0,78
RPL35A	ALL-DFS	0,0339	0,93	RPS27A	ALL-OS	0,0001	0,77
RPL35A	ALL-DMFS	0,0106	0,88	RPS28	ALL-DFS	0,0001	0,81
RPL35A	ALL-OS	0,0001	0,84	RPS28	ALL-DMFS	0,0001	0,78
RPL38	ALL-DFS	0,112	1,06	RPS28	ALL-OS	0,0001	0,81
RPL38	ALL-DMFS	0,5231	1,03	RPS3	ALL-DFS	0,05	0,93
RPL38	ALL-OS	0,0809	1,08	RPS3	ALL-DMFS	0,0075	0,88
RPL5	ALL-DFS	0,0001	0,85	RPS3	ALL-OS	0,0064	0,88
RPL5	ALL-DMFS	0,0001	0,82	RPS3A	ALL-DFS	0,0001	0,8
RPL5	ALL-OS	0,0001	0,83	RPS3A	ALL-DMFS	0,0001	0,77
RPL7	ALL-DFS	0,0004	0,88	RPS3A	ALL-OS	0,0001	0,76
RPL7	ALL-DMFS	0,0018	0,86	RPS4X	ALL-DFS	0,0001	0,77
RPL7	ALL-OS	0,0031	0,87	RPS4X	ALL-DMFS	0,0001	0,72
RPS12	ALL-DFS	0,0039	0,9	RPS4X	ALL-OS	0,0001	0,74
RPS12	ALL-DMFS	0,002	0,86	RPS7	ALL-DFS	0,6242	0,98
RPS12	ALL-OS	0,0039	0,87	RPS7	ALL-DMFS	0,6075	0,98
RPS13	ALL-DFS	0,0001	0,84	RPS7	ALL-OS	0,1279	0,93
RPS13	ALL-DMFS	0,0001	0,79	RPSA	ALL-DFS	0,802	1,01
RPS13	ALL-OS	0,0001	0,82	RPSA	ALL-DMFS	0,333	1,05
RPS15A	ALL-DFS	0,1209	0,95	RPSA	ALL-OS	0,0318	0,9
RPS15A	ALL-DMFS	0,236	0,94				

SUPPLEMENTARY TABLE 3B

RS gene	Subtype/Survival parameter	P-value	Hazard Ratio	RS gene	Subtype/Survival parameter	P-value	Hazard Ratio
RPL11	HER2-DFS	0,0987	0,86	RPS15A	LBBC-DFS	0,5534	0,95
RPL11	HER2-DMFS	0,0763	0,81	RPS15A	LBBC-DMFS	0,3423	0,89
RPL11	HER2-OS	0,1742	0,85	RPS15A	LBBC-OS	0,6281	1,06
RPL11	LABC-DFS	0,0063	0,84	RPS15A	TNBC-DFS	0,2029	1,12
RPL11	LABC-DMFS	0,0611	0,83	RPS15A	TNBC-DMFS	0,2184	1,15
RPL11	LABC-OS	0,0205	0,82	RPS15A	TNBC-OS	0,2069	1,17
RPL11	LBBC-DFS	0,114	0,86	RPS16	HER2-DFS	0,242	0,9
RPL11	LBBC-DMFS	0,007	0,72	RPS16	HER2-DMFS	0,3034	0,89
RPL11	LBBC-OS	0,7318	1,05	RPS16	HER2-OS	0,5458	1,07
RPL11	TNBC-DFS	0,5885	1,05	RPS16	LABC-DFS	0,0675	0,89
RPL11	TNBC-DMFS	0,5742	1,07	RPS16	LABC-DMFS	0,0855	0,85
RPL11	TNBC-OS	0,9895	1	RPS16	LABC-OS	0,0009	0,76

RPL13A	HER2-DFS	0,1752	0,88	RPS16	LBBC-DFS	0,4386	0,93
RPL13A	HER2-DMFS	0,3863	0,9	RPS16	LBBC-DMFS	0,1926	0,85
RPL13A	HER2-OS	0,7955	0,97	RPS16	LBBC-OS	0,6538	0,94
RPL13A	LABC-DFS	0,0115	0,85	RPS16	TNBC-DFS	0,7549	0,97
RPL13A	LABC-DMFS	0,1101	0,85	RPS16	TNBC-DMFS	0,7795	0,97
RPL13A	LABC-OS	0,0151	0,81	RPS16	TNBC-OS	0,4726	0,92
RPL13A	LBBC-DFS	0,687	0,96	RPS18	HER2-DFS	0,0171	0,8
RPL13A	LBBC-DMFS	0,2135	0,86	RPS18	HER2-DMFS	0,0766	0,81
RPL13A	LBBC-OS	0,6213	0,94	RPS18	HER2-OS	0,1138	0,83
RPL13A	TNBC-DFS	0,0586	1,19	RPS18	LABC-DFS	0,0001	0,77
RPL13A	TNBC-DMFS	0,2423	1,14	RPS18	LABC-DMFS	0,2442	0,89
RPL13A	TNBC-OS	0,0537	1,28	RPS18	LABC-OS	0,0039	0,79
RPL22	HER2-DFS	0,1112	0,86	RPS18	LBBC-DFS	0,7326	1,03
RPL22	HER2-DMFS	0,2349	0,87	RPS18	LBBC-DMFS	0,7582	0,96
RPL22	HER2-OS	0,5708	0,93	RPS18	LBBC-OS	0,914	1,01
RPL22	LABC-DFS	< 0,0001	0,77	RPS18	TNBC-DFS	0,5971	1,05
RPL22	LABC-DMFS	0,0003	0,7	RPS18	TNBC-DMFS	0,4	1,1
RPL22	LABC-OS	0,001	0,76	RPS18	TNBC-OS	0,071	1,25
RPL22	LBBC-DFS	0,006	0,77	RPS19	HER2-DFS	0,839	0,98
RPL22	LBBC-DMFS	0,0038	0,7	RPS19	HER2-DMFS	0,8995	0,99
RPL22	LBBC-OS	0,086	0,8	RPS19	HER2-OS	0,2292	1,16
RPL22	TNBC-DFS	0,7394	1,03	RPS19	LABC-DFS	0,0742	0,89
RPL22	TNBC-DMFS	0,513	1,08	RPS19	LABC-DMFS	0,8237	1,02
RPL22	TNBC-OS	0,9853	1	RPS19	LABC-OS	0,0831	0,87
RPL23	HER2-DFS	0,8363	1,02	RPS19	LBBC-DFS	0,8794	0,99
RPL23	HER2-DMFS	0,7434	1,04	RPS19	LBBC-DMFS	0,4365	0,91
RPL23	HER2-OS	0,7981	1,03	RPS19	LBBC-OS	0,783	0,96
RPL23	LABC-DFS	0,0028	0,82	RPS19	TNBC-DFS	0,9281	1,01
RPL23	LABC-DMFS	0,001	0,72	RPS19	TNBC-DMFS	0,6689	0,95
RPL23	LABC-OS	0,0042	0,79	RPS19	TNBC-OS	0,2787	1,15
RPL23	LBBC-DFS	0,0733	0,84	RPS2	HER2-DFS	0,7694	1,03
RPL23	LBBC-DMFS	0,0321	0,77	RPS2	HER2-DMFS	0,9571	0,99
RPL23	LBBC-OS	0,6031	0,93	RPS2	HER2-OS	0,4439	1,1
RPL23	TNBC-DFS	0,9992	1	RPS2	LABC-DFS	0,0319	0,87
RPL23	TNBC-DMFS	0,4758	1,09	RPS2	LABC-DMFS	0,4204	1,08
RPL23	TNBC-OS	0,7854	0,97	RPS2	LABC-OS	0,0059	0,8
RPL23A	HER2-DFS	0,375	0,92	RPS2	LBBC-DFS	0,0828	0,85
RPL23A	HER2-DMFS	0,1341	0,84	RPS2	LBBC-DMFS	0,0421	0,78
RPL23A	HER2-OS	0,2514	1,15	RPS2	LBBC-OS	0,465	0,91
RPL23A	LABC-DFS	0,0051	0,83	RPS2	TNBC-DFS	0,8783	0,99
RPL23A	LABC-DMFS	0,6223	0,95	RPS2	TNBC-DMFS	0,4254	0,91
RPL23A	LABC-OS	0,0425	0,85	RPS2	TNBC-OS	0,7081	1,05
RPL23A	LBBC-DFS	0,3956	0,92	RPS20	HER2-DFS	0,0119	0,79
RPL23A	LBBC-DMFS	0,9919	1	RPS20	HER2-DMFS	0,0023	0,7
RPL23A	LBBC-OS	0,9094	0,99	RPS20	HER2-OS	0,0224	0,76
RPL23A	TNBC-DFS	0,12	0,87	RPS20	LABC-DFS	0,022	0,86
RPL23A	TNBC-DMFS	0,181	0,86	RPS20	LABC-DMFS	0,0275	0,81
RPL23A	TNBC-OS	0,7214	0,96	RPS20	LABC-OS	0,0191	0,82
RPL27	HER2-DFS	0,3401	0,91	RPS20	LBBC-DFS	0,5613	0,95
RPL27	HER2-DMFS	0,1089	0,83	RPS20	LBBC-DMFS	0,0311	0,77
RPL27	HER2-OS	0,4115	0,91	RPS20	LBBC-OS	0,5085	0,92
RPL27	LABC-DFS	0,3176	0,94	RPS20	TNBC-DFS	0,116	1,15

RPL27	LABC-DMFS	0,515	0,94	RPS20	TNBC-DMFS	0,2124	1,15
RPL27	LABC-OS	0,2984	0,92	RPS20	TNBC-OS	0,3259	1,13
RPL27	LBBC-DFS	0,3006	0,91	RPS21	HER2-DFS	0,8766	0,99
RPL27	LBBC-DMFS	0,7228	0,96	RPS21	HER2-DMFS	0,6333	0,94
RPL27	LBBC-OS	0,2772	0,87	RPS21	HER2-OS	0,8826	0,98
RPL27	TNBC-DFS	0,5003	1,06	RPS21	LABC-DFS	0,2768	1,07
RPL27	TNBC-DMFS	0,4733	0,92	RPS21	LABC-DMFS	0,226	1,13
RPL27	TNBC-OS	0,3649	1,12	RPS21	LABC-OS	0,2543	0,91
RPL3	HER2-DFS	0,0049	0,77	RPS21	LBBC-DFS	0,3398	1,1
RPL3	HER2-DMFS	0,02	0,76	RPS21	LBBC-DMFS	0,7723	1,04
RPL3	HER2-OS	0,0775	0,8	RPS21	LBBC-OS	0,7232	0,95
RPL3	LABC-DFS	0,1869	0,92	RPS21	TNBC-DFS	0,22	1,12
RPL3	LABC-DMFS	0,0155	0,79	RPS21	TNBC-DMFS	0,0963	1,21
RPL3	LABC-OS	0,0367	0,84	RPS21	TNBC-OS	0,9859	1
RPL3	LBBC-DFS	0,1065	0,85	RPS25	HER2-DFS	0,0352	0,82
RPL3	LBBC-DMFS	0,0186	0,75	RPS25	HER2-DMFS	0,0256	0,77
RPL3	LBBC-OS	0,2492	0,86	RPS25	HER2-OS	0,1974	0,86
RPL3	TNBC-DFS	0,4708	1,07	RPS25	LABC-DFS	< 0,0001	0,68
RPL3	TNBC-DMFS	0,3728	1,11	RPS25	LABC-DMFS	< 0,0001	0,48
RPL3	TNBC-OS	0,5498	0,93	RPS25	LABC-OS	< 0,0001	0,68
RPL30	HER2-DFS	0,1915	0,89	RPS25	LBBC-DFS	0,0001	0,69
RPL30	HER2-DMFS	0,055	0,8	RPS25	LBBC-DMFS	< 0,0001	0,6
RPL30	HER2-OS	0,4076	0,91	RPS25	LBBC-OS	0,0412	0,77
RPL30	LABC-DFS	0,7864	0,98	RPS25	TNBC-DFS	0,6306	0,96
RPL30	LABC-DMFS	0,7293	1,03	RPS25	TNBC-DMFS	0,5114	1,08
RPL30	LABC-OS	0,3613	0,93	RPS25	TNBC-OS	0,6795	0,95
RPL30	LBBC-DFS	0,9101	1,01	RPS26	HER2-DFS	0,8155	0,97
RPL30	LBBC-DMFS	0,6839	0,95	RPS26	HER2-DMFS	0,3517	0,87
RPL30	LBBC-OS	0,756	1,04	RPS26	HER2-OS	0,9875	1
RPL30	TNBC-DFS	0,8931	1,01	RPS26	LABC-DFS	0,1102	0,88
RPL30	TNBC-DMFS	0,564	0,94	RPS26	LABC-DMFS	0,009	0,71
RPL30	TNBC-OS	0,5779	1,07	RPS26	LABC-OS	0,0301	0,83
RPL32	HER2-DFS	0,4989	1,06	RPS26	LBBC-DFS	0,8478	0,98
RPL32	HER2-DMFS	0,5925	0,94	RPS26	LBBC-DMFS	0,8752	1,03
RPL32	HER2-OS	0,5275	1,08	RPS26	LBBC-OS	0,4462	0,9
RPL32	LABC-DFS	0,4086	0,95	RPS26	TNBC-DFS	0,1987	0,85
RPL32	LABC-DMFS	0,5333	0,94	RPS26	TNBC-DMFS	0,0901	0,75
RPL32	LABC-OS	0,1499	0,89	RPS26	TNBC-OS	0,9143	0,99
RPL32	LBBC-DFS	0,2886	0,91	RPS27	HER2-DFS	0,0734	0,85
RPL32	LBBC-DMFS	0,1336	0,83	RPS27	HER2-DMFS	0,0324	0,77
RPL32	LBBC-OS	0,9225	0,99	RPS27	HER2-OS	0,2629	0,87
RPL32	TNBC-DFS	0,7514	1,03	RPS27	LABC-DFS	0,0069	0,84
RPL32	TNBC-DMFS	0,4136	1,1	RPS27	LABC-DMFS	0,1645	0,87
RPL32	TNBC-OS	0,9056	0,99	RPS27	LABC-OS	0,0087	0,8
RPL35	HER2-DFS	0,8352	1,02	RPS27	LBBC-DFS	0,3887	0,92
RPL35	HER2-DMFS	0,2423	0,87	RPS27	LBBC-DMFS	0,2349	0,86
RPL35	HER2-OS	0,4626	1,09	RPS27	LBBC-OS	0,7159	0,95
RPL35	LABC-DFS	0,5187	1,04	RPS27	TNBC-DFS	0,9996	1
RPL35	LABC-DMFS	0,7658	0,97	RPS27	TNBC-DMFS	0,7879	0,97

RPL35	LABC-OS	0,1853	0,9	RPS27	TNBC-OS	0,9994	1
RPL35	LBBC-DFS	0,8357	0,98	RPS27A	HER2-DFS	0,023	0,81
RPL35	LBBC-DMFS	0,5277	0,93	RPS27A	HER2-DMFS	0,0008	0,67
RPL35	LBBC-OS	0,292	0,87	RPS27A	HER2-OS	0,0202	0,76
RPL35	TNBC-DFS	0,2274	1,12	RPS27A	LABC-DFS	0,0453	0,88
RPL35	TNBC-DMFS	0,3062	1,12	RPS27A	LABC-DMFS	0,0017	0,73
RPL35	TNBC-OS	0,4515	1,1	RPS27A	LABC-OS	0,0046	0,79
RPL35A	HER2-DFS	0,9542	1,01	RPS27A	LBBC-DFS	0,0086	0,78
RPL35A	HER2-DMFS	0,7232	1,04	RPS27A	LBBC-DMFS	0,0057	0,71
RPL35A	HER2-OS	0,5669	0,93	RPS27A	LBBC-OS	0,0172	0,73
RPL35A	LABC-DFS	0,7884	0,98	RPS27A	TNBC-DFS	0,1755	0,88
RPL35A	LABC-DMFS	0,8376	0,98	RPS27A	TNBC-DMFS	0,3959	0,91
RPL35A	LABC-OS	0,2977	0,92	RPS27A	TNBC-OS	0,5095	0,92
RPL35A	LBBC-DFS	0,033	0,82	RPS28	HER2-DFS	0,1947	0,89
RPL35A	LBBC-DMFS	0,0265	0,76	RPS28	HER2-DMFS	0,4621	0,92
RPL35A	LBBC-OS	0,0717	0,79	RPS28	HER2-OS	0,0586	0,8
RPL35A	TNBC-DFS	0,524	1,06	RPS28	LABC-DFS	0,0231	0,86
RPL35A	TNBC-DMFS	0,5708	1,07	RPS28	LABC-DMFS	0,0653	0,84
RPL35A	TNBC-OS	0,7327	0,96	RPS28	LABC-OS	0,0198	0,82
RPL38	HER2-DFS	0,52	1,06	RPS28	LBBC-DFS	0,0006	0,72
RPL38	HER2-DMFS	0,3858	1,11	RPS28	LBBC-DMFS	0,0007	0,66
RPL38	HER2-OS	0,5523	1,07	RPS28	LBBC-OS	0,0654	0,79
RPL38	LABC-DFS	0,7652	0,98	RPS28	TNBC-DFS	0,3071	0,91
RPL38	LABC-DMFS	0,557	0,94	RPS28	TNBC-DMFS	0,3037	0,89
RPL38	LABC-OS	0,2729	0,91	RPS28	TNBC-OS	0,6289	0,94
RPL38	LBBC-DFS	0,9126	1,01	RPS3	HER2-DFS	0,2475	0,9
RPL38	LBBC-DMFS	0,8803	1,02	RPS3	HER2-DMFS	0,0259	0,77
RPL38	LBBC-OS	0,3876	1,12	RPS3	HER2-OS	0,6965	1,05
RPL38	TNBC-DFS	0,8087	1,02	RPS3	LABC-DFS	0,3946	0,95
RPL38	TNBC-DMFS	0,4522	0,92	RPS3	LABC-DMFS	0,6518	0,96
RPL38	TNBC-OS	0,143	1,2	RPS3	LABC-OS	0,1572	0,89
RPL5	HER2-DFS	0,0055	0,77	RPS3	LBBC-DFS	0,2983	0,91
RPL5	HER2-DMFS	0,0038	0,71	RPS3	LBBC-DMFS	0,1046	0,82
RPL5	HER2-OS	0,182	0,85	RPS3	LBBC-OS	0,5352	0,92
RPL5	LABC-DFS	0,0027	0,82	RPS3	TNBC-DFS	0,2416	1,11
RPL5	LABC-DMFS	0,0181	0,79	RPS3	TNBC-DMFS	0,2504	1,14
RPL5	LABC-OS	0,2352	0,91	RPS3	TNBC-OS	0,4188	1,11
RPL5	LBBC-DFS	0,0323	0,82	RPS3A	HER2-DFS	0,0887	0,85
RPL5	LBBC-DMFS	0,0061	0,71	RPS3A	HER2-DMFS	0,3668	0,9
RPL5	LBBC-OS	0,4977	0,92	RPS3A	HER2-OS	0,9746	1
RPL5	TNBC-DFS	0,4089	1,08	RPS3A	LABC-DFS	0,0004	0,79
RPL5	TNBC-DMFS	0,5864	1,06	RPS3A	LABC-DMFS	0,0438	0,82
RPL5	TNBC-OS	0,9142	0,99	RPS3A	LABC-OS	< 0,0001	0,69
RPL7	HER2-DFS	0,1182	0,86	RPS3A	LBBC-DFS	0,107	0,86
RPL7	HER2-DMFS	0,0587	0,8	RPS3A	LBBC-DMFS	0,0424	0,78
RPL7	HER2-OS	0,8962	0,98	RPS3A	LBBC-OS	0,3264	0,88
RPL7	LABC-DFS	0,0074	0,84	RPS3A	TNBC-DFS	0,3868	0,92
RPL7	LABC-DMFS	0,0039	0,76	RPS3A	TNBC-DMFS	0,5896	0,94
RPL7	LABC-OS	0,0021	0,77	RPS3A	TNBC-OS	0,2176	0,86
RPL7	LBBC-DFS	0,0395	0,82	RPS4X	HER2-DFS	0,0072	0,78
RPL7	LBBC-DMFS	0,0329	0,77	RPS4X	HER2-DMFS	0,0076	0,73
RPL7	LBBC-OS	0,1257	0,82	RPS4X	HER2-OS	0,2849	0,88

RPL7	TNBC-DFS	0,0922	1,17	RPS4X	LABC-DFS	< 0,0001	0,73
RPL7	TNBC-DMFS	0,1255	1,19	RPS4X	LABC-DMFS	0,0016	0,73
RPL7	TNBC-OS	0,4296	1,11	RPS4X	LABC-OS	< 0,0001	0,71
RPS12	HER2-DFS	0,8492	0,98	RPS4X	LBBC-DFS	0,1424	0,87
RPS12	HER2-DMFS	0,8298	0,97	RPS4X	LBBC-DMFS	0,336	0,89
RPS12	HER2-OS	0,6159	0,94	RPS4X	LBBC-OS	0,2965	0,87
RPS12	LABC-DFS	0,0903	0,9	RPS4X	TNBC-DFS	0,9713	1
RPS12	LABC-DMFS	0,0027	0,75	RPS4X	TNBC-DMFS	0,8391	1,02
RPS12	LABC-OS	0,0027	0,78	RPS4X	TNBC-OS	0,7063	1,05
RPS12	LBBC-DFS	0,5065	0,94	RPS7	HER2-DFS	0,051	0,83
RPS12	LBBC-DMFS	0,0916	0,81	RPS7	HER2-DMFS	0,0086	0,73
RPS12	LBBC-OS	0,6529	0,94	RPS7	HER2-OS	0,0483	0,79
RPS12	TNBC-DFS	0,0854	0,85	RPS7	LABC-DFS	0,0015	0,81
RPS12	TNBC-DMFS	0,6205	0,94	RPS7	LABC-DMFS	0,0135	0,78
RPS12	TNBC-OS	0,182	0,85	RPS7	LABC-OS	0,0006	0,75
RPS13	HER2-DFS	0,3734	0,92	RPS7	LBBC-DFS	0,7815	0,97
RPS13	HER2-DMFS	0,3665	0,9	RPS7	LBBC-DMFS	0,3603	0,89
RPS13	HER2-OS	0,4255	0,91	RPS7	LBBC-OS	0,8558	0,98
RPS13	LABC-DFS	0,0093	0,85	RPS7	TNBC-DFS	0,8385	1,02
RPS13	LABC-DMFS	0,0133	0,79	RPS7	TNBC-DMFS	0,8874	0,98
RPS13	LABC-OS	0,0048	0,79	RPS7	TNBC-OS	0,8615	0,98
RPS13	LBBC-DFS	0,3597	0,92	RPSA	HER2-DFS	0,782	1,03
RPS13	LBBC-DMFS	0,6477	0,95	RPSA	HER2-DMFS	0,7093	0,96
RPS13	LBBC-OS	0,4957	0,92	RPSA	HER2-OS	0,6998	0,95
RPS13	TNBC-DFS	0,6859	1,04	RPSA	LABC-DFS	0,1237	0,9
RPS13	TNBC-DMFS	0,8514	0,98	RPSA	LABC-DMFS	0,3202	0,91
RPS13	TNBC-OS	0,8722	1,02	RPSA	LABC-OS	0,0403	0,84
RPS15A	HER2-DFS	0,3032	0,91	RPSA	LBBC-DFS	0,4729	1,07
RPS15A	HER2-DMFS	0,4721	0,92	RPSA	LBBC-DMFS	0,6834	1,05
RPS15A	HER2-OS	0,2442	0,87	RPSA	LBBC-OS	0,4102	1,12
RPS15A	LABC-DFS	0,5152	0,96	RPSA	TNBC-DFS	0,1473	1,14
RPS15A	LABC-DMFS	0,3119	0,91	RPSA	TNBC-DMFS	0,0062	1,37
RPS15A	LABC-OS	0,0099	0,8	RPSA	TNBC-OS	0,4197	1,11

SUPPLEMENTARY TABLE 3C

RS gene	Subtype/Survival parameter	P-value	Hazard Ratio	RS gene	Subtype/Survival parameter	P-value	Hazard Ratio
RPL11	HER2-DFS	0,1567	0,89	RPS15A	LBBC-DFS	0,18	0,91
RPL11	HER2-DMFS	0,0375	0,79	RPS15A	LBBC-DMFS	0,1527	0,87
RPL11	HER2-OS	0,5656	0,94	RPS15A	LBBC-OS	0,0543	0,83
RPL11	LABC-DFS	0,001	0,78	RPS15A	TNBC-DFS	0,1688	1,12
RPL11	LABC-DMFS	0,0586	0,8	RPS15A	TNBC-DMFS	0,1802	1,15
RPL11	LABC-OS	0,0139	0,79	RPS15A	TNBC-OS	0,3208	1,11
RPL11	LBBC-DFS	0,6317	0,97	RPS16	HER2-DFS	0,8136	0,98
RPL11	LBBC-DMFS	0,0841	0,84	RPS16	HER2-DMFS	0,7346	1,04
RPL11	LBBC-OS	0,9781	1	RPS16	HER2-OS	0,2376	1,13
RPL11	TNBC-DFS	0,7825	1,02	RPS16	LABC-DFS	0,0025	0,8
RPL11	TNBC-DMFS	0,4504	0,93	RPS16	LABC-DMFS	0,2846	0,88

RPL11	TNBC-OS	0,7533	1,03	RPS16	LABC-OS	0,0001	0,69
RPL13A	HER2-DFS	0,4316	0,93	RPS16	LBBC-DFS	0,6283	1,04
RPL13A	HER2-DMFS	0,8591	0,98	RPS16	LBBC-DMFS	0,5773	0,95
RPL13A	HER2-OS	0,4012	0,91	RPS16	LBBC-OS	0,9373	0,99
RPL13A	LABC-DFS	0,0193	0,84	RPS16	TNBC-DFS	0,5114	0,95
RPL13A	LABC-DMFS	0,0176	0,75	RPS16	TNBC-DMFS	0,3389	0,91
RPL13A	LABC-OS	0,0147	0,78	RPS16	TNBC-OS	0,9773	1
RPL13A	LBBC-DFS	0,649	0,97	RPS18	HER2-DFS	0,1219	0,88
RPL13A	LBBC-DMFS	0,2128	0,88	RPS18	HER2-DMFS	0,063	0,81
RPL13A	LBBC-OS	0,2259	0,89	RPS18	HER2-OS	0,2518	0,88
RPL13A	TNBC-DFS	0,1123	1,14	RPS18	LABC-DFS	< 0,0001	0,72
RPL13A	TNBC-DMFS	0,5036	1,07	RPS18	LABC-DMFS	0,6493	0,95
RPL13A	TNBC-OS	0,0947	1,2	RPS18	LABC-OS	0,0001	0,68
RPL22	HER2-DFS	0,7238	0,97	RPS18	LBBC-DFS	0,5243	0,95
RPL22	HER2-DMFS	0,5495	0,93	RPS18	LBBC-DMFS	0,4616	0,93
RPL22	HER2-OS	0,6474	0,95	RPS18	LBBC-OS	0,3795	0,92
RPL22	LABC-DFS	0,0002	0,77	RPS18	TNBC-DFS	0,4175	1,07
RPL22	LABC-DMFS	0,0042	0,71	RPS18	TNBC-DMFS	0,5077	1,07
RPL22	LABC-OS	< 0,0001	0,67	RPS18	TNBC-OS	0,0882	1,2
RPL22	LBBC-DFS	0,0005	0,78	RPS19	HER2-DFS	0,3434	0,92
RPL22	LBBC-DMFS	0,0003	0,7	RPS19	HER2-DMFS	0,3555	0,9
RPL22	LBBC-OS	0,0047	0,77	RPS19	HER2-OS	0,2525	0,89
RPL22	TNBC-DFS	0,659	0,97	RPS19	LABC-DFS	0,0034	0,81
RPL22	TNBC-DMFS	0,9158	1,01	RPS19	LABC-DMFS	0,8013	0,97
RPL22	TNBC-OS	0,6621	0,96	RPS19	LABC-OS	0,1062	0,86
RPL23	HER2-DFS	0,2414	0,91	RPS19	LBBC-DFS	0,5992	1,04
RPL23	HER2-DMFS	0,4843	0,93	RPS19	LBBC-DMFS	0,8204	1,02
RPL23	HER2-OS	0,1565	0,86	RPS19	LBBC-OS	0,572	0,95
RPL23	LABC-DFS	0,0001	0,75	RPS19	TNBC-DFS	0,8997	1,01
RPL23	LABC-DMFS	0,0002	0,63	RPS19	TNBC-DMFS	0,6172	0,95
RPL23	LABC-OS	0,0001	0,69	RPS19	TNBC-OS	0,9561	1,01
RPL23	LBBC-DFS	0,2134	0,91	RPS2	HER2-DFS	0,691	0,97
RPL23	LBBC-DMFS	0,2864	0,9	RPS2	HER2-DMFS	0,8257	0,98
RPL23	LBBC-OS	0,6865	0,96	RPS2	HER2-OS	0,7911	1,03
RPL23	TNBC-DFS	0,6149	0,96	RPS2	LABC-DFS	0,0003	0,76
RPL23	TNBC-DMFS	0,8651	0,98	RPS2	LABC-DMFS	0,0596	0,8
RPL23	TNBC-OS	0,9226	0,99	RPS2	LABC-OS	0,0001	0,68
RPL23A	HER2-DFS	0,6274	0,96	RPS2	LBBC-DFS	0,1812	0,91
RPL23A	HER2-DMFS	0,0735	0,82	RPS2	LBBC-DMFS	0,8808	0,99
RPL23A	HER2-OS	0,9412	1,01	RPS2	LBBC-OS	0,4451	0,93
RPL23A	LABC-DFS	0,0035	0,8	RPS2	TNBC-DFS	0,2956	0,92
RPL23A	LABC-DMFS	0,007	0,73	RPS2	TNBC-DMFS	0,086	0,84
RPL23A	LABC-OS	0,0046	0,76	RPS2	TNBC-OS	0,7025	1,04
RPL23A	LBBC-DFS	0,0058	0,81	RPS20	HER2-DFS	0,0489	0,85
RPL23A	LBBC-DMFS	0,088	0,84	RPS20	HER2-DMFS	0,0008	0,69
RPL23A	LBBC-OS	0,3988	0,92	RPS20	HER2-OS	0,0438	0,81
RPL23A	TNBC-DFS	0,1131	0,88	RPS20	LABC-DFS	0,0255	0,85
RPL23A	TNBC-DMFS	0,2773	0,89	RPS20	LABC-DMFS	0,1615	0,85
RPL23A	TNBC-OS	0,9041	0,99	RPS20	LABC-OS	0,0168	0,79
RPL27	HER2-DFS	0,1752	0,9	RPS20	LBBC-DFS	0,1678	0,9
RPL27	HER2-DMFS	0,0174	0,77	RPS20	LBBC-DMFS	0,0819	0,84

RPL27	HER2-OS	0,8589	0,98	RPS20	LBBC-OS	0,1648	0,87
RPL27	LABC-DFS	0,0199	0,85	RPS20	TNBC-DFS	0,455	1,06
RPL27	LABC-DMFS	0,6721	0,95	RPS20	TNBC-DMFS	0,8478	1,02
RPL27	LABC-OS	0,0246	0,82	RPS20	TNBC-OS	0,9938	1
RPL27	LBBC-DFS	0,6527	0,97	RPS21	HER2-DFS	0,0654	1,17
RPL27	LBBC-DMFS	0,3025	0,9	RPS21	HER2-DMFS	0,4427	1,09
RPL27	LBBC-OS	0,5009	0,94	RPS21	HER2-OS	0,576	1,06
RPL27	TNBC-DFS	0,7362	1,03	RPS21	LABC-DFS	0,4855	0,95
RPL27	TNBC-DMFS	0,8601	0,98	RPS21	LABC-DMFS	0,7796	0,97
RPL27	TNBC-OS	0,6129	1,05	RPS21	LABC-OS	0,3136	0,91
RPL3	HER2-DFS	0,0114	0,81	RPS21	LBBC-DFS	0,8559	1,01
RPL3	HER2-DMFS	0,0034	0,72	RPS21	LBBC-DMFS	0,9696	1
RPL3	HER2-OS	0,0319	0,79	RPS21	LBBC-OS	0,701	0,96
RPL3	LABC-DFS	0,0196	0,84	RPS21	TNBC-DFS	0,4222	1,07
RPL3	LABC-DMFS	0,0655	0,81	RPS21	TNBC-DMFS	0,1664	1,15
RPL3	LABC-OS	0,0292	0,81	RPS21	TNBC-OS	0,7689	0,97
RPL3	LBBC-DFS	0,0568	0,87	RPS25	HER2-DFS	0,0265	0,83
RPL3	LBBC-DMFS	0,0246	0,8	RPS25	HER2-DMFS	0,0688	0,82
RPL3	LBBC-OS	0,0079	0,76	RPS25	HER2-OS	0,8254	0,98
RPL3	TNBC-DFS	0,8686	1,01	RPS25	LABC-DFS	< 0,0001	0,69
RPL3	TNBC-DMFS	0,9377	1,01	RPS25	LABC-DMFS	0,0001	0,63
RPL3	TNBC-OS	0,3957	0,91	RPS25	LABC-OS	0,0001	0,68
RPL30	HER2-DFS	0,9924	1	RPS25	LBBC-DFS	< 0,0001	0,72
RPL30	HER2-DMFS	0,1107	0,84	RPS25	LBBC-DMFS	< 0,0001	0,61
RPL30	HER2-OS	0,8031	1,03	RPS25	LBBC-OS	0,0086	0,77
RPL30	LABC-DFS	0,0653	0,87	RPS25	TNBC-DFS	0,9182	1,01
RPL30	LABC-DMFS	0,2815	0,88	RPS25	TNBC-DMFS	0,844	0,98
RPL30	LABC-OS	0,0338	0,82	RPS25	TNBC-OS	0,3229	1,11
RPL30	LBBC-DFS	0,1967	0,91	RPS26	HER2-DFS	0,1896	0,87
RPL30	LBBC-DMFS	0,7547	0,97	RPS26	HER2-DMFS	0,2315	0,84
RPL30	LBBC-OS	0,556	0,94	RPS26	HER2-OS	0,8516	0,98
RPL30	TNBC-DFS	0,4055	1,07	RPS26	LABC-DFS	0,3322	0,92
RPL30	TNBC-DMFS	0,7948	1,03	RPS26	LABC-DMFS	0,2944	0,85
RPL30	TNBC-OS	0,5171	1,07	RPS26	LABC-OS	0,248	0,89
RPL32	HER2-DFS	0,6321	0,96	RPS26	LBBC-DFS	0,5116	0,94
RPL32	HER2-DMFS	0,341	0,9	RPS26	LBBC-DMFS	0,6733	0,95
RPL32	HER2-OS	0,9312	0,99	RPS26	LBBC-OS	0,8847	0,98
RPL32	LABC-DFS	0,3594	0,93	RPS26	TNBC-DFS	0,4982	1,07
RPL32	LABC-DMFS	0,6333	0,95	RPS26	TNBC-DMFS	0,6119	0,93
RPL32	LABC-OS	0,0468	0,83	RPS26	TNBC-OS	0,2016	1,16
RPL32	LBBC-DFS	0,0107	0,83	RPS27	HER2-DFS	0,0316	0,83
RPL32	LBBC-DMFS	0,016	0,79	RPS27	HER2-DMFS	0,0904	0,83
RPL32	LBBC-OS	0,0375	0,82	RPS27	HER2-OS	0,12	0,85
RPL32	TNBC-DFS	0,5644	1,05	RPS27	LABC-DFS	0,0149	0,83
RPL32	TNBC-DMFS	0,3593	1,1	RPS27	LABC-DMFS	0,665	0,95
RPL32	TNBC-OS	0,4718	1,08	RPS27	LABC-OS	0,0002	0,69
RPL35	HER2-DFS	0,4332	1,07	RPS27	LBBC-DFS	0,1309	0,89
RPL35	HER2-DMFS	0,1843	0,86	RPS27	LBBC-DMFS	0,0748	0,84
RPL35	HER2-OS	0,5417	1,07	RPS27	LBBC-OS	0,1406	0,86
RPL35	LABC-DFS	0,028	0,85	RPS27	TNBC-DFS	0,9055	0,99

RPL35	LABC-DMFS	0,0969	0,82	RPS27	TNBC-DMFS	0,4101	0,92
RPL35	LABC-OS	0,0081	0,78	RPS27	TNBC-OS	0,9143	0,99
RPL35	LBBC-DFS	0,1463	1,11	RPS27A	HER2-DFS	0,0671	0,86
RPL35	LBBC-DMFS	0,1398	1,16	RPS27A	HER2-DMFS	0,1721	0,86
RPL35	LBBC-OS	0,8908	0,99	RPS27A	HER2-OS	0,5099	0,93
RPL35	TNBC-DFS	0,4363	1,06	RPS27A	LABC-DFS	0,014	0,83
RPL35	TNBC-DMFS	0,729	1,04	RPS27A	LABC-DMFS	0,035	0,78
RPL35	TNBC-OS	0,5756	1,06	RPS27A	LABC-OS	0,0009	0,73
RPL35A	HER2-DFS	0,3639	0,93	RPS27A	LBBC-DFS	0,0338	0,85
RPL35A	HER2-DMFS	0,6149	0,95	RPS27A	LBBC-DMFS	0,1229	0,86
RPL35A	HER2-OS	0,9716	1	RPS27A	LBBC-OS	0,0006	0,72
RPL35A	LABC-DFS	0,4089	0,94	RPS27A	TNBC-DFS	0,3564	0,93
RPL35A	LABC-DMFS	0,6506	0,95	RPS27A	TNBC-DMFS	0,1835	0,87
RPL35A	LABC-OS	0,005	0,77	RPS27A	TNBC-OS	0,6676	1,05
RPL35A	LBBC-DFS	0,5809	0,96	RPS28	HER2-DFS	0,1369	0,88
RPL35A	LBBC-DMFS	0,7791	0,97	RPS28	HER2-DMFS	0,3729	0,91
RPL35A	LBBC-OS	0,0903	0,85	RPS28	HER2-OS	0,0219	0,78
RPL35A	TNBC-DFS	0,468	1,06	RPS28	LABC-DFS	0,008	0,82
RPL35A	TNBC-DMFS	0,2529	1,12	RPS28	LABC-DMFS	0,2044	0,86
RPL35A	TNBC-OS	0,7974	1,03	RPS28	LABC-OS	0,0013	0,73
RPL38	HER2-DFS	0,5931	1,04	RPS28	LBBC-DFS	0,0102	0,83
RPL38	HER2-DMFS	0,9761	1	RPS28	LBBC-DMFS	0,0065	0,76
RPL38	HER2-OS	0,1534	1,16	RPS28	LBBC-OS	0,2157	0,89
RPL38	LABC-DFS	0,5104	1,05	RPS28	TNBC-DFS	0,5827	0,96
RPL38	LABC-DMFS	0,8363	0,98	RPS28	TNBC-DMFS	0,3104	0,9
RPL38	LABC-OS	0,5207	0,94	RPS28	TNBC-OS	0,853	1,02
RPL38	LBBC-DFS	0,5414	0,96	RPS3	HER2-DFS	0,2992	0,92
RPL38	LBBC-DMFS	0,921	1,01	RPS3	HER2-DMFS	0,0385	0,79
RPL38	LBBC-OS	0,9088	1,01	RPS3	HER2-OS	0,6837	0,96
RPL38	TNBC-DFS	0,7526	1,02	RPS3	LABC-DFS	0,0239	0,85
RPL38	TNBC-DMFS	0,714	0,96	RPS3	LABC-DMFS	0,1507	0,85
RPL38	TNBC-OS	0,1724	1,15	RPS3	LABC-OS	0,031	0,81
RPL5	HER2-DFS	0,019	0,82	RPS3	LBBC-DFS	0,4587	1,06
RPL5	HER2-DMFS	0,0046	0,73	RPS3	LBBC-DMFS	0,9922	1
RPL5	HER2-OS	0,1581	0,86	RPS3	LBBC-OS	0,9193	0,99
RPL5	LABC-DFS	0,0037	0,81	RPS3	TNBC-DFS	0,3722	1,07
RPL5	LABC-DMFS	0,1876	0,86	RPS3	TNBC-DMFS	0,4678	1,08
RPL5	LABC-OS	0,01	0,78	RPS3	TNBC-OS	0,4168	1,09
RPL5	LBBC-DFS	0,5867	0,96	RPS3A	HER2-DFS	0,1344	0,88
RPL5	LBBC-DMFS	0,1048	0,85	RPS3A	HER2-DMFS	0,3696	0,91
RPL5	LBBC-OS	0,6641	1,04	RPS3A	HER2-OS	0,8438	0,98
RPL5	TNBC-DFS	0,9615	1	RPS3A	LABC-DFS	< 0,0001	0,71
RPL5	TNBC-DMFS	0,3649	0,91	RPS3A	LABC-DMFS	0,0042	0,71
RPL5	TNBC-OS	0,6144	0,95	RPS3A	LABC-OS	< 0,0001	0,64
RPL7	HER2-DFS	0,9037	1,01	RPS3A	LBBC-DFS	0,004	0,81
RPL7	HER2-DMFS	0,156	0,86	RPS3A	LBBC-DMFS	0,0105	0,78
RPL7	HER2-OS	0,8312	1,02	RPS3A	LBBC-OS	0,0041	0,75
RPL7	LABC-DFS	0,0009	0,78	RPS3A	TNBC-DFS	0,1148	0,88
RPL7	LABC-DMFS	0,0025	0,7	RPS3A	TNBC-DMFS	0,1299	0,86
RPL7	LABC-OS	0,0001	0,68	RPS3A	TNBC-OS	0,0898	0,84
RPL7	LBBC-DFS	0,0007	0,77	RPS4X	HER2-DFS	0,0036	0,78

RPL7	LBBC-DMFS	0,0063	0,76	RPS4X	HER2-DMFS	0,1295	0,85
RPL7	LBBC-OS	0,1066	0,85	RPS4X	HER2-OS	0,0476	0,81
RPL7	TNBC-DFS	0,1157	1,14	RPS4X	LABC-DFS	< 0,0001	0,69
RPL7	TNBC-DMFS	0,4223	1,09	RPS4X	LABC-DMFS	0,0122	0,74
RPL7	TNBC-OS	0,0951	1,2	RPS4X	LABC-OS	< 0,0001	0,6
RPS12	HER2-DFS	0,5374	1,05	RPS4X	LBBC-DFS	0,0113	0,83
RPS12	HER2-DMFS	0,9142	1,01	RPS4X	LBBC-DMFS	0,0532	0,82
RPS12	HER2-OS	0,4201	1,09	RPS4X	LBBC-OS	0,0179	0,79
RPS12	LABC-DFS	0,0321	0,85	RPS4X	TNBC-DFS	0,9786	1
RPS12	LABC-DMFS	0,0751	0,81	RPS4X	TNBC-DMFS	0,6658	0,96
RPS12	LABC-OS	0,0021	0,74	RPS4X	TNBC-OS	0,7615	0,97
RPS12	LBBC-DFS	0,095	0,88	RPS7	HER2-DFS	0,482	0,94
RPS12	LBBC-DMFS	0,0233	0,8	RPS7	HER2-DMFS	0,1592	0,86
RPS12	LBBC-OS	0,0207	0,79	RPS7	HER2-OS	0,9168	0,99
RPS12	TNBC-DFS	0,1306	0,89	RPS7	LABC-DFS	0,0003	0,77
RPS12	TNBC-DMFS	0,5994	0,95	RPS7	LABC-DMFS	0,0117	0,74
RPS12	TNBC-OS	0,2986	0,89	RPS7	LABC-OS	< 0,0001	0,62
RPS13	HER2-DFS	0,2427	0,91	RPS7	LBBC-DFS	0,3437	0,93
RPS13	HER2-DMFS	0,0935	0,83	RPS7	LBBC-DMFS	0,2712	0,9
RPS13	HER2-OS	0,4638	0,92	RPS7	LBBC-OS	0,423	0,92
RPS13	LABC-DFS	0,0017	0,79	RPS7	TNBC-DFS	0,9599	1
RPS13	LABC-DMFS	0,0896	0,82	RPS7	TNBC-DMFS	0,611	0,95
RPS13	LABC-OS	0,0012	0,73	RPS7	TNBC-OS	0,9154	0,99
RPS13	LBBC-DFS	0,4269	0,94	RPSA	HER2-DFS	0,0582	1,18
RPS13	LBBC-DMFS	0,2514	0,89	RPSA	HER2-DMFS	0,8312	1,02
RPS13	LBBC-OS	0,1793	0,88	RPSA	HER2-OS	0,4549	1,09
RPS13	TNBC-DFS	0,4517	0,94	RPSA	LABC-DFS	0,0354	0,85
RPS13	TNBC-DMFS	0,2564	0,89	RPSA	LABC-DMFS	0,3231	0,89
RPS13	TNBC-OS	0,8716	0,98	RPSA	LABC-OS	0,039	0,82
RPS15A	HER2-DFS	0,7049	0,97	RPSA	LBBC-DFS	0,4851	0,95
RPS15A	HER2-DMFS	0,4671	0,92	RPSA	LBBC-DMFS	0,4753	0,93
RPS15A	HER2-OS	0,6245	1,05	RPSA	LBBC-OS	0,0867	0,84
RPS15A	LABC-DFS	0,1542	0,9	RPSA	TNBC-DFS	0,2314	1,1
RPS15A	LABC-DMFS	0,2823	0,88	RPSA	TNBC-DMFS	0,0203	1,27
RPS15A	LABC-OS	0,0177	0,79	RPSA	TNBC-OS	0,1695	1,16

SUPPLEMENTARY TABLE 4 – PREDICTED REGULATORY INTERACTIONS BETWEEN MIRNAS AND TRANSCRIPTION FACTORS (TFs). (A) TF-RP REGULATORY PAIRS PREDICTED IN AT LEAST TWO DATABASES. (B) miRNA-RP REGULATORY PAIRS WITH AGREEMENT SCORE > 1.

SUPPLEMENTARY TABLE 4A

Ribosomal protein	Transcription factors	Ribosomal protein	Transcription Factor
<i>RPL11</i>	<i>E2F1, ETS1, FLI1, FOXA1, MYC, SPI1</i>	<i>RPS16</i>	<i>ATF3, E2F1, EGR1, ETS1, FLI1, MYC, SPI1, STAT1</i>
<i>RPL13A</i>	<i>ETS1, MYBL2, MYC, SPI1</i>	<i>RPS18</i>	<i>MYC</i>
<i>RPL22</i>	<i>ETS1, MYC, SPI1, TRIM28</i>	<i>RPS19</i>	<i>ETS1, FLI1, MYBL2, MYC, SPI1, STAT1, TAL1, TCF3, TRIM28</i>
<i>RPL23</i>	<i>ETS1, FOXA1, JUN, MYBL2, MYC, SPI1</i>	<i>RPS2</i>	<i>EGR1, MYC, TCF3, TRIM28</i>
<i>RPL23A</i>	<i>ATF3, E2F1, EGR1, ETS1, FOXA1, JUN, MYBL2, MYC, TCF3</i>	<i>RPS20</i>	<i>E2F1, ETS1, MYC, SPI1, STAT1, TRIM28</i>
<i>RPL27</i>	<i>ATF3, ETS1, FLI1, JUN, MYC, SPI1, TAL1</i>	<i>RPS21</i>	<i>ATF3, ETS1, FLI1, MYBL2, MYC, SPI1, TAL1</i>
<i>RPL3</i>	<i>ATF3, E2F1, EGR1, ETS1, FLI1, MYB, MYBL2, MYC, TAL1, TCF3</i>	<i>RPS25</i>	<i>ETS1, FLI1, MYBL2, MYC, SPI1, TAL1, ETS1, FLI1, MYBL2, MYC, SPI1, TAL1</i>
<i>RPL30</i>	<i>E2F1, FLI1, MYC, SPI1</i>	<i>RPS26</i>	<i>ETS1, FLI1, FOXA1, MYBL2, MYC, SPI1, TAL1, TCF3</i>
<i>RPL32</i>	<i>ATF3, E2F1, ETS1, FLI1, MYC, SPI1, STAT1, TAL1, TRIM28</i>	<i>RPS27</i>	<i>E2F1, ETS1, FLI1, FOXA1, MYB, MYC, SPI1</i>
<i>RPL35</i>	<i>E2F1, ETS1, FOXA1, MYBL2, MYC, SPI1, TRIM28</i>	<i>RPS27A</i>	<i>EGR1, ETS1, JUN, MYB, MYC, SPI1</i>
<i>RPL35A</i>	<i>ATF3, MYBL2, STAT1, TAL1</i>	<i>RPS28</i>	<i>E2F1, FOXA1, MYBL2, MYC</i>
<i>RPL38</i>	<i>E2F1, EGR1, ETS1, FLI1, MYC, SPI1</i>	<i>RPS3</i>	<i>ETS1, FLI1, FOXA1, MYC, SPI1, TAL1</i>
<i>RPL5</i>	<i>ETS1, JUN, MYC, SPI1, STAT1, E2F1, ETS1, MYC, SPI1, STAT1</i>	<i>RPS3A</i>	<i>ETS1, FLI1, JUN, MYB, MYC, SPI1</i>
<i>RPL7</i>	<i>E2F1, ETS1, MYC, SPI1</i>	<i>RPS4X</i>	<i>EGR1, ETS1, MYC, TCF3, TRIM28</i>
<i>RPS12</i>	<i>EGR1, ETS1, MYBL2, MYC, SPI1</i>	<i>RPS7</i>	<i>ETS1, FLI1, MYB, MYC, TAL1</i>
<i>RPS13</i>	<i>ATF3, ETS1, MYC, SPI1</i>	<i>RPSA</i>	<i>E2F1, FLI1, MYC, SPI1</i>
<i>RPS15A</i>	<i>E2F1, ETS1, MYC, STAT1, TRIM28</i>		

SUPPLEMENTARY TABLE 4B

Ribosomal protein	miRNA
<i>RPL13A</i>	hsa-let-7c-3p, hsa-mir-204-5p, hsa-mir-32-5p hsa-mir-421, hsa-mir-486-5p, hsa-mir-5683 hsa-mir-625-5p, hsa-mir-96-5p
<i>RPL22</i>	hsa-let-7c-5p, hsa-mir-125b-2-3p, hsa-mir-1301-3p, hsa-mir-130b-5p, hsa-mir-139-3p, hsa-mir-140-3p, hsa-mir-141-3p, hsa-mir-145-5p, hsa-mir-148b-3p, hsa-mir-15b-3p, hsa-mir-16-5p, hsa-mir-181b-3p, hsa-mir-187-3p, hsa-mir-195-5p, hsa-mir-199b-5p, hsa-mir-200a-3p, hsa-mir-210-3p, hsa-mir-223-3p, hsa-mir-224-5p, hsa-mir-29a-3p, hsa-mir-29b-3p, hsa-mir-301a-3p, hsa-mir-3065-3p, hsa-mir-30a-3p, hsa-mir-30c-2-3p, hsa-mir-30e-3p, hsa-mir-323a-3p, hsa-mir-337-3p hsa-mir-33a-3p, hsa-mir-33a-5p, hsa-mir-33b-5p hsa-mir-342-3p, hsa-mir-370-3p, hsa-mir-376c-3p, hsa-mir-381-3p, hsa-mir-421, hsa-mir-450b-5p, hsa-mir-452-5p, hsa-mir-483-3p, hsa- mir-485-3p, hsa-mir-494-3p, hsa-mir-497-5p, hsa-mir-539-5p, hsa-mir-5683, hsa-mir-590-3p, hsa- mir-671-5p, hsa-mir-6892-5p hsa-mir-7-5p, hsa-mir-9-5p, hsa-mir-98-5p, hsa-mir-99a-3p
<i>RPL23</i>	hsa-let-7c-5p, hsa-let-7d-5p, hsa-mir-106a-5p, hsa-mir-106b-5p, hsa- mir-10a-5p, hsa-mir-10b-5p, hsa-mir-126-5p, hsa-mir-130a-3p, hsa- mir-130b-3p, hsa-mir-133a-3p, hsa-mir-140-5p, hsa-mir-141-3p, hsa- mir-141-5p, hsa-mir-142-5p, hsa-mir-143-5p, hsa-mir-145-3p, hsa-mir- 150-3p, hsa-mir-150-5p, hsa-mir-15b-5p, hsa-mir-16-1-3p, hsa-mir-16- 5p, hsa-mir-17-3p, hsa-mir-17-5p, hsa-mir-181a-3p, hsa-mir-181a-5p hsa-mir-181b-5p, hsa-mir-181d-5p, hsa-mir-182-5p, hsa-mir-18a-5p, hsa-mir-191-5p, hsa-mir-19a-3p, hsa-mir-200a-3p, hsa-mir-205-5p, hsa-mir-20b-3p, hsa-mir-20b-5p, hsa-mir-27a-5p, hsa-mir-301a-3p, hsa-mir-301a-5p, hsa-mir-3065-5p, hsa-mir-30a-3p, hsa-mir-30e-3p, hsa-mir-3170, hsa-mir-323a-3p, hsa-mir-335-3p, hsa-mir-340-5p, hsa- mir-342-3p, hsa-mir-3614-5p, hsa-mir-365a-3p, hsa-mir-365b-3p, hsa- mir-376c-3p, hsa-mir-377-5p, hsa-mir-378a-5p, hsa-mir-381-3p hsa-mir-3913-5p, hsa-mir-411-5p, hsa-mir-425-5p, hsa-mir-450b-5p, hsa-mir-4668-3p, hsa-mir-493-5p, hsa-mir-494-3p, hsa-mir-495-3p, hsa-mir-503-5p, hsa-mir-5683, hsa-mir-615-3p, hsa-mir-616-5p, hsa- mir-654-3p, hsa-mir-654-5p, hsa-mir-655-3p, hsa-mir-708-3p, hsa-mir- 708-5p, hsa-mir-7-1-3p, hsa-mir-758-5p, hsa-mir-7-5p, hsa-mir-93-5p, hsa-mir-940, hsa-mir-98-5p
<i>RPL23A</i>	hsa-let-7a-2-3p, hsa-let-7g-3p, hsa-mir-106a-5p, hsa-mir-106b-5p, hsa-mir-148b-5p, hsa-mir-17-5p, hsa-mir-190a-5p, hsa-mir-20b-5p, hsa-mir-3170, hsa-mir-345-5p, hsa-mir-378a-3p, hsa-mir-378c, hsa- mir-431-5p, hsa-mir-4677-3p, hsa-mir-485-3p, hsa-mir-487b-3p, hsa- mir-511-5p, hsa-mir-584-5p
<i>RPL27</i>	hsa-mir-103a-3p, hsa-mir-107, hsa-mir-142-3p, hsa-mir-655-3p
<i>RPL30</i>	hsa-mir-129-5p
<i>RPL32</i>	hsa-let-7c-5p, hsa-mir-103a-2-5p, hsa-mir-103a-3p, hsa-mir-106a-5p, hsa-mir-106b-5p, hsa-mir-107, hsa-mir-1301-3p, hsa-mir-143-3p, hsa- mir-145-3p, hsa-mir-145-5p, hsa-mir-15a-5p, hsa-mir-15b-5p, hsa-mir- 16-5p, hsa-mir-17-5p, hsa-mir-181b-3p, hsa-mir-182-5p, hsa-mir-188- 5p, hsa-mir-18a-5p, hsa-mir-192-5p, hsa-mir-195-5p, hsa-mir-20b-3p, hsa-mir-20b-5p, hsa-mir-215-5p, hsa-mir-27a-5p, hsa-mir-27b-5p, hsa-mir-28-3p, hsa-mir-30c-2-3p, hsa-mir-335-5p, hsa-mir-340-3p, hsa-mir-378a-3p, hsa-mir-378c, hsa-mir-452-3p, hsa-mir-493-5p, hsa-

	mir-497-5p, hsa-mir-539-5p, hsa-mir-92b-3p, hsa-mir-93-5p, hsa-mir-98-5p
<i>RPL35</i>	hsa-mir-200c-5p, hsa-mir-218-5p, hsa-mir-3614-5p
<i>RPL35A</i>	hsa-mir-196a-5p, hsa-mir-4668-3p
<i>RPL38</i>	hsa-mir-651-5p, hsa-mir-655-3p, hsa-mir-7-5p
<i>RPL7</i>	hsa-mir-452-5p, hsa-mir-539-5p
<i>RPS12</i>	hsa-mir-4677-3p
<i>RPS13</i>	hsa-mir-15b-3p, hsa-mir-190a-5p, hsa-mir-190b, hsa-mir-494-3p
<i>RPS15A</i>	hsa-let-7a-3p, hsa-let-7c-5p, hsa-let-7d-5p, hsa-mir-101-5p, hsa-mir-10a-5p, hsa-mir-10b-3p, hsa-mir-10b-5p, hsa-mir-1247-3p, hsa-mir-1271-5p, hsa-mir-1296-5p, hsa-mir-134-5p, hsa-mir-141-3p, hsa-mir-143-5p, hsa-mir-145-3p, hsa-mir-146b-5p, hsa-mir-148b-5p, hsa-mir-150-3p, hsa-mir-150-5p, hsa-mir-181b-3p, hsa-mir-18a-5p, hsa-mir-194-3p, hsa-mir-200a-5p, hsa-mir-200b-3p, hsa-mir-200b-5p, hsa-mir-200c-3p, hsa-mir-20b-3p, hsa-mir-27a-5p, hsa-mir-29a-3p, hsa-mir-29b-3p, hsa-mir-3065-3p, hsa-mir-3065-5p, hsa-mir-30c-2-3p, hsa-mir-3170, hsa-mir-32-5p, hsa-mir-326, hsa-mir-335-3p, hsa-mir-335-5p, hsa-mir-3614-5p, hsa-mir-362-3p, hsa-mir-369-5p, hsa-mir-377-5p, hsa-mir-378a-3p, hsa-mir-378c, hsa-mir-3913-5p, hsa-mir-421, hsa-mir-429, hsa-mir-432-5p, hsa-mir-452-5p, hsa-mir-455-3p, hsa-mir-4668-3p, hsa-mir-5683, hsa-mir-590-3p, hsa-mir-616-5p, hsa-mir-654-5p, hsa-mir-7-1-3p, hsa-mir-758-3p, hsa-mir-758-5p, hsa-mir-769-5p, hsa-mir-92b-3p, hsa-mir-940, hsa-mir-96-5p, hsa-mir-98-3p, hsa-mir-98-5p
<i>RPS16</i>	hsa-let-7a-3p, hsa-mir-145-3p, hsa-mir-381-3p, hsa-mir-452-5p, hsa-mir-4668-3p, hsa-mir-616-5p, hsa-mir-98-3p
<i>RPS18</i>	hsa-mir-331-5p
<i>RPS19</i>	hsa-mir-616-5p
<i>RPS2</i>	hsa-mir-10b-3p
<i>RPS20</i>	hsa-mir-100-5p, hsa-mir-145-3p, hsa-mir-16-1-3p, hsa-mir-199b-5p, hsa-mir-421, hsa-mir-99a-5p
<i>RPS21</i>	hsa-mir-335-3p
<i>RPS26</i>	hsa-mir-335-3p hsa-mir-454-3p
<i>RPS27A</i>	hsa-mir-187-3p, hsa-mir-218-5p, hsa-mir-451a, hsa-mir-485-3p, hsa-mir-511-5p
<i>RPS3</i>	hsa-mir-106a-5p, hsa-mir-106b-5p, hsa-mir-142-3p, hsa-mir-148b-5p, hsa-mir-17-5p, hsa-mir-181b-5p, hsa-mir-193a-5p, hsa-mir-205-5p, hsa-mir-20b-5p, hsa-mir-23a-5p, hsa-mir-296-5p, hsa-mir-30c-2-3p, hsa-mir-3127-5p, hsa-mir-3170, hsa-mir-323a-3p, hsa-mir-337-3p, hsa-mir-33a-3p, hsa-mir-654-3p, hsa-mir-708-5p, hsa-mir-93-5p
<i>RPS3A</i>	hsa-mir-93-5p
<i>RPS4X</i>	hsa-mir-103a-3p, hsa-mir-107, hsa-mir-125b-5p, hsa-mir-130a-3p, hsa-mir-130b-3p, hsa-mir-130b-5p, hsa-mir-146b-5p, hsa-mir-148a-3p, hsa-mir-148b-3p, hsa-mir-182-5p, hsa-mir-195-3p, hsa-mir-19b-1-5p, hsa-mir-20b-5p, hsa-mir-301a-3p, hsa-mir-335-3p, hsa-mir-421 hsa-mir-450b-5p, hsa-mir-454-3p, hsa-mir-5683, hsa-mir-93-5p, hsa-mir-9-5p

13. REFERÊNCIAS (CAPÍTULO 1 & CAPÍTULO 2)

AJORE, R.; RAISER, D.; MCCONKEY, M.; JOUD, M.; BOIDOL, B.; MAR, B.; SAKSENA, G.; WEINSTOCK, D, M.; ARMSTRONG, S.; ELLIS, S, R.; ELBERT, B, L.; NILSSON, B. Deletion of ribosomal protein genes is a common vulnerability in human cancer, especially in concert with TP53 mutations. **EMBO Molecular Medicine**, 9(4):498-507, 2017.

AKTER, K.; MANSOUR, M, A.; HYODO, T.; SENG, T. FAM98A associates with DDX1-C14orf166-FAM98B in a novel complex involved in colorectal cancer progression. **Biochemistry & cell biology**, 2017.

AL-WAJEE, A, S.; SALHIMI, S, M.; AL-MANSOUB, A, A.; KHALID, I, A.; HARVEY, T, M.; LATIFF, A.; ISMAIL, M, A. Comparative proteomic analysis of different stages of breast cancer tissues using ultra high performance liquid chromatography tandem mass spectrometer. **Plos One**, 2020.

AMAT, S.; PENAULT-LLORCA, F.; CURE, H.; BOUEDEC, G.; ACHARD, J, L.; PRAGH, I, V.; FEILLEL, V.; MUORET-REYNIER, M, A.; DAUPLAT, J.; CHOLLET, P. Scarff-Bloom-Richardson (SBR) grading: a pleiotropic marker of chemosensitivity in invasive ductal breast carcinomas treated by neoadjuvant chemotherapy. **International journal of oncology**, 20:791-796, 2002.

AN, X.; LIN, X.; YANG, A.; JIANG, Q.; GENG, B.; HUANG, M.; LU, J.; XIANG, Z.; YUAN, Z.; WANG, S.; SHI, Y.; ZHU, J. Cavin3 Suppresses Breast Cancer Metastasis via Inhibiting AKT Pathway. **Frontiers in pharmacology**, 11, 2020.
ANAND, S.; SAMUEL, M.; ANG, C, S.; KEERTHIKUMAR, S.; MATHIVANAN, S. Label-based and label-free strategies for protein quantitation. **Proteome Bioinformatics, Methods in Molecular Biology**, 1549, 2017.

ANDRADE, L, J, O.; OLIVEIRA, G, C, M.; BITTENCOURT, A, M, V.; MELO, P, R, S. Protein molecular modeling of genetic markers for thyroid cancer. **J Bras Patol Med Lab**, 52(4), 324-337, 2016.

ANDRÉS-LEÓN, E.; PEÑA, D, G.; GÓMEZ-LÓPEZ, G.; PISANO, D, G. miRGate: a curated database of human, mouse and rat miRNA–mRNA targets. **Database**, 2015.

AZIZI, E.; CARR, A, J.; PLITAS, G.; MAZUTIS, L.; RUDENSKY, A, Y.; PE'ER, S. Single-cell map of diverse immune phenotypes in the breast tumor microenvironment. **Cell** 174, 1293–1308, 2018.

BAI, L.; DENG, X.; LI, Q.; WANG, M.; AN, W.; A, D.; GAO, Z.; XIE, Y.; DAI, Y.; CONG, Y, S. Down-Regulation of the Cavin Family Proteins in Breast Cancer. **Journal of Cellular Biochemistry**, 113:322–328, 2012.

BALIC, M.; THOMSEN, C.; WURSTLEIN, R.; GNANT, M.; HARBECK, N. St. Gallen/Vienna 2019: A Brief Summary of the Consensus Discussion on the Optimal Primary Breast Cancer Treatment. **Breast care**, 14:103-110, 2019.

BARH, D. Omics Approaches in Breast Cancer. Springer, 2014.

BEE, A.; BREWER, D.; BEESLEY, C.; DODSON, A.; FOROOTAN, S.; DICKISON, T.; GERARD, P.; LANE, B.; YAO, S.; COOPER, C, S.; DJAMGOZ, M, B, A.; GOSDEN, C, M.; KE, Y.; FOSTER, C, S. siRNA Knockdown of Ribosomal Protein Gene RPL19 Abrogates the Aggressive Phenotype of Human Prostate Cancer. **PlosOne**, 6(7), 2011.

BEE, A.; KE, Y.; FOROOTAN, S.; LIN, K.; BEESLEY, C.; FORREST, S, E.; FOSTER, C, S. Ribosomal Protein L19 Is a Prognostic Marker for Human Prostate Cancer. **Clinical cancer research**, 12(7), 2006.

BOUCHAL, P.; SCHUBERT, O, T.; FAKTOR, J.; BUDINSKA, E.; NENUTIL, R.; AEBERSOLD, R. Breast Cancer Classification Based on Proteotypes Obtained by SWATH Mass Spectrometry. **Cell Reports**, 28, 832–843, 2019.

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

BRUNO, P.; LIU, Y.; PARK, G, Y.; MURAI, J.; KOCH, C, E.; EISEN, T, J.; PRITCHARD, J, R.; POMMIER, Y.; LIPPARD, S, J.; HEMANN, M, T. A subset of platinum-containing chemotherapeutic agents kills cells by inducing ribosome biogenesis stress. **Nature medicine**, 2018.

BRUNORO, G, V, F.; CARVALHO, P, C.; BARBOSA, V, C.; PAGNONCELLI, D.; GALLO, G, V, D, M.; PERALES, J.; ZAHEDI, R, P.; VALENTE, R, H.; NEVES-FERREIRA, A, G, C. Differential proteomic comparison of breast cancer secretome using a quantitative paired analysis workflow. **BMC Cancer**, 19:365, 2019.

CAO, B.; FANG, Z.; LIAO, P.; ZHOU, X.; XIONG, J.; XENG, S.; LU, H. Cancer-mutated ribosome protein L22 (RPL22/eL22) suppresses cancer cell survival by blocking p53-MDM2 circuit. **Oncotarget**, 8(53): 90651-90661, 2017.

CAPPELLETTI, V.; IORIO, E.; MIODINI, P.; SILVESTRI, M.; DUGO, M.; DAIDONE, M, G. Metabolic footprints and molecular subtypes in breast cancer. **Disease markers**, 10, 2017.

CARITHERS, L, J.; ARDLIE, K.; BARCUS, M.; BRANTON, P, A.; BRITTON, A.; BUIA, S, A.; COMPTON, C, C.; DELUCA, D, S.; PETER-DEMCHOK, J.; GELFAND, E, T.; GUAN, P.; KORZENIEWSKI, G, E.; LOCKHART, N, C.; RABINER, C, A.; RAO, A, K.; ROBINSON, K, L.; ROCHE, N, V.; SAWYER, S, J.; SEGRÈ, A, V.; SHIVE, C, E.; SMITH, A, M.; SOBIN, L, H.; UNDALE, A, H.; VALENTINO, K, M.; VAUGHT, J.; YOUNG, T, R.; MOORE, H, M. A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. **BIOPRESERVATION AND BIOBANKING**, 13(5), 2015.

CARLSSON, H.; PETERSSON, S.; ENERBACK, C. Cluster analysis of S100 gene expression and genes correlating to psoriasin (S100A7) expression at different stages of breast cancer development. **International journal of oncology** 27: 1473-1481, 2005.

CARVALHO, A, S.; BAETA, H.; SILVA, B, C.; MORAES, M, C, S.; BODO, C.; BECK, H, C.; RODRIGUEZ, M, S.; SARASWAT, M.; PANDEY, A.; MATTHIESEN, R. Extra-cellular vesicles carry proteome of cancer hallmarks. *Frontiers in Bioscience*, **Landmark**, 25, 398-436, 2020.

CASTRO, R, X. Adesão das usuárias das unidades básicas de saúde do município de Rubelita – Minas Gerais, aos métodos de detecção precoce de câncer de mama. Trabalho de conclusão de curso, Araçuí, 2011.

CELA, I.; MATTEO, A, D.; FEDERICI, L.; Nucleophosmin in Its Interaction with Ligands. **International Journal of Molecular Science**, 21(4885), 2020.

CERAMI, E.; GAO, J.; DOGRUSOZ, U.; GROSS, B, E.; SUMER, S, O.; AKSOY, B, A.; JACOBSEN, A.; BYRNE, C, J.; HEUER, M, L.; LARSSON, E.; ANTIPIN, Y.; REVA, B.; GOLDBERG, A, P.; SANDER, C.; SCHULTZ, N. The cBio Cancer Genomics Portal: An Open Platform for exploring Multidimensional Cancer Genomics Data. **American association of cancer research**, 2(5):401-404, 2012.

CGHFBC. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. **Lancet Oncol**, 13: 1141–51, 2012.

CHAKAFANA, G.; SHONHAI, A. The Role of Non-Canonical Hsp70s (Hsp110/Grp170) in Cancer. **Cells**. 28;10(2):254, 2021.

CHAN, L, W, C.; LIN, X.; YUNG, G.; LUI, T.; CHIU, Y, M.; WANG, F.; TSUI, N, B, Y.; CHO, W, C, S.; YIP, S, P.; SIU, P, M.; WONG, S, C, C.; TUNG, B. Novel structural co-expression analysis linking the NPM1- associated ribosomal

biogenesis network to chronic myelogenous leukemia. **Scientific reports**, 5(10973), 2015.

CHANG, H, Y.; JUAN, H, F. Proteomic Data Analysis: Functional Enrichment. In: *A Practical Guide to Cancer Systems Biology*, 2017.

CHEN, C.; HOU, J.; TANNER, J, J.; CHENG, J. Bioinformatics methods for mass spectrometry-based proteomics data analysis. **Int. J. Mol. Sci.**, 21, 2873, 2020.

CHENG, Z.; DAI, Y.; HUANG, W.; ZHONG Q.; ZHU, P.; ZHANG, W.; WU, Z.; LIN, Q.; ZHU, H.; CUI, L.; QIAN, T.; DENG, C.; FU, L.; LIU, Y.; ZENG, T. Prognostic Value of MicroRNA-20b in Acute Myeloid Leukemia. **Frontiers in oncology**, 10, 2021.

CHUNG, L.; SHIBLI, S.; MOORE, K.; ELDER, E, E.; BOYLE, F, M.; MARSH, D, J.; BAXTER, R, C. Tissue biomarkers of breast cancer and their association with conventional pathologic features. *British Journal of Cancer*, 108, 351–360, 2013.

CONGLONG, X.; ZHANG, Z.; YU, H.; YU, M.; YANG, T.; MIAO, M.; SHI, H. Up-regulation of CNDP2 facilitates the proliferation of colon cancer. **BMC Gastroenterology**, 14:96, 2014.

CORRÊA, S.; PANIS, C.; BINATO, R.; HERRERA, A, C.; PIZZATTI, L.; ABDELHAY, E. Identifying potential markers in Breast Cancer subtypes using plasma label-free proteomics. **Journal of Proteomics**, 151, 33–42, 2017.

DA COSTA, G, G.; GOMIG, T, H, B.; KAVISKI, R.; SOUSA, K, S.; KUKOLJ, C.; LIMA, R, S.; URBAN, C, A.; CAVALLI, I, J.; RIBEIRO, E, M, S, F. Comparative proteomics of tumor and paired normal breast tissue highlights potential biomarkers in breast cancer. **Cancer genomics & proteomics**, 12: 251-262, 2015.

DAI, X.; LI, T.; BAI, Z.; YANG, Y.; LIU, X.; ZHANG, J.; SHI, B. Breast cancer intrinsic subtype classification, clinical use and future trends. **Am J Cancer Res**; 5(10):2929-2943, 2015.

DALAL, N.; JALANDRA, R.; SHARMA, M.; PRAKASH, H.; MAKHARIA, G, K.; SOLANKI, P, R.; SINGH, R.; KUMAR, A. Omics technologies for improved diagnosis and treatment of colorectal cancer: Technical advancement and major perspectives. **Biomedicine & Pharmacotherapy**, 131, 110648, 2020.

DAVE, B.; GRANADOS-PRINCIPAL, S.; ZHU, R.; BENZ, S.; RABIZADEH, S.; SOON-SHION, P.; YU, K, D.; SHAO, Z.; LI, X.; GILCREASE, M.; LAI, Z.; CHEN, Y.; HUANG, T, H, M.; SHEN, H.; LIU, X.; FERRARI, M.; ZHAN, M.; WONG, S, T, C.; KUMARASWAMI, M.; MITTAL, V.; CHEN, X.; GROSS, S, S.; CHANG, J, C. Targeting RPL39 and MLF2 reduces tumor initiation and metastasis in breast cancer by inhibiting nitric oxide synthase signaling. **PNAS**, 111(24), 8838-8843, 2014.

DIAZ, M, I.; DÍAZ, P.; BENNETT, J, C.; URRÁ, H.; ORTIZ, R.; ORELLANA, P, C.; HETZ, .; QUEST, A, F, G. Caveolin-1 suppresses tumor formation through the inhibition of the unfolded protein response. **Cell Death and Disease**,11:648, 2020.

DIREITO, I.; MONTEIRO, L.; MELO, T.; FIGUEIRA, D.; LOBO, J.; ENES, V.; MOURA, G.; HENRIQUE, R.; SANTOS, M, A, S.; JERÓNIMO, C.; AMADO, F.; FARSILHA, M.; HELGUERO, L, A. Protein aggregation patterns inform about breast cancer response to antiestrogens and reveal the RNA ligase RTCB as mediator of acquired tamoxifen resistance. **Cancers**, 13, 3195, 2021.

DONG, Z.; JIANG, H.; LIANG, S.; WANG, Y.; JIANG, W.; ZHU, C. Ribosomal Protein L15 is involved in Colon Carcinogenesis. **International Journal of Medical Sciences**, 16(8): 1132-1141, 2019.

DRYGIN, D.; SIDDIQUI-JAIN, A.; O'BRIEN, S. Anticancer Activity of CX-3543: A Direct Inhibitor of rRNA biogenesis. **Cancer research**, 69:7653-7661, 2009.

DUTTON-REGESTER, K.; GARTNER, J, J.; EMMANUEL, R.; QUTOB, N.; DAVIES, M, A.; GERSHENWALD, J, E.; ROBINSON, W.; ROBINSON, S.; ROSENBERG, S, A.; SCOLYER, R, A.; MANN, G, J.; THOMPSON, J, F.; HAYWARD, N, K.; SAMUELS, Y. A highly recurrent RPS27 5'UTR mutation in melanoma. **Oncotarget**, 5(10):2912-2917, 2014.

EBRIGHT, R, Y.; LEE, S.; WITTNER, B, S. NIEDERHOFFER, K, L.; NICHOLSON, B, T.; BARDIA, A.; TRUESDELL, S.; WILEY, D, F.; WESLEY, B.; LI, S.; MAI, A.; ACETO, N.; VINCENT-JORDAN, N.; SZABOLCS, A.; CHIRN, B.; KREUZER, J.; COMAILLS, V.; KALINNICH, M.; HAAS, W.; TING, D, T.; TONER, M.; VASUDEVAN, S.; HABER, D, A.; MASHESWARAN, S.; MICALIZZI, D, S. Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. **Science**, 367(6485): 1468-1473, 2020.

ELLIS, H.; MAHADEVAN, V. Anatomy and physiology of the breast. **Surgery (Oxford)**, 31(1), 11-14, 2013.

EL-SHARKAWY, A, M. Principles and Practice of Cancer Prevention and Control. **OMICS Group eBooks**, 2014.

ERCAN, C.; VAN DIEST, P, J.; VOOIJS, M. Mammary development and breast cancer: The Role of Stem Cells. **Curr Mol Med**; 11(4): 270–285, 2011.

FANG, Y, F.; LEE, Y, C.; WANG, Y, Y.; WANG, C, H.; HOU, M, F.; YUAN, S, S.; YWHAE promotes proliferation, metastasis, and chemoresistance in breast cancer cells. **J Med Sci**.;35:408–416, 2019.

FONG, Y.; EVANS, J.; BROOK, D.; KENKRE, J.; JARVIS, P.; GOWER-THOMAS, K. The Nottingham Prognostic Index: five- and ten-year data for all-cause survival within a screened population. **Ann R Coll Surg Engl**, 97:137-139, 2015.

FRAGOMENI, S, M.; SCIALLI, A.; JERUSS, J, S. Molecular subtypes and local-regional control of breast cancer. **Surg Oncol Clin N Am**; 27(1): 95–120. 2019.

FRAGOMENI, S, M.; SCIALLI, A.; JERUSS, J, S. Molecular subtypes and local-regional control of breast cancer. **Surg Oncol Clin N Am**; 27(1): 95–120. 2019.

FU, W.; SUN, H.; ZHAO, Y.; CHEN, M.; YANG, X.; LIU, Y.; JIN, W. BCAP31 drives TNBC development by modulating ligand-independent EGFR trafficking and spontaneous EGFR phosphorylation. **Theranostics**, 9(22), 6468-6484, 2019.

FUNG, S, M.; WONG, X, Y.; LEE, S, X.; MIAO, H.; HARTMAN, M.; WEE, H, M. Performance of single-nucleotide polymorphisms in breast cancer risk prediction models: A systematic review and meta-analysis. **Cancer Epidemiol Biomarkers Prev**; 28(3), 2019.

GALEA, M, H.; BLAMEY, R, W.; ELSON, C, E.; ELLIS, I, O. The Nottingham Prognostic Index in primary breast cancer. **Breast cancer research and treatment**, 22:207-219, 1992.

GAM, L, H. Breast cancer and protein biomarkers. **World J Exp Med**, 20(2): 86-91, 2012.

GAMPLE, P.; JANOENSRI, R.; WANG, H.; TAN, F.; MORAN, M.; BROWN, T.; FLAMENT-AUVIGNE, I.; RAKHA, E, A.; TOSS, M.; DABBS, D, J.; REGITNIG, P.; OLSON, N.; WREN, J, H.; ROBINSOM ,C.; CORRADO, G, S.; PENG, L, H.; LIU, Y.; MERMEL, G, H.; STEINER, D, F.; CHEN, P, H, C. Determining breast cancer biomarker status and associated morphological features using deep learning. **COMMUNICATIONS MEDICINE**, 1(14), 2021.

GAN, Y.; YE, F.; HE, X, X.; The role of YWHAZ in cancer: A maze of opportunities and challenges. **Journal of Cancer**, 11, 2020.

GERMAIN, D, R.; GRAHAM, K.; GLUBRECHT, S, D.; HUGH, J, C; MACKEY, J, R.; GODBOUT R. DEAD box 1: a novel and independent prognostic marker for early recurrence in breast cancer. **Breast Cancer Res Treat**, 127:53–63, 2011.

GOLDHIRSCH, A.; WINER, E, P.; COATES, A, S.; GELBER, R, D.; PICCART-GEBHART, M.; THURLIMANN, B.; SENN, H, J. Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. **Annals of Oncology**, 2:2206-2223, 2013.

GOMIG, T, H, B.; CAVALLI, I, J.; SOUZA, L, R, R.; VIEIRA, E.; LUCENA, A, C, R.; BATISTA, M.; MACHADO, K, C.; MARCHINI, F, K.; MARCHI, F, A.; LIMA,

R, S.; URBAN, C, A.; CAVALLI, L, R.; RIBEIRO, E, M, S, F. Quantitative label-free mass spectrometry using contralateral and adjacent breast tissues reveal differentially expressed proteins and their predicted impacts on pathways and cellular functions in breast cancer. **Journal of Proteomics**, 199, 1–14, 2019.

GOUDARZI, K, M.; LINDSTROM, M, S. Role of ribosomal protein mutations in tumor development (Review). **International journal of oncology**, 48(1313-1324), 2016.

GRAVES, P, S.; HAYSTEAD, T, A, J.; Molecular Biologist's Guide to Proteomics. **Microbiology and molecular biology reviews**, 39–63, 2002.

GROSSET, A, A.; LABRIE, M.; VLADOIU, M, C.; YOUSEF, E, M.; GABOURY, L.; PIERRE, Y, S. Galectin signatures contribute to the heterogeneity of breast cancer and provide new prognostic information and therapeutic targets. **Oncotarget**, 7(14), 2016.

HADDAD, C, F. Hereditary breast cancer: review and current approach. *Mastology*;30:e20200042, 2020.

HANAHAN, D.; WEINBERG, R, A. Hallmarks of Cancer: The Next Generation. **Cell**, 144, 4, 2011.

HARRELL, F, J. With contributions from Dupont C and many others. Hmisc: Harrell Miscellaneous. R package version 3.17–1. CRAN.R-project.org/package=Hmisc. **Accessed 31 dez, 2021.**

HAUKAAS, T H.; EUCEDA, L,R.; GISKEØDEGÅRD, G,F.; LAMICHHANE, S.; KROHN, M.; JERNSTRÖM, S.; AURE, M,R.; LINGJÆRDE, O,C.; SCHLICHTING, E.; GARRED, Ø.; DUE, E, U.; MILLS, G, B.; SAHLBERG, K, K.; BØRRESEN-DALE, A, L.; BATHEN, T, F. Metabolic clusters of breast cancer in relation to gene- and protein expression subtypes. **Cancer & Metabolism**, 4:12, 2016.

HE, Z.; XU, Q.; WANG, X.; WANG, J.; MU, X.; CAI, Y.; QIAN, Y.; SHAO, W.; SHAO, Z. RPLP1 promotes tumor metastasis and is associated with a poor prognosis in triple-negative breast cancer patients. **Cancer cell int.**, 18, 2018.

HEBERLE, H.; MEIRELLES, G, V.; SILVA, F, R, da.; TELLES, G, P.; MINGHIM, R. InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics**, 16:169, 2015.

HESKE, C, M. Beyond energy metabolism: Exploiting the additional roles of NAMPT for cancer therapy. **Frontiers in oncology**, 9(1514), 2020.
HOFMAN, I, J, F.; PATCHETT, S.; DUIN, M, V,. GEERDENS, E.; VERBEECK, J.; MICHAUX, L.; DELFORGE, M.; SONNEVELD, P.; JOHANSON, A, W.;

KEERSMAECKER, K. Low frequency mutations in ribosomal proteins RPL10 and RPL5 in multiple myeloma. **HAEMATOLOGICA**, 102, 2017.

HONG, W, H.; SHABIHKHANI, M.; TELESCA, D.; YANG, S.; TSO, J, L.; MENKIVAR, J, C.; WEI, B.; LUCEY, G, M.; MARENINOV, S.; CHEN, Z.; LIAU, L, M.; LAI, A.; NELSON, S, F.; CLOUGHESY, T, F.; TSO, C, L. Ribosomal Proteins RPS11 and RPS20, Two Stress-Response Markers of Glioblastoma Stem Cells, Are Novel Predictors of Poor Prognosis in Glioblastoma Patients. **PlosOne**, 2015.

HUANG, D, W.; SHERMAN, B, T.; LEMPICKI, R, A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. **Nature protocols**, 4(1), 44-57, 2009.

HUANG, X, P.; ZHAO, C, X.; LI, Q, J.; CAI, Y.; LIU, F, X.; H, H. XU, X.; HAN, Y, L.; WU, M.; ZHAN, Q, M.; WANG, M, R. Alteration of RPL14 in squamous cell carcinomas and preneoplastic lesions of the esophagus. **Gene**, 366:161-168, 2006.

INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA. Estimativa 2020: Incidência de câncer no Brasil. Rio de Janeiro, **INCA**, 2019.

IQBAL, M, J.; JAVED, Z.; SADIA, H.; QURESHI, I, A.; IRSHAD, A.; AHMED, R.; MALIK, K.; ABBAS, A.; PEZZANI, R.; SHARIFI-RAD, J. Clinical applications of artificial intelligence and machine learning in cancer diagnosis: looking into the future. **Cancer Cell Int**, 21:270, 2021.

JAIN, K, K.; Role of proteomics in diagnosis of cancer. **Technology in Cancer Research & Treatment**, 1(4), 2002.

JERBY, L.; WOLF, L.; DENKERT, G.; STEIN, G, Y.; HILVO, M.; ORESIC, M.; GEIGER, T.; RUPPIN, R. Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer. **Cancer Res**; 72(22), 2012.

JÉZÉQUEL, P.; GOURAUD, W.; AZZOUZ, F, B.; GUÉRIN-CHARBONNEL, C.; JUIN, P, P.; LASLA, H.; COMPONE, M. bc-GenExMiner 4.5: new mining module computes breast cancer differential gene expression analyses. **DATABASE**, 00: 1-10, 2021.

JUBAN, G.; GIRAUD, G.; GUYOT, B.; BELIN, S.; DIAZ, J, J.; STARCK, J.; GUILLOUF, C.; MOREAU-GACHELIN, F.; MORLÉ, F. Spi-1 and Fli-1 Directly Activate Common Target Genes Involved in Ribosome Biogenesis in Friend Erythroleukemic Cells. **Molecular and cellular biology**, 29(10):2852-2864, 2009.

JUNAID, M, A.; CLARK, G, M.; PULLARKAT, R, K. A lysosomal pepstatin-insensitive proteinase as a novel biomarker for breast carcinoma. **The International Journal of Biological Markers**, 15(2), 129-134, 2000.

KAMINKÁS, M.; CISZEWKI, T.; LOPACKA-SZATAN, K.; MIOTLA, P.; STAROSLAWSKA, E. Breast cancer risk factors. **Prz Menopauzalny**; 14(3): 196-202, 2015.

KANG, G, J.; PARK, M, K.; BYUN, H, J.; KIM, H, J.; KIM, E, J.; YU, L.; KIM, B.; SHIM, J, G.; LEE, H.; LEE, C, H. SARNP, a participant in mRNA splicing and export, negatively regulates E-cadherin expression via interaction with pinin. **J Cell Physiol**. 1–13. 2019.

KANG, J.; BRAJANOVSKI, N.; CHAN, K, T.; XUAN, J.; PEARSON, R, B.; SANJI, E. Ribosomal proteins and human diseases: molecular mechanisms and targeted therapy. **Signal Transduction and Targeted Therapy**. 6(323), 2021.

KANHAIYA, K.; CZEIZLER, E.; GRATIE, C.; PETRE, I. Controlling Directed Protein Interaction Networks in Cancer. *Scientific Reports*, 7: 10327, 2017.
KIM, J.; GAO, L.; TAN, K. Multi-Analyte Network Markers for Tumor Prognosis. **Plos ONE**, 7(12): e52973, 2012.

KOOMEN, J, M.; HAURA, E, B.; BEPLER, G.; SUTPHEN, R.; REMILY-WOOD, E, R.; BENSON, K.; HUSSEIN, M.; HAZLEHURST, L, A.; YEATMAN, T, J.; HILDRETH, L, T.; SELLERS, T, A.; JACOBSEN, P, B.; FENSTERMACHER, D, A.; DALTON, W, S. Proteomic Contributions to Personalized Cancer Care. **Molecular & Cellular Proteomics**, 7.10, 2008.

KOUROU, K.; EXARCHOS, T, P.; EXARCHOS, K, P.; KARAMOUZIS, M, V.; FOTIADIS, D, I. Machine learning applications in cancer prognosis and prediction. **Computational and Structural Biotechnology Journal**, 13, 8–17, 2015.

KOVACS, I, A.; LUCK, K.; SPIROHN, K.; WANG, Y.; POLLIS, C.; SCHLABACH, S.; BIAN, W.; KIM, D, K.; KISHORE, N.; HAO, T.; CALDERWOOD, M, A.; VIDAL, M.; BARABÁSI, A, L. Network-based prediction of protein interactions. **Nature communications**,10:1240, 2019.

KRUG, K.; JAEHNIG, E, J.; SATPATHY, S.; ELLIS, M, J.; GILLETTE, M, A.; Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. **Cell**, 183, 1436–1456, 2020.

LANG, G, T.; JIANG, Y, Z.; SHI, J, X.; YANG, F.; LI, X, G.; PEI, Y, C.; ZHANG, C, H.; MA, D.; XIAO, Y.; HU, P, C.; WANG, H.; YANG, Y, S.; GUO, L, W.; LU, X, X.; XUE, M, Z.; WANG, P.; CAO, A, W.; LING, W.; WNAG, Z, W.; YU, K, D.; DI, G, H.; LI, D., Q.; WANG, Y, J.; YU, Y.; SHI, L, M.; HU, X.; HUANG, W.; SHAO, Z, M. Characterization of the genomic landscape and actionable mutations in Chinese breast cancers by clinical sequencing. **Nature communications**, 11:5679, 2020.

LAWRENCE, M, S.; STOJANOV, P.; MERMEL, C, H.; ROBINSON, J, T.; GARRAWAY, L, A.; GOLUB, T, R.; MEYERSON, M.; GABRIEL, S, B.; LANDER, E, S.; GETZ, G. Discovery and saturation analysis of cancer genes across 21 tumour types. **Nature**, 505(495), 2014.

LEE, S, B.; BOSE, S.; AHN, S, H.; SON, B, H.; KO, B, S.; KIM, H, J.; CHUNG, I, J.; KIM, J.; LEE, W.; KO, M, S.; LEE, K.; CHANG, S.; PARK, H, S.; LEE, J, W.; KIM, D, C. Breast cancer diagnosis by analysis of serum N-glycans using MALDI-TOF mass spectroscopy. **Plos One**, 2021.

LI, H.; ZHANG, H.; HUANG, G.; BING, Z.; LUO, H.; XU, D.; LIU, J.; AN, X. Loss of RPS27a expression regulates cell cycle, apoptosis, and proliferation via the RPS27a-RPL11-MDM2-p53 pathway in lung adenocarcinoma cells. **Research Square**, 2020.

LI, J.; GUAN, X.; FAN, Z.; CHING, L, M.; LI, Y.; WANG, Z.; CAO, W, M.; LIU, D, X. Non-invasive biomarkers for early detection of breast cancer. **Cancers**, 12, 2767, 2020.

LI, R.; QU, H.; WANG, S.; WEI, J.; ZHANG, L.; MA, R.; LU, J.; ZHU, J.; DHONG, W, D.; KIA, Z. GDCRNATools: an R/Bioconductor package for integrative analysis of lncRNA, miRNA, and mRNA data in GDC. **Bioinformatics**, 2018.

LI, Y.; HUANG, J.; SUN, J.; XIANG, S.; YANG, D.; YING, X.; LI, H.; REN, G. The transcription levels and prognostic values of seven proteasome alpha subunits in human cancers. **Oncotarget**, 8(3), 4501-4519, 2017.

LIBERZON, A.; SUBRAMANIAN, A.; PINCHBACK, R.; PABLO, T.; MESIROV, J, P. Molecular signatures database (MSigDB) 3.0. **Bioinformatics**, 27(12), 1739–1740, 2011.

LIN, Z.; PENG, R.; SUN, Y.; ZHANG, L.; ZHANG, Z. Identification of ribosomal protein family in triple-negative breast cancer by bioinformatics analysis. **Bioscience Reports**, 41, 2021.

LIN, Z.; PENG, R.; SUN, Y.; ZHANG, L.; ZHANG, Z. Identification of ribosomal protein family in triple-negative breast cancer by bioinformatics analysis. **Bioscience Reports**, 41, 2021.

LIU, R, A.; GRAHAM, K.; GLUBRECHT, D, D.; GERMAIN, D, R.; MACHEY, J, R.; GODBOUT, R. Association of FABP5 expression with poor survival in triple-negative breast cancer. **AJP**, 178(3), 2011.

LOBO, M, D, P.; MORENO, F, B, M, B.; SOUZA, G, H, M, F, S.; VERDE, S, M, M, L.; MOREIRA, R, A.; MONTEIRO-MOREIRA, A, C, O. Label-Free Proteome analysis of Plasma from Patients with Breast cancer: stage-specific Protein expression. **Front. Oncol.** 7:14, 2017.

LÓPEZ-CORTÉS, A.; CABRERA-ANDRADE, A.; VÁZQUEZ-NAYA, J, M.; PAXOS, A.; GONZÁLES-DÍAZ, H.; PAZ-Y-MIÑO, C.; GUERRERO, S.; PÉREZ-CASTILLO, Y.; TEJERA, E.; MUNTEANU, C, R. Prediction of breast cancer proteins involved in immunotherapy, metastasis, and RnA-binding using molecular descriptors and artificial neural networks. **Scientific Reports**,10:8515, 2020.

LUALDI, M.; FASANO, M. Statistical analysis of proteomics data: A review on feature selection. **Journal of Proteomics**;198(30), 18-26, 2019.

LUCK, K.; SHEYNKMAN, G, M.; ZHANG, I.; VIDAL, M. Proteome-scale human interactomics. **Trends Biochem Sci**; 42(5): 342–354, 2017.

MACKLIN, A.; KHAN, S.; KISLINGER, T. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. **Clin Proteom**, 17:17, 2020.

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

MANNELL, A. The role of Ki-67 in breast cancer. **SAJS**, 54(2), 2016.
MCVEIGH, T, P.; KERIN, M, J. Clinical use of the Oncotype DX genomic test to guide treatment decisions for patients with invasive breast cancer. **Breast Cancer - Targets and Therapy**:9 393–400, 2017.

MERWE, D, E, V, der.; OIKONOMOPOULOU, K.; MARSHALL, J.; DIAMANDIS, E, P. Mass spectrometry: uncovering the cancer proteome for diagnostics. **Advances in cancer research**, 2007.

MILIOLI, H, H.; SOUSA, K, S.; KAVISKI, R.; OLIVEIRA, N, C, D, S.; URBAN, C, A.; LIMA, R, S.; CAVALLI, I, J.; RIBEIRO, E, M, S, F. Comparative Proteomics of Primary Breast Carcinomas and Lymph Node Metastases Outlining Markers of Tumor Invasion. **Cancer genomics & proteomics**, 12: 89-102 (2015).

MIYAKE, T.; NAKAYAMA, T.; NAOI, Y.; TAMAMOTO, N.; OTANI, Y.; KIM, S, J.; SHIMAZU, K.; SHIMOMURA, A.; MARUYAMA, N.; TAMAKI, Y.; NOGUCHI, S. GSTP1 expression predicts poor pathological complete response to neoadjuvant chemotherapy in ER-negative breast cancer. **Cancer science**, 103(5), 913-920, 2012.

MOLINIE, N.; GAUTREAU, A. The Arp2/3 regulatory system and its deregulation in cancer. **Physiol Rev**, 98: 215–238, 2018.
MOMENIMOVAHED, Z.; SALEHINIA, H. Epidemiological characteristics of and risk factors for breast cancer in the world. **Breast Cancer - Targets and Therapy**:11 151–164, 2019.

MUTHU, M.; CHERIYAN, V, T.; RISHI, A, K. CARP-1 / CCAR1: A biphasic regulator of cancer cell growth and apoptosis. **Oncotarget**, 6(9), 6490-6510, 2015.

NARLA, A.; EBERT, B, L. Ribosomopathies: human disorders of ribosome dysfunction. **Blood**, 15(16), 2010.

NASCIMENTO, R, G.; OTONI, K, N. Histological and molecular classification of breast cancer: what do we know? **Mastology**;30: e20200024, 2020.

NEVE, R,M.; CHIN ,K.; FRIDLAND, J.; YEH, J.; BAEHNER, F, L.; FEVR, T.; CLARK, L.; BAYANI, N.; COPPE, J, P.; TONG, F.; SPEED, T.; SPELLMAN, P, T.; DEVRIES, S.; LAPUK, A.; WANG, N,J.; KUO, W, L.; STILWELL, J, L.; PINKEL, D.; ALBERTSON, D, G.; WALDMAN, F, M.; MCCORMICK, F.; DICKSON, R, B.; JOHNSON, M,D.; LIPPMAN, M.; ETHIER, S.; GAZDAR, A.; GRAY, J, W.; A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. **Cancer cell** 10, 515–527, 2006.

NIEMINEN, T, T.; O'DONOHUE, M, F.; WU, Y.; LOHI, H.; SCHERER, S, W.; PATERSON, A, D.; ELLONEN, P.; ABDEL-RAHMAN, W, M.; CALO, S.; MECKLIN, J, P.; JARVINEN, H, J.; GLEIZES, P, E.; PELTOMAKI, P. Germline Mutation of RPS20, Encoding a Ribosomal Protein, Causes Predisposition to Hereditary Nonpolyposis Colorectal Carcinoma Without DNA Mismatch Repair Deficiency. **Gastroenterology**, 147:595-598, 2014.

NISHIMURA, Y.; KOMATSU, S.; ICHIKAWA, D.; NAGATA, H.; HIRAJIMA, S.; TAKESHIDA, H.; KAWAGUCHI, T.; ARITA, T.; KONISHI, H.; KASHIMOTO, K.; SHIOZAKI, A.; FUJIWARA, H.; OKAMOTO, K.; TSUDA, H.; OTSUJI, E. Overexpression of YWHAZ relates to tumor cell proliferation and malignant outcome of gastric carcinoma. **British Journal of Cancer**, 108, 1324–1331, 2013.

NOTTROTT, S.; SIMARD, M, J.; RICHTER, J, D. Human let-7a miRNA blocks protein production on actively translating polyribosomes. **Nature structural & molecular biology**, 13(12), 2006.

ORSOLIC, I.; BURSAC, S.; JURADA, D.; HOFMAN, I, D.; DEMBIC, Z.; BARTEK, J.; MIHALEK, I.; VOLAREVIC, S. Cancer-associated mutations in the ribosomal protein L5 gene dysregulate the HDM2/p53-mediated ribosome biogenesis checkpoint. **Oncogene**, 2020.

OSZ, A.; LANCZKY, A.; GYORFFY, B. Survival analysis in breast cancer using proteomic data from four independent datasets. PREPRINT, 2020.

OTTEWELL, P, D.; O'DONNELL, L.; HOLEN, I. Molecular alterations that drive breast cancer metastasis to bone. **BoneKey Reports** 4(643), 2015.

PALAMIDESSI, A.; FRITTOLI, E.; GARRÉ, M.; FARETTA, M.; MIONE, M.; TESTA, I.; DIASPRO, A.; LANZETTI, L.; SCITA, G.; FIORE, P, D, D.; Endocytic

trafficking of RAC is required for the spatial restriction of signaling in cell migration. **Cell** 134, 135–147, 2008.

PANIS, C.; PIZZATTI, L.; HERRERA, A, C.; CORRÊA, S.; BINATO, E.; ABDELHAY, E. Label-Free Proteomic Analysis of Breast Cancer Molecular Subtypes. **J. Proteome Res.**, 13, 4752–4772, 2014.

PAPALAZAROU, V.; MACHESKY, L, M. The cell pushes back: The Arp2/3 complex is a key orchestrator of cellular responses to environmental forces. **Current Opinion in Cell Biology**, 68:37–44, 2021.

PARISE, C, A.; CAGGIANO, V. Breast Cancer Survival Defined by the ER/PR/HER2 Subtypes and a Surrogate Classification according to Tumor Grade and Immunohistochemical Biomarkers. **Journal of Cancer Epidemiology**, 2014.

PARSONS, J.; FRANCAVILLA, C. ‘Omics approaches to explore the breast cancer landscape. **Cell Dev. Biol.** 7:395, 2020.

PEREIRA, B.; CHIN, S, F.; RUEDA, O, M.; VOLLAN, H, K, M.; PROVENZANO, E.; BARDWELL, H, A.; PUGH, M.; JONES, L.; RUSSELL, R.; SAMMUT, S, J.; TSUI, D, W, Y.; LIU, B.; DAWSON, S, J.; ABRAHAM, J.; NORTHEN, H.; PEDEN, J, F.; MUKHERJEE, A.; TURASHBILI, G.; GREEM A, R.; MCKINNEY, D, R.; OLOUMI, A.; SHAH, S.; ROSENFELD, N.; MURPHY, L.; BENTLEY, D, R.; ELLIS, I, O.; PURUSHOTHAM, A.; PINDER, S, E.; BORRESEN-DALE, A, L.; EARL, H, M.; PHAROAH, P, D.; ROSS, M, T.; APARICIO, S.; CALDAS, C. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. **Nature communications**, 7, 2016.

PEROU, C, M.; SORLIE, T.; EISEN, M, B.; de RIJN, M, V.; JEFFREY, S, S.; REES, C, A.; POLLACK, J, R.; ROSS, D, T.; JOHNSEN, H.; AKSLEN, L, A.; FLUGE, O.; PERGAMENSHIKOV, A.; WILLIAMS, C.; ZHU, S, X.; LONNING, P, E.; BORRESEN-DALE, A, L.; BROWN, P, O.; BOTSTEIN, D. Molecular portraits of human breast tumors. **Nature**, 406, 2000.

PLAYDON, M, C.; ZIEGLER, R, G.; SOMPSON, J, N.; STOLZENBERG-SOLOMON, R.; THOMPSON, H, J.; IRWIN, M, L.; MAYNE, S, T.; HOOVER, R, N.; MOORE, S, C. Nutritional metabolomics and breast cancer risk in a prospective study. **Am J Clin Nutr**;106: 637–49, 2017.

POLYAK, A. Breast cancer: origins and evolution. **J. Clin. Invest.** 117:3155–3163, 2007.

POPAY, T, M.; WANG, J.; ADAMS, C, M.; HOWARD, G, C.; CODREANU, S, G.; SHERROD, S, D.; MCLEAN, J, A.; THOMAS, L, R.; LOREY, S, .; MACHIDA, Y, J.; WEISSMILLER, A, M.; EISCHEN, C, M.; LIU, Q.; TANSEY,

W, P. MYC regulates ribosome biogenesis and mitochondrial gene expression programs through its interaction with host cell factor-1. **eLife**, 10, 2019.

POWELL, C, A.; NASSER, M, W.; ZHAO, H.; WOCHNA, J, C.; ZHANG, X.; SHAPIRO, S.; SHILO, K.; GANJU, R, G. Fatty acid binding protein 5 promotes metastatic potential of triple negative breast cancer cells through enhancing epidermal growth factor receptor stability. **Oncotarget**, 6(8), 2015.

PRIETO, D, A. Mass spectrometry in cancer biomarker research: a case for immunodepletion of abundant blood-derived proteins from clinical tissue specimens. **Biomark Med**; 8(2): 269–286, 2014.

RAMDAS, P.; RADHAKRISHNAN, A, K.; SANI, A, A, A.; KUMARI, M.; RAO, J, S, A.; RAHMAN, P, S, A. Advancing the role of gamma-tocotrienol as proteasomes inhibitor: a quantitative proteomic analysis of MDA-MB-231 human breast cancer cells. **Biomolecules**, 10, 19; 2020.

RAO, S.; CAI, K, Q.; STADANLICK, J, E.; GREENBERG-KUSHNIR, N.; SOLANKI-PATEL, N.; LEE, S, Y.; FAHL, S, P.; TESTA, J, R.; WIEST, D, L. Ribosomal Protein Rpl22 Controls the Dissemination of T-cell Lymphoma. **Cancer research**, 76(11), 2016.

RAO, S.; LEE, S, Y.; GUTIERREZ, A.; PERRIGOUE, J.; THAPA, R, J.; TU, Z.; JEFFERS, J, R.; RHODES, M.; ANDERSON, S.; ORAVECZ, T.; HUNGER, S, P.; TIMAKHOV, R, A.; ZHANG, R.; BALACHANDRAN, S.; ZAMBETTI, G, P.; TESTA, J, R.; LOOK, A, T.; WIEST, D, L. Inactivation of ribosomal protein L22 promotes transformation by induction of the stemness factor, Lin28B. **Blood**, 120: 3764-3773, 2013.

RICHARD, A, L.; ECKHARDT, M.; KROGAN, N, J. Mass spectrometry-based protein–protein interaction networks for the study of human diseases. **Mol Syst Biol**.17: e8792, 2021.

RISHA, Y.; MIMIC, Z.; GHOBADLOO, S, M.; BEREZOVSKI, M, V. The proteomic analysis of breast cell line exosomes reveals disease patterns and potential biomarkers. **Scientific Reports**, 10:13572, 2020.

ROUILLARD, A.; GUNDERSEN, G, W.; FERNANDEZ, N.; WANG, Z.; MONTEIRO, C, D.; MCDERMOTT, M, G.; MA1YAN, A. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. **Database**, 2016.

SAHARA, S.; AOTO, M.; EGUCHI, Y.; IMAMOTO, N.; YONEDA, Y.; TSUJIMOTO, Y. Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation. **Nature**, 401, 168-173, 1999.

SAHIN, F.; QIU, W.; WILENTZ, R, E.; IACOBUZIO-DONAHUE, C, A.; GROSMARK, A.; SU, G, H. RPL38, FOSL1, and UPP1 Are Predominantly Expressed in the Pancreatic Ductal Epithelium. **Pancreas**, 30(2):158-167, 2005.

SCHMID, S, L. Reciprocal regulation of signaling and endocytosis: Implications for the evolving cancer cell. **J. Cell Biol.**, 216(9), 2623–2632, 2016.

SERRANO-CARBAJAL, E, A.; ESPINAL-ENRIQUEZ, J.; HERNÁNDEZ-LEMUS, E. Targeting metabolic deregulation landscapes in breast cancer subtypes. **Front. Oncol.** 10:97, 2020.

SHANNON, P.; MARKIEL, A.; OZIER, O.; BALIGA, N, S.; WANG, J, T.; RAMAGE, D.; AMIN, N, SCHWIKOWSKI, B.; IDEKER, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. **Genome research**, 12:2498-2504, 2003.

SIBIO, A.; ABRIATA, G.; BUFFA, S.; VINIEGRA, M.; FORMAN, D.; SIERRA, M, S. Etiology of breast cancer (c50) in central and south america. In: cancer in central and south America. Lyon: International agency for research on cancer. Disponível em: http://www-dep.iarc.fr/CSU_resources.htm, 2016.

SIMPSON, P, T.; REIS-FILHO, J, S.; GALE, T.; LAKHANI, S, R. Molecular evolution of breast cancer. **J Pathol**; 205: 248–254, 2005.

SOLIMAN, H.; SHAH, V.; SRKALOVIC, G.; MAHTANI, R.; LEVINE, E.; MAVROMATIS, B.; SRINIVASIAH, J.; KASSAR, M.; GABORDI, R.; QAMAR, R.; UNTCH, S.; KLING, H, M.; TREECE, T.; AUDEN, W. MammaPrint guides treatment decisions in breast Cancer: Results of the IMPACT trial. **BMC Cancer**; 20:81, 2020.

SØRLIE, T.; TIBSHIRANI, R.; PARKER, J.; HASTIE, T.; MARRON, J, S.; NOBEL, S.; DENG, S.; JOHNSEN, H.; PESICH, R.; GEISLER, S.; DEMETER, J.; PEROU, C, M.; LØNNING, P, E.; BROWN, P, O.; BØRRESEN-DALE, A, L.; BOTSTEIN, D. Repeated observation of breast tumor subtypes in independent gene expression data sets. **PNAS**, 100(14), 8418-8423, 2003.

SOYSAL, S.; PTP1B expression is an independent positive prognostic factor in human breast cancer. **Breast Cancer Res Treat.** 137(2), 2013.

STEGNER, H, R.; BRORKEN, W.; BAHNSEN, J. Relation Between Receptor Status and Ultrastructure in Breast Cancer: E+P+ versus E-P-. **Path. Res. Pract.**, 181:377-381, 1986.

STINGL, J. Estrogen and Progesterone in normal mammary gland development and in cancer. **Horm canc**, 2:85–90, 2010.

STRICKER, T, P.; BROWN, C, D.; BANDLAMUDI, C.; MCNERNEY, M.; KITTLER, R.; MONTOYA, V.; PETERSON, A.; GROSSMAN, R.; WHITE, K, P; Robust stratification of breast cancer subtypes using differential patterns of transcript isoform expression. **Plos Genet**, 13(3), 2017.

SU, H.; XU, T.; GANAPATHY, S.; SHADAFAN, M.; LONG, M.; HUANG, T, H, M.; THOMPSON, I.; YUAN, Z, M. Elevated snoRNA biogenesis is essential in breast cancer. **Oncogene**, 1-11, 2013.

SUÁREZ-ARROYO, I, J.; FELIZ-MOSQUEA, Y, R.; PÉREZ-LASPIUR, J.; ARJU, R.; GIASHUDDIN, S.; MALDONADO, -MARTÍNEZ, G.; CUBANO, L, A.; SCHNEIDER, R, J.; MARTÍNEZ-MONTEMAYOR, M, M, M. The proteome signature of the inflammatory breast cancer plasma membrane identifies novel molecular markers of disease. **Am J Cancer Res**, 6(8):1720-1740, 2016.

SUBRAMANIAN, A.; TAMAYO, P.; MOOTHA, V, K.; MUKHERJEE, S.; ELBERT, B, J.; GILLETTE, M, A.; PAULOVICH, A.; POMEROY, S,. GOLUB, T, R.; LANDER E, S.; MERISOV, J, P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. **PNAS**, 102 (43):15545-155505, 2005.

SULIMA, S, O.; HOFMAN, I, J, H.; KEERSMAECKER, K.; DINMAN, J, D. How ribosomes translate cancer. **Cancer Discovery**, 7(10):1069-1087, 2018.

SUNG, H.; FERLAY, J.; SIEGEL, R, L.; LAVERSANNE, M.; SOERJOMATARAM, I.; JEMAL, A.; BRAY, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. **CA Cancer**, 71:209-249, 2021.

SUZUKI, R.; ORSINI, N.; SAJI, S.; KEY, T, J.; WOLK, A.; Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status—A meta-analysis. **Int. J. Cancer**: 124, 698–712, 2009.

TANG, Z.; DING, S.; HUANG, H.; LUO, P.; QING, B.; ZHANG, S.; TANG, R. HDAC1 triggers the proliferation and migration of breast cancer cells via upregulation of interleukin-8. **Biological Chemistry** 'Just Accepted' paper, 2018.

TELKERSEN, T.; PERNEMALM, M.; GROMOV, P.; BORRENSEN-DALE, A, L; KROGH, A.; HAAKENSEN, V, D.; LETHIO, J.; PAPALEO, E.; GROMOVA, I. High-throughput proteomics of breast cancer interstitial fluid: identification of tumor subtype-specific serologically relevant biomarkers. **Molecular Oncology** 15, 429–461, 2021.

TENGA, M, J.; LAZAR, L, M.; A Proteomic Snapshot of Breast Cancer Cell Cycle: The G1/S Transition Point. **Proteomics**. 13(1): 48–60, 2013.

TWEEDIE, S.; BRASCHI, B.; GRAY, K.; JONES, T, E, M.; SEAL, R, L.; YATES, B.; BRUFORD, E, A. Genenames.org: the HGNC and VGNC resources in 2021. **Nucleic Acids Research**, 49:939-946, 2021.

TWIMAN, R, M. Proteomics. **Elsevier**, 2021.

TYANOVA, S.; COX, J. Perseus: A Bioinformatics Platform for Integrative Analysis of Proteomics Data in Cancer Research. **Cancer Systems Biology: Methods and Protocols, Methods in Molecular Biology**, 1711, 2018.

VASAIKAR, S, V.; STRAUB, P.; WANG, J.; ZHANG, B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. **Nucleic Acids Research**, Vol. 46, 2018.

VENTICINQUE, L.; JAMISON, K, .; MERUELO, D. Interactions Between Laminin Receptor and the Cytoskeleton During Translation and Cell Motility. **PlosOne**, 6(1), 2011.

WALLDEN, B.; STORHOFF, J.; NIELSEN, T.; DOWIDAR, N.; SCHAPER, C.; FERREE, S.; LIU, S.; LEUNG, S.; GEISS, G.; SNIDER, J.; VICKERY, T.; DAVIES, S, R.; MARDIS, E, R.; GNANT, M.; SESTAK, I.; ELLIS, M, J.; PEROU, C, M.; BERNARD, P, S.; PARKER, J, S.; Development and verification of the PAM50-based Prosigna breast cancer gene signature assay. **BMC Medical Genomics**; 8:54, 2015.

WALTER, W.; SÁNCHEZ-CABO, R.; RICOTE, M. GOplot: an R package for visually combining expression data with functional analysis. **Bioinformatics**, 31(17), 2912–2914, 2015.

WANG, C, Q, LI, Y.; HUANG, B, F.; ZHAO, Y, M.; YUAN, H.; GOU, D.; SU, C, M, U, G, N, WNAG, Q.; LONG, T.; WNAG, Y.; TANG, C, H.; LI, X. EGFR conjunct FSCN1 as a novel therapeutic strategy in triple negative breast cancer. **Scientific Reports**, 15654, 2017.

WANG, C, Q.; TANG, C, H.; CHANG, H, T.; LI, X, N.; ZHAO, Y, M.; SU, N, X, X.; ZENG, Y.; DU, Z.; WANG, Y.; HUANG, B, F. Fascin -1 as a novel diagnostic marker of triple-negative breast cancer. **Cancer Medicine**; 5(8):1983–1988, 2015.

WANG, H.; XIE, B.; KONG, Y.; TAO, Y.; YANG, G.; GAO, M.; XU, H.; ZHAN, F.; SHI, J.; ZHANG, Y.; WU, X. Overexpression of RPS27a contributes to enhanced chemoresistance of CML cells to imatinib by the transactivated STAT3. **Oncotarget**, 7(14), 2015.

WANG, W.; NAG, S.; ZHANG, X.; WANG, M, H.; WANG, H.; ZHOU, J.; ZHANG, R. Ribosomal Proteins and Human Diseases: Pathogenesis, Molecular

Mechanisms, and Therapeutic Implications. **Medicinal Research Reviews**, 35(2):225-285, 2015.

WANG, Z.; HOU, L.; QI, Z.; SUN, J.; GAO, W.; MENG, J.; WANG, Y.; SUN, H.; GU, H.; XIN, Y.; GUO, Z.; YANG, G. Small Ribosomal Protein Subunit S7 Suppresses Ovarian Tumorigenesis through Regulation of the PI3K/AKT and MAPK Pathways. **PlosOne**, 8(11), 2013.

WARNER, J, R.; MCLINTOSH, K, B. How Common Are Extraribosomal Functions of Ribosomal Proteins? **Molecular cell**, 34, 2009.

WAZIR, U.; AHMED, M, H.; BRIDGER, J, M.; HARVEY, A.; KIANG, W, G.; SHARMA, A, K.; MOKBEL, K. The Clinicopathological Significance of Lamin A/C, Lamin B1 and Lamin B receptor mRNA Expression in Human Breast Cancer. **Cellular & Molecular Biology Letters**, 2013.

WEINSTEIN, J, N.; COLLISSON, E, A.; MILLS, G, B.; SHAW, K, M.; OZENBERGER, B, A.; ELLROTT, K.; SHMULEVICH, I.; SANDER, C.; STUART, J, M.; CANCER GENOME ATLAS RESEARCH NETWORK. The Cancer Genome Atlas Pan-Cancer Analysis Project. **Nature genetics**, 45(10):1113-1120, 2013.

WISNIEWSKI, J, R.; GOUGAZ, F, Z. Fast and sensitive total protein and peptide assays for proteomic analysis. **Anal. Chem**, 87, 4110–4116, 2015.

WU, J.; CHEN, J.; ZHANG, Y.; DAI, W.; LIN, X.; YANG, X.; CHEN, M. ADH2 is downregulated by methylation and acts as a novel biomarker for breast carcinoma prognosis. **Annals of Clinical & Laboratory Science**, 51(1), 2021.

WU, Q.; GOU, Y.; WANG, Q.; JIN, H.; CUI, L.; ZHANG, Y.; HE, L.; WANG, J.; NIE, Y.; SHI, Y.; FAN, D. Downregulation of RPL6 by siRNA Inhibits Proliferation and Cell Cycle Progression of Human Gastric Cancer Cell Lines. **PlosOne**, 6(10), 2011.

WU, X.; HASAN, M, A.; CHEN, J, Y. Pathway and network analysis in proteomics. **Journal of Theoretical Biology**, 362, 44–52, 2014.
YAN, T, T.; FU, X, L.; LI, J.; BIAN, Y, N.; LIU, D, J.; HUA, R.; REN, L, L.; LI, C, T.; SUN, Y, W.; CHEN, H, Y.; FANG, J, Y.; HONG, J. Downregulation of RPL15 may predict poor survival and associate with tumor progression in pancreatic ductal adenocarcinoma. **Oncotarget**, 6(35):37028-37042, 2015.

YANG, S.; CUI, J.; YANG, Y.; LIU, Z.; YAN, H.; TANG, C.; WANG, H.; QIN, H.; LI, X.; WANG, W.; HUANG, Y.; GAO, H. Over-expressed RPL34 promotes malignant proliferation of non-small cell lung cancer cells. **Gene**, 576:421-428, 2016.

- YANOVICH, G.; AGMON, H.; HAREL, M.; SONNENBLICK, A.; PERETZ, T.; GEIGER, T. Clinical Proteomics of Breast Cancer Reveals a Novel Layer of Breast Cancer Classification. **Cancer Res**, 78(20), 6001-6010, 2018.
- YERSAL, O.; BARUTCA, S. Biological subtypes of breast cancer: Prognostic and therapeutic implications. **World J Clin Oncol**, 10; 5(3): 412-424, 2014.
- ZAMANIAN-AZODI, M.; RAZAEI-TAVIRANI, R.; RAHMAT-RAD, S.; HASANZADEH, H.; TAVINANI, M, R.; SEYYDI, S, S. Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer. **Gastroenterol Hepatol**; 8(3):215-224, 2015.
- ZHANG, L, Q.; YANG, H, Y.; YANG, S, Q.; WANG, Y.; CHEN, X, J.; LU, H, S.; ZHAO, L, P. CNDP2 acts as an activator for human ovarian cancer growth and metastasis via the PI3K/AKT pathway. **Technology in Cancer Research & Treatment**, 18: 1-10, 2019.
- ZHANG, M.; LEE, A, V.; ROSEN, J, M. The Cellular Origin and Evolution of Breast Cancer. **Cold Spring Harb Perspect Med**;7:a027128, 2017.
- ZHAO, W.; LI, X.; NIAN, W.; WANG, J.; WANG, X.; SUN, L.; ZHU, Y.; TONG, Z. Ribosome Proteins Represented by RPL27A Mark the Development and Metastasis of Triple-Negative Breast Cancer in Mouse and Human. **Frontiers in Cell and Developmental Biology**, 9, 2021.
- ZHAO, Y.; YANG, X.; XU, X.; ZHANG, L.; YU, H.; MIAO, Z.; LI, D.; WANG, S. Deubiquitinase PSMD7 regulates cell fate and is associated with disease progression in breast cancer. **Am J Transl Res**, 12(9):5433-5448, 2020.
- ZHENG, L.; DAI, H.; ZHOU, N.; LI, M.; SINGH, P.; QIU, J.; TSARK, W.; HUANG, Q.; KERNSTINE, K.; ZHANG, X.; LIN, D.; SHEN, B. Fen1 mutations result in autoimmunity, chronic inflammation and cancers. **Nature medicine**, 13(7), 821-819, 2007.
- ZHU, Y.; GUO, M.; ZHANG, L.; XU, T.; WANG, L, XU, G. Biomarker triplet NAMPT/VEGF/HER2 as a de novo detection panel for the diagnosis and prognosis of human breast cancer. **Oncology reports** 35: 454-462, 2016.
- ZIANG, Y, Z.; MA, D.; SUO, C.; SHI, L.; HUANG, W.; SHAO, Z, M. Genomic and transcriptomic landscape of triple negative breast cancers: Subtypes and treatment strategies. **Cancer Cell** 35, 428–440, 2019.
- ZONNEVILLE, J.; WONG, V.; LIMOGE, M.; NIKIFOROV, M.; BAKIN, A, V. TAK1 signaling regulates p53 through a mechanism involving ribosomal stress. **Scientific reports**, 10(2517), 2020.