UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau.

Guilherme Pauperio Lanfredi

Ribeirão Preto
2022

Guilherme Pauperio Lanfredi

Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau.

VERSÃO CORRIGIDA

Tese apresentada ao Programa de Pós-Graduação em Bioquímica como parte dos requisitos para obtenção do título de Doutor em Ciências, área de concentração Bioquímica.

Versão corrigida da tese de doutorado apresentada ao programa de pós-graduação em Bioquímica em 23/06/2022. A versão original encontra-se disponível no departamento de Bioquímica da Faculdade de Medicina de Ribeirão Preto.

Orientador: Vitor Marcel Faça

Ribeirão Preto/SP

2022

Guilherme Pauperio Lanfredi

Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau.

Tese apresentada ao Programa de Pós-Graduação em Bioquímica como parte dos requisitos para obtenção do título de Doutor em Ciências, área de concentração Bioquímica.

Orientador: Vitor Marcel Faça

Ribeirão Preto/SP

2022

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

Lanfredi, Guilherme Pauperio.

Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau. Ribeirão Preto, 2022.

112 p. : il. ; 30 cm

Tese apresentada à Faculdade de Medicina de Ribeirão Preto/USP. Área de concentração: Bioquímica.

Orientador: Faça, Vitor Marcel.

1. Proteômica. 2. Espectrometria de Massas. 3. Assinatura molecular. 4. Câncer de ovário.

FOLHA DE APROVAÇÃO

GUILHERME PAUPERIO LANFREDI

Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau.

Tese apresentada ao Programa de Pós-Graduação em Bioquímica como parte dos requisitos para obtenção do título de Doutor em Ciências, área de concentração Bioquímica.

Aprovado em ____/____/____
Orientador:
Prof. Dr. Vitor Marcel Faça Instituição: FMRP/USP

Assinatura:_____

Banca Examinadora:

Prof. Dr.:_____Instituição:_____

Julgamento:_____Assinatura:_____

Prof. Dr(a).:_____Instituição:_____

Julgamento:_____Assinatura:_____

Prof. Dr(a).:_____Instituição:_____

Julgamento:_____Assinatura:_____

Agradecimentos

Agradeço à minha família que sempre me apoiou, incentivou e contribuiu para a conquista dos meus objetivos, avós, tios, primos, irmã e principalmente meus pais, fundamentais em todas as etapas.

À minha parceira Luciana, fonte de cuidado e de inspiração constante em todos os aspectos, encorajando minha dedicação.

Ao professor e orientador Dr. Vitor Marcel Faça pelas inúmeras oportunidades, orientação e ensinamentos, visando nosso desenvolvimento profissional em meio às adversidades da pesquisa nacional.

À amiga pessoal e especialista Ma. Ana Paula Masson pela ajuda dentro e fora do laboratório, auxiliando nos experimentos e na vida.

Aos funcionários do departamento de Bioquímica da FMRP pelo suporte, principalmente Silvia pela dedicação, Ivone, Rose e Ana, pelo empenho e, muitas vezes, paciência, na manutenção do programa de pós-graduação.

Aos amigos e colegas de laboratório, principalmente ao atual time do laboratório de proteômica do câncer, Virgínia, John, Alessandra, Guilherme, Beatriz, Maria Luiza, Izadora e tantos outros que passaram por lá, contribuindo para meu desenvolvimento científico.

Ao professor Dr. Francisco J. C. dos Reis e doutores Aline, Mariana, Carolina e Germano por iniciarem e proporcionarem meios para que eu seguisse esta linha de pesquisa.

Às agências de fomento CAPES, CNPq e FAPESP pelo auxílio financeiro e suporte ao laboratório.

Resumo

Lanfredi, G.P. Análise do secretoma de células de câncer de ovário Caov-3 induzidas ao processo de transição epitélio-mesenquimal por fator de crescimento EGF e sua associação com assinaturas de tumores de alto grau. 112 p. Tese de doutorado – Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo. 2022.

O câncer de ovário (CaOV) é a mais letal das neoplasias ginecológicas, em particular, o câncer de ovário seroso de alto grau (HGSC) tem um prognóstico ruim, principalmente pela ocorrência de metástases. A reprogramação molecular da célula, ocorrida na transição epitelial-mesenquimal (EMT) representa um modelo para estágios avançados de progressão tumoral. A EMT promove mudanças celulares importantes na adesão celular e comunicação intercelular, que são particularmente dependentes de sinais do microambiente tumoral. Considerando a importância da comunicação celular durante EMT e metástase, aqui analisamos as alterações induzidas pelo fator de crescimento epidérmico (EGF) no secretoma da linhagem celular de câncer de ovário Caov-3. Usando uma combinação de GEL-LC-MS/MS e marcação metabólica isotópica estável (SILAC), identificamos candidatos superexpressos durante a EMT como partida na identificação de proteínas relevantes para HGSC. Com base em bancos de dados públicos, compreendendo estudos dos consórcios TCGA e CPTAC, nossas proteínas candidatas foram comparadas e priorizadas para análise posterior. As análises ressaltaram que vários dos candidatos foram associados a vesículas extracelulares, importantes para a comunicação célula-célula e metástase. Além disso, associamos as proteínas candidatas com dados de expressão gênica e proteica do subtipo mesenquimal de câncer de ovário para obtenção de assinaturas para classificação de tumores. Com estas assinaturas moleculares relevantes para a agressividade do câncer ovário, apoiadas por dados de expressão de proteínas e genes, desenvolvemos métodos de proteômica dirigida para avaliar amostras clínicas individuais de CaOV. As informações quantitativas obtidas para peptídeos, representativos de proteínas, foi capaz de agrupar HGSC entre outros tipos de tumor e seus diferentes estágios. Nosso estudo destacou a riqueza do secretoma para revelar proteínas relevantes para HGSC, que poderiam ser usadas em estudos futuros com maiores coortes como uma potencial assinatura de estratificação para tumores de ovário, a fim de orientar a conduta clínica para terapia do paciente.

Palavras-chave: Proteômica, Espectrometria de Massas, MRM, Assinatura molecular, Câncer de ovário.

Abstract

Lanfredi, G.P. Analysis of Caov-3 ovarian cancer cells secretome induced to epithelial-mesenchymal transition process by EGF and its association with high-grade tumor signatures. 112 p. PhD thesis – Ribeirão Preto Medical School, University of São Paulo. 2022.

Ovarian cancer (CaOV) is the most lethal of gynecological neoplasms, in particular, high-grade serous ovarian cancer (HGSC) has a poor prognosis, mainly due to the occurrence of metastases. The molecular reprogramming of the cells at the epithelial-mesenchymal transition (EMT) represents a model for advanced stages of tumor progression. EMT promotes important cellular changes in cell adhesion and intercellular communication, which are particularly dependent on signals from the tumor microenvironment. Considering the importance of cell communication during EMT and metastasis, here we analyze the changes induced by epidermal growth factor (EGF) in the secretome of the ovarian cancer cell line Caov-3. Using a combination of GEL-LC-MS/MS and stable isotopic metabolic labeling (SILAC), we identified candidates overexpressed during EMT as a starting point for identifying HGSC-relevant proteins. Based on public databases, comprising studies from the TCGA and CPTAC consortia, our candidate proteins were compared and prioritized for further analysis. The analyzes highlighted that several target candidates were associated with extracellular vesicles, important for cell-cell communication and metastasis. In addition, we associated target proteins with gene and protein expression data of the mesenchymal molecular subtype of ovarian cancer to obtain signatures for tumor classification. With these molecular signatures relevant to ovarian cancer aggressiveness, supported by protein and gene expression data, we developed targeted proteomics methods to evaluate individual clinical samples of CaOV. The quantitative information obtained for peptides, representative of proteins, was able to group HGSC among other tumor types and their different stages. Our study highlighted the richness of the secretome to reveal HGSC-relevant proteins, which could be used in future studies with larger cohorts as a potential stratification signature for ovarian tumors in order to guide clinical management for patient therapy.

Keywords: Proteomics, Mass spectrometry, MRM, Molecular signature, Ovarian cancer.

Sumário

1 - Introdução..... 12
1.1 - Câncer de Ovário..... 12
1.2 - A Transição Epitélio Mesenquimal – EMT..... 16
1.3 - Busca por biomarcadores e assinaturas no câncer de ovário..... 20
1.4 - Abordagens proteômicas para o estudo do câncer..... 24
2 – Objetivo..... 28
2.1 - Objetivo Geral..... 28
2.2 - Objetivos específicos..... 28
3 – Métodos..... 29
3.1 - Cultivo de células de câncer de ovário e indução da EMT..... 30
3.1.1 - Fracionamento subcelular..... 31
3.1.2 - Obtenção e preparação da amostra de secretoma de células Caov-3..... 33
3.2 - Análise Gel-LC-MS/MS do secretoma..... 35
3.3 - Validação de marcadores da EMT por Western blotting..... 36
3.5 - Preparo das amostras para análise proteômica dirigida..... 38
3.6 – Métodos de proteômica dirigida..... 38
3.6.1 - Método de proteômica dirigida para avaliação de marcadores epiteliais e mesenquimais utilizando peptídeos sintéticos..... 38
3.6.2 - Método de proteômica dirigida para painel de proteínas relevantes à assinatura mesenquimal..... 40
3.7 - Comparação entre modelo de progressão tumoral e dados do CPTAC..... 42
3.7.1 - Agrupamento hierárquico..... 42
3.8 - Análise de componentes principais (PCA) e k-means..... 43
3.8.1 – Construção de classificador com dados do CPTAC (análise supervisionada) 44
4 – Resultados..... 45
4.1 - Considerações iniciais..... 45
4.2- Painel de biomarcadores da EMT..... 46
4.3 - Avaliação do modelo de Indução da EMT..... 51
4.4 - Identificação de proteínas no secretoma de células Caov-3 induzidas à EMT. .55
4.5 - Construção de uma assinatura molecular proteo-genômica do subtipo mesenquimal para o Câncer de Ovário..... 61
4.6 - Análise de amostras clínicas para a assinatura proteômica do subtipo mesenquimal..... 66
4.7 - Evolução da assinatura proteômica do subtipo mesenquimal para o câncer de ovário..... 71
4.8 - Análise por agrupamento e classificadores das assinaturas moleculares..... 78
5 -Discussão..... 84
5.1 - Identificação de proteínas no secretoma de células Caov-3..... 84
5.2 - Painéis de assinatura mesenquimal partindo de proteínas superexpressas no secretoma..... 90
5.3 - Análise por agrupamento e classificadores das assinaturas moleculares..... 95
5.4 – Considerações finais..... 97
6 – Conclusões..... 99
7 – Figuras suplementares..... 100

8 – Referências bibliográficas.....102

Lista de figuras

Figura 1: Origens do câncer de ovário.....15
Figura 2: Estágios da transição Epitélio-mesenquimal.....18
Figura 3: Metodologias utilizadas para obtenção das assinaturas mesenquimais.....29
Figura 4: Otimização de energia de colisão de peptídeos representativos de biomarcadores clássicos da EMT.....49
Figura 5: Imagens representativas das alterações morfológicas de células da linhagem Caov-3 tratadas com fator de crescimento EGF.....52
Figura 6: Quantificação relativa por Western blotting e LC-MRM de biomarcadores clássicos da EMT em cultura de células Caov-3 tratadas com fator de crescimento.. 54
Figura 7: Proteínas identificadas por LC-MS/MS no secretoma de células Caov-3.. 57
Figura 8: Diagrama de Venn com proteínas superexpressas. Identificadas em cada subfração celular obtida de cultura de células da linhagem Caov-3 e secretoma.….59
Figura 9: Rede de interação obtida pelo *software* String contendo proteínas superexpressas no secretoma obtido de células Caov-3.....60
Figura 10: Diagrama de sobreposição de proteínas superexpressas no secretoma com transcritos mais expressos no subtipo mesenquimal (TCGA).....61
Figura 11: Método de proteômica dirigida para painel de proteínas da assinatura mesenquimal.....63
Figura 12: Resultados individuais por proteína do método de proteômica dirigida desenvolvido para assinatura mesenquimal.....67
Figura 13: Avaliação da assinatura molecular de expressão gênica de proteínas em amostras tumorais de pacientes.....70
Figura 14: Identificação de proteínas por assinaturas proteômicas de subtipos moleculares.....72
Figura 15: Diagrama de sobreposição de proteínas superexpressas no secretoma para determinação com proteínas mais expressas no subtipo mesenquimal (CPTAC).....74
Figura 16: Avaliação da assinatura molecular proteômica referente ao subtipo mesenquimal em amostras tumorais de pacientes com câncer de ovário.....76
Figura 17: Projeção de PCA dos diferentes subtipos moleculares do câncer de ovário de alto grau. a)JHU. b)PNNL.....79
Figura 18: Projeção em PCA de amostras do CPTAC classificadas por algoritmo Random trees.....81
Figura 19: Associação de proteínas selecionadas identificadas no secretoma com funções do reactoma.....88
Figura 20: Rede de interações montada com proteínas selecionadas identificadas no secretoma.....89
Figura 21: Curvas de sobrevivência por subtipo molecular.....91

Figura S1: Avaliação molecular da indução da EMT……100

Figura S2: Análises do conjunto de proteínas superexpressas nas células Caov-3 após indução da EMT……101

Lista de tabelas

Tabela 1: Exemplos de marcadores relacionados ao câncer de ovário já aprovados pelo FDA....22
Tabela 2: Otimização de energia de colisão para painel de biomarcadores clássicos da EMT....48
Tabela 3: Parâmetros utilizados no monitoramento de biomarcadores clássicos da EMT....50
Tabela 4: Painel de peptídeos para o monitoramento de proteínas relevantes à assinatura mesenquimal....64
Tabela 5: Identificação de amostras de tumores de pacientes submetidos à análise por MRM....66
Tabela 6: Variáveis utilizadas pelo algoritmo Random trees na composição do modelo probabilístico....82
Tabela 7: Principais proteínas utilizadas superexpressas no secretoma que compuseram as assinaturas....94

Lista de Abreviaturas

CaOV – Câncer de ovário
CCD – Dispositivo de carga acoplada
CPTAC – Clinical Proteomic Tumor Analysis Consortium
CT - experimento controle
Ctrl – experimento controle
DDA – Análise dependente de dados
DIA – Análise independente de dados
DNA – Ácido desoxirribonucleico
DTT - Ditiotreitol
ECM – Matriz extracelular
EDTA - Ácido etilenodiamino tetra-acético
EGF – Fator de crescimento epidermal
EGFR – Receptor de fator de crescimento epidermal
EMT – Transição Epitélio-mesenquimal
FBS – Soro fetal bovino
FDA – Food and Drug Administration
FDR – Taxa de falsa descoberta
FSH – Hormônio folículo estimulante
HGSC – Câncer de ovário seroso de alto grau
IAA - Iodoacetoamida
IGF – Fator de crescimento semelhante à Insulina
JHU – Johns Hopkins University
LC-MS – Cromatógrafo líquido – espectrômetro de massas
MRM – Monitoramento de reações múltiplas
mRNA – ácido ribonucleico mensageiro
NK – *Natural Killer*
PBS – Solução tampão de fosfato
PCA - Análise de componentes principais
PNNL – Pacific Northwestern National Laboratory
PTM – Modificação pós-traducional
PVDF - Fluoreto de polivinilideno
Q-TOF – Quadrupolo – analisador por tempo de voo
ROMA – The Risk of Ovarian Malignancy Algorithm
SDS-PAGE – Eletroforese em gel de poliacrilamida com dodecil-sulfato de sódio
SILAC – Marcação com isótopos estáveis com aminoácidos em cultura de células
SRM – Monitoramento de reação única
SSRCalc – *Sequence specific retention calculator*
TCGA – The Cancer Genome Atlas Program
TGF – Fator de crescimento transformador
TOF - Analisador por tempo de voo

1 - Introdução

1.1 - Câncer de Ovário

O câncer de ovário (CaOV) é a segunda neoplasia maligna ginecológica mais comum, sendo a mais letal e oitava mais incidente no mundo [1,2]. No Brasil, em 2020 estimaram-se 6650 novos casos, enquanto a mortalidade em 2019 foi de 4123 casos registrados [3], representando assim uma preocupação relevante para a pesquisa e órgãos de saúde.

A letalidade do câncer de ovário se deve principalmente ao seu diagnóstico tardio, em graus de difícil tratamento, com o estabelecimento de metástases (aproximadamente 75% em estágios III ou IV), em que a taxa de sobrevivência é de 5 anos [4]. Em contraste, se diagnosticadas nos estágios iniciais, quando a doença ainda se limita aos ovários, 90% das pacientes podem ser curadas, com taxa de sobrevivência de aproximadamente 95%, incentivando o estudo de formas para o diagnóstico precoce e compreensão dos mecanismos moleculares envolvidos na doença [5].

O diagnóstico tardio que culmina na alta letalidade se deve ao fato dos sintomas iniciais do câncer de ovário serem vagos e inespecíficos, podendo indicar outras patologias de menor preocupação, tornando a triagem por este critério difícil e não recomendada [6,7].

Outro fator que contribui para a letalidade do CaOV é a falta de dados disponíveis sobre a incidência de subtipos dos tumores no nível diagnóstico, o que pode ser determinante para o correto diagnóstico, prognóstico e direcionamento à terapia. Os carcinomas ovarianos se manifestam de cinco formas principais: carcinoma seroso de alto grau, os mais frequentes e prevalentes; carcinoma endometrial; de células claras; de

baixo grau e carcinoma mucinoso. Havendo ainda a classificação em relação aos estágios destes, que contempla tamanho, comprometimento de linfonodos e metástases, sendo este último o fator determinante para a atribuição do mais alto grau [8].

A heterogeneidade observada nos tumores também se reflete em suas origens, onde as neoplasias de ovário mais comuns relacionadas aos tumores serosos de alto grau permanece em debate, pois apesar de se apresentarem como tumores ovarianos, estes raramente são encontrados contidos nos ovários e sim compreendendo adjacências[9].

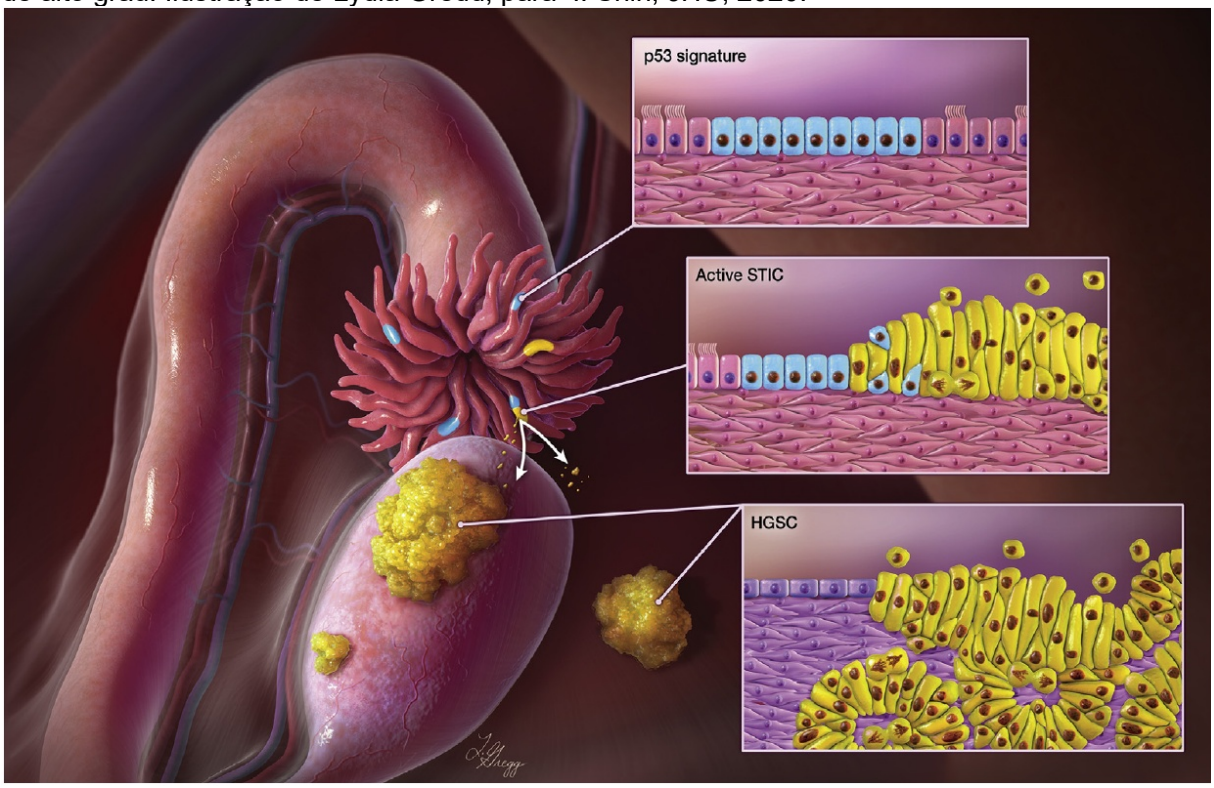
Há duas teorias principais sobre esta origem. Na primeira, os tumores seriam originados na superfície do epitélio ovariano, ou mesotélio pélvico, quando lesões iniciais originam inclusões cistiformes, formados por células com potencial de diferenciação estimuladas pelo microambiente pró-inflamatório. A segunda teoria propõe a origem partindo das trompas de falópio, ou especificamente de suas fímbrias, ou seja, uma origem extra ovariana. Estas duas propostas são suportadas por achados e observações igualmente válidas e não mutuamente exclusivas, porém de difícil confirmação pelo baixo nível de diagnóstico e acompanhamento precoce [10].

Alguns indícios levantados para as hipóteses de origem dos tumores vêm de suas assinaturas moleculares e consequentemente também de suas mutações no nível genômico, sendo a mais observada nos tumores de alto grau a mutação no gene TP53, acompanhada de várias outras (BRCA1, BRCA2, NF1, RB1, etc), nem sempre comum entre outros tipos de câncer e geralmente ligadas à instabilidade genômica e comprometimento do reparo de DNA. Já, outros tipos de CaOV de menor grau apresentam outras mutações: Endometrióide e Células claras (PTN, PIK3CA, ARID1A, CTNNB1 para Endometrióide); Mucinoso e Baixo grau (KRAS mútuo, HER2 e BRAF respectivamente), caracterizando assim enfermidades diferentes desde sua origem [11]. A

ocorrência de mutações também é associada à diferentes níveis de risco, exemplo disso é a presença de variantes do gene BRCA1 em pacientes, que corresponde a chance do desenvolvimento de câncer de ovário de 49%, já variantes no gene BRCA2, 21% para o mesmo câncer, até os 80 anos [12].

As variantes características de cada patologia manifestam uma vulnerabilidade alvo para as ações provenientes da desregulação do microambiente tumoral, este fator é relevante no surgimento e manutenção dos tumores. No ovário os eventos responsáveis pela iniciação dos tumores de alto grau envolvem as interações entre células epiteliais, imunes e do estroma. As células epiteliais, especialmente as localizadas no fim das fímbrias são diretamente expostas aos fluídos foliculares da ovulação e estes contém vários fatores considerados carcinogênicos por conter substâncias genotóxicas, como espécies reativas de oxigênio. Assim, com a repetição desses episódios de contínuo dano e reparo ao DNA, mutações e mudanças epigenéticas com potencial cancerígeno podem surgir, como esquematizado na figura 1. Este processo carcinogênico pode ainda ser promovido por fatores de crescimento, como IGF (Fator de Crescimento Insulínico) [13].

Figura 1: **Origens do câncer de ovário.** Células prejudicadas por mutações no gene TP53, lesões originadas nas trompas de falópio ou carcinomas intra-epiteliais serosos tubários (STIC) somados ao microambiente pró-inflamatório são possíveis origens do carcinoma ovariano seroso de alto grau. Ilustração de Lydia Gredd, para I. Shih, JHU, 2020.



A associação do CaOV têm sido verificada com diversos fatores de crescimento, estes achados ainda acompanham a identificação de importantes mecanismos da sinalização celular. Na família dos fatores TGFβ (Fator de Crescimento Transformador), TGFβ1, produzido principalmente por macrófagos encontrados em ascites, são associados à reincidência precoce, enquanto a expressão dos fatores TGFβ2 e TGFβ3 são fortes indicadores de mau prognóstico clínico quando expresso pelas células tumorais [14]. Outro fator de crescimento que surge de forma relevante na correlação e mecanismos ligados ao câncer de ovário é o EGF (Fator de Crescimento Epidermal), responsável por diversas vias de ativação intracelular, como Ras/Raf/MAPK, Jak/Stat e AKT/PI3K, participa da regulação da proliferação e migração celular. Os receptores de EGF, cuja

ativação é associada à pior prognóstico e aumento da malignidade, são superexpressos na taxa de 30 a 98% dos tumores epiteliais de ovário [15].

Apesar da boa resposta com a aplicação de terapias antitumorais baseadas em platina em pacientes de tumores de ovário de alto grau, há evidências de que a ativação de receptores de EGF ocorra em resposta ao quimioterápico. Assim, culminando na resistência a este fármaco, devido a possível reversão ligada às recorrentes mutações em BRCA1 e 2 que restaura os mecanismos de reparo de danos ao DNA, aumentando a eficiência destes. Além do efeito direto da supra regulação da via de sinalização Akt e promoção da sobrevivência celular [11].

Por fim, o comportamento mais agressivo dos tumores e ocorrência de metástase é fortemente associado à superexpressão da família EGFR desencadeando a transição epitelial para mesenquimal (EMT), uma das várias características do complexo desenvolvimento do câncer. Esta promove alterações celulares e do microambiente resultando em invasão e migração, favorecendo a metástase [16]. Assim, a associação dos mecanismos desencadeados pelo processo de transição epitélio mesenquimal relacionados ao EGF com a evolução e metástase do câncer de ovário podem abrir novas oportunidades para a inibição da progressão tumoral, ou mesmo indicar possíveis assinaturas moleculares que indiquem o potencial de agressividade do tumor.

1.2 - A Transição Epitélio Mesenquimal – EMT

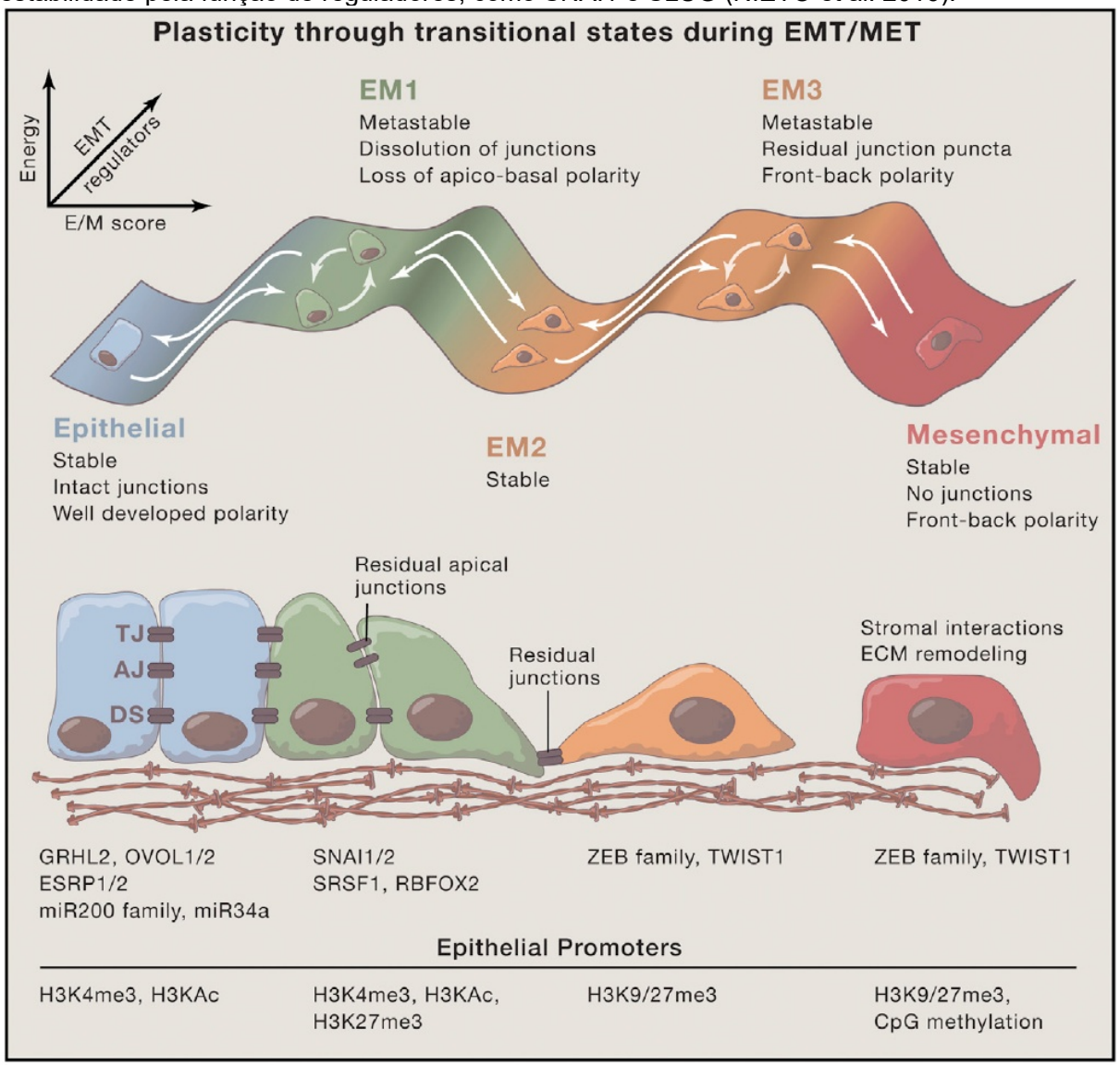
As propriedades invasivas de células tumorais resultam da reprogramação e transformação intensa durante a progressão do câncer. Estas podem ser observadas morfologicamente, como a perda da polaridade ápico-basal, perda de adesão e aquisição de mobilidade, sendo estes os passos que levam posteriormente ao processo da cascata de invasão metastática habilitando estas células a deixar o tumor primário. Estes

processos estão associados ao fenômeno denominado transição epitélio-mesenquimal (EMT) [17].

Particularmente, o padrão de disseminação de células metastáticas de tumores provenientes do ovário difere do comumente observado em outros tipos de tumores epiteliais sólidos, cuja disseminação ocorre predominantemente pela corrente sanguínea. As células do câncer de ovário se disseminam partindo do tumor primário por meio do fluido peritoneal para outros locais propícios para evolução das metástases, incluindo aí a própria cavidade peritoneal. Por sua vez a cavidade peritoneal pode proporcionar um ambiente favorável para a fixação das células metastáticas por ser composta de componentes da matriz extracelular e fibroblastos revestidos por uma camada de células mesoteliais como única barreira [18].

O processo da EMT é um mecanismo biológico regulado, preliminar para a migração de células derivadas do epitélio na morfogênese e notado no desenvolvimento embriológico normal. Entretanto, é bem estabelecido na literatura que a EMT também pode desempenhar função importante na metástase, registrado por levar à consequências semelhantes à progressão tumoral em vários modelos de câncer. Neste processo células de fenótipo epitelial deixam algumas de suas principais características e adquirem outras, mais voltadas ao fenótipo mesenquimal. Assim, o processo de EMT não significa em si um fenômeno patogênico. Os programas da EMT habilitam a migração celular e invasão pela membrana basal e matriz extracelular para grupos multicelulares induzindo uma migração coletiva, incluindo células com fenótipo epitelial. Atualmente é mais expressiva a compreensão do caráter de “transição”, pois há células com taxa parcial de indução, misturando características dos fenótipos epitelial e mesenquimal (figura 2), sendo estas essenciais à disseminação e invasão [19].

Figura 2: **Estágios da transição Epitélio-mesenquimal.** Gradiente de estágios possíveis do processo de transição epitélio-mesenquimal, ilustrados pelos fenômenos da perda de adesão célula-célula, polaridade ápico-basal e remodelamento, havendo conformações de maior estabilidade pela função de reguladores, como SNAI1 e SLUG (NIETO et al. 2016).



A regulação da EMT, assim como as alterações que levam à metástase, em modelos tradicionais aparece como eventos tardios na progressão tumoral. Entretanto, estudos apontam que esses fenômenos podem ocorrer mais cedo, e até mesmo de forma evidente na fase de lesões pré-neoplásicas [20]. Uma série de fatores de transcrição com potencial de indução são diretamente relacionadas à EMT, notadamente proteínas SNAI1, SLUG, TWIST e ZEB1 já foram explorados em detalhe na literatura. Outros como ZEB2,

FOXC2, PRRX1 são reportados, mas suas respectivas funções específicas no câncer são menos documentadas. A variedade específica de fatores de transcrição relacionados à EMT expressos depende de características como contexto celular, nível de diferenciação, origem celular e grau em que se encontra a própria transição [21].

A EMT também é regulada de outras formas e as modificações pós-traducionais são relevantes em várias etapas, incluindo a degradação de proteínas diretamente relacionadas ao processo mediadas pela ubiquitinação e a fosforilação das proteínas TWIST1 e SNAI1 por MAP quinases e GSK3b, respectivamente. O esforço para construção de uma mapa que determine a compreensão da causalidade das assinaturas da EMT se mostra possível de acordo com as demonstrações da ocorrência de assinaturas comuns entre diferentes tipos de câncer [22]. Ainda, os complexos processos da transição durante a EMT acontecem com o ajuste de vias metabólicas para responder corretamente à grande demanda bioenergética requerida , este ajuste vem diretamente dos fatores de transcrição que regulam todo o processo, inclusive o ajuste das vias metabólicas, que são um componente inevitável da mudança fenotípica da EMT. No entanto, o metabolismo também é descrito como um regulador da adaptabilidade celular, e a reprogramação deste pode ser outro modulador para a EMT [23].

Também é notória a competência de células induzidas à EMT na supressão imune, prevenindo ataques contra os tumores através da expressão de citocinas, quimiocinas e fatores de crescimento para repressão da atividade citotóxica de células T, citolítica de células NK e apresentação de antígenos por células dendríticas. Esta competência também é relacionada à resistência a drogas como observado para, os carcinomas com maior proporção de células mesenquimais, que são menos sensíveis à quimioterápicos, juntamente com aqueles com mais propriedades de células tronco [24].

Os estudos da EMT evidenciam contribuições desta à progressão tumoral que incluem aquisição de propriedades de células tronco, instabilidade genômica, invasão e disseminação, evasão da senescência e apoptose, sensibilidade às terapias e efeitos associados ao microambiente tumoral. Fazendo-se necessária a verificação de algumas destas alterações, avaliando seus resultados em células neoplásicas e normais. Também são necessários mais estudos para determinação precisa de como fenótipos celulares distintos associados a programas parciais de EMT contribuem para a evolução do câncer [19].

Além dos marcadores EMT clássicos envolvendo proteínas de adesão célula-célula e fatores de transcrição, outros marcadores potenciais podem estar envolvidos, dada a variação biológica. Portanto, visando condições patogênicas específicas, busca-se biomarcadores que desempenhem um papel importante no processo metastático, como será discutido nas próximas seções.

1.3 - Busca por biomarcadores e assinaturas no câncer de ovário.

Múltiplos biomarcadores de tumores, séricos, técnicas radiográficas e algoritmos de estratificação de risco têm sido o foco de extensa pesquisa no diagnóstico precoce do câncer de ovário. Pacientes com câncer de ovário em estágio inicial têm poucos sintomas e a progressão da doença pode estar relacionada com sintomas inespecíficos mais pronunciados, incluindo inchaço e dor abdominal. Como já discutido, tumores com maior probabilidade de serem detectados em um estágio inicial estão associados a melhor sobrevida devido à características moleculares e não apenas ao estágio de diagnóstico, em contrapartida uma grande proporção de HGSCs agressivos são descobertos quando já invadiram a cavidade peritoneal. Atualmente, biomarcadores proteicos de câncer de ovário são usados para monitorar a progressão da doença, e alguns biomarcadores

séricos de câncer de ovário também são usados como estratégias de triagem em populações de alto risco (tabela 1) [25].

Os níveis séricos de proteínas CA125 (Mucin-16), HE4 e o Risk of Ovarian Malignancy Algorithm (ROMA), combinação de CA125 e HE4, adquiridos por técnicas de imunoensaio são os marcadores mais utilizados na investigação diagnóstica de pacientes com suspeita de massa pélvica maligna. No entanto, a indicação obtida do monitoramento do marcador CA125 geralmente não é suficiente para iniciar o tratamento, a menos que o paciente seja sintomático ou esteja se inscrevendo em um ensaio clínico. Além disso, alguns pacientes com recorrência podem não apresentar níveis aumentados de CA125. Níveis de HE4 também estão sendo avaliados para o monitoramento da recorrência e progressão da doença. Atualmente não há marcadores aprovados na clínica para a predição da resposta ao tratamento e prognóstico, sendo os citados reservados para resposta ao tratamento, recorrência e malignidade de massa pélvica [26].

Outros testes multivariados foram desenvolvidos para o manejo clínico do câncer de ovário. Ova1 composto por 5 indicadores, compreendendo proteínas inflamatórias e de transporte; Transferrina, Microglobulina β-2, Apolipoproteína A-1, Transtirretina e CA125 indicando o risco para o câncer de ovário foi aprovado para pré-operatórios, trazendo um incremento em sensibilidade [27]. A segunda geração de testes multivariados mostrando esta tendência de avaliação de múltipla fontes, por exemplo, Overa combinando marcadores (Apolipoproteína A-1, HE4, Hormônio folículo estimulante (FSH), TRansferrina e CA125) mantendo a sensibilidade e aumentando a especificidade. É importante destacar que estes testes não são verdadeiramente diagnósticos e sim para triagem e referência, dois requisitos são necessários para sua indicação, uma massa confirmada por teste de imagem e a determinação de cirurgia pela presença de tumor [28].

Tabela 1: Exemplos de marcadores relacionados ao câncer de ovário já aprovados pelo FDA (BRADBURY et al. 2021).

Biomarker	Type	Specimen	Method	Clinical Use	Year
CA125	Protein	Serum, plasma	Immunoassay	Monitoring treatment response	1997
HE4	Protein	Serum	Immunoassay	Monitoring disease recurrence or progression	2008
ROMA (HE4 + CA125)	Protein	Serum	Immunoassay	Prediction of pelvic mass malignancy	2011
OVA1 Next Generation	Protein	Serum	Immunoassay	Prediction of pelvic mass malignancy	2009

CA125: cancer antigen 125; HE4: human epididymis protein 4; ROMA: Risk of Ovarian Malignancy Algorithm; PCR: polymerase chain reaction.

Avanços nas ciências ômicas (genômica, proteômica, metabolômica) têm potencial para desenvolver novas ferramentas para diagnóstico precoce, acompanhamento de tratamento e triagem. A análise proteômica detalhada de várias amostras biológicas (por exemplo, soro, plasma, urina, tecido) obtidas de pacientes com câncer contribui para estudos de tumorigênese, monitoramento terapêutico e desenvolvimento de novas terapias direcionadas. No entanto, a proposição de um novo painel de biomarcadores é um desafio, considerando que os biomarcadores devem atender requisitos, como aprimorar os métodos diagnósticos atualmente utilizados, aumentar sua sensibilidade e especificidade, corresponder ao estágio da doença e estar presentes em fluidos biológicos fáceis de obter [29].

Uma avaliação molecular abrangente do câncer é necessária para compreensão dos mecanismos da doença, para, principalmente, um melhor prognóstico e orientação do tratamento. O consórcio Cancer Genome Atlas (TCGA) realizou uma extensa caracterização genômica e transcriptômica dos carcinomas de ovário do tipo seroso de alto grau (HGSC) com o objetivo de obter um panorama genômico e contribuir para o desenvolvimento de terapias para esta neoplasia maligna altamente letal [30,31] . Embora os os resultados obtidos destas análises genômicas sejam substanciais, as funções codificadas no genoma são geralmente executados no nível das proteínas e muitas vezes ainda são modulados por modificações pós-traducionais (PTMs). Para obter uma avaliação mais abrangente do complexo fenótipo ovariano HGSC, o Clinical Proteomic

Tumor Analysis Consortium (CPTAC) buscou então realizar uma extensa caracterização proteômica quantitativa baseada em espectrometria de massa (MS) de tumores HGSC caracterizados pelo TCGA [31].

Os dados proteômicos gerados pelo CPTAC enriqueceram a compreensão da heterogeneidade do câncer de ovário e permitiram a classificação dos subgrupos HGSC com base primeiramente nas alterações genéticas usando o banco de dados do TCGA e posteriormente com a composição do proteoma pelos dados do próprio CPTAC.

Foram revelados, por exemplo, grupo de quinases com atividade elevada em tumores, incluindo quinases mitóticas e quinases dependentes de ciclina, e potenciais pontos de intervenção para a terapia do HGSC. O perfil proteômico dos tecidos tumorais permitiu a subtipagem de tumores. Por exemplo, o perfil proteômico global dos HGSCs correspondeu em parte aos subtipos transcriptômicos do TCGA, mesenquimal, proliferativo, imunorreativo e diferenciativo, superando as informações anteriores para identificar subtipos. Juntos, esses estudos identificaram alvos clinicamente tangíveis [32].

Os resultados do estudo do TCGA usando expressão de mRNA mostraram quatro subtipos de HGSC definidos como diferenciativo, imunorreativo, mesenquimal e proliferativo, já no estudo proteômico do CPTAC os subtipos foram determinados como diferenciativo, metabólico, mesenquimal, proliferativo e estromal.

Os quatro subtipos molecularmente distintos definidos pelo TCGA foram determinados com base em assinaturas moleculares. No entanto, esses subtipos moleculares ainda não possuem um significado biológico e clínico claro e confiável no HGSC. O subtipo mesenquimal de tumores foi relatado como tendo o pior prognóstico clínico dos quatro subtipos. Esses tumores foram definidos por baixa pureza tumoral, componentes estromais reativos e infiltração. Fenotipicamente, esses tumores exibem EMT aumentada, angiogênese, remodelação da matriz extracelular (ECM) e proteólise.

Consistente com esses achados, os tumores do subtipo mesenquimal exibiram aumento da expressão de genes HOX, bem como TGFβ aberrante associado ao estroma, conforme indicado pelas assinaturas de expressão gênica. As análises proteômicas pelo projeto CPTAC demonstraram também que esses tumores apresentam aumento da expressão de ECM e sinalização de citocinas [33].

Utilizar as proteínas como alvos clínicos apresentam vantagens. A especificidade do proteoma a tipos celulares e condições específicas é maior que o genoma, pois sofrem mais influência ambiental refletindo melhor estas condições. O nível de expressão proteico é consequência de vários processos celulares anteriores, abrangendo assim os resultados de suas alterações. Também sendo o maior efetore nas funções celulares [34].

As abordagens proteômicas revelaram milhares de potenciais biomarcadores de câncer, a maioria destas ainda precisam ser validadas. Desta forma, ainda apresentam desafios para a descoberta de biomarcadores, a padronização e otimização de protocolos de obtenção e processamento das amostras clínicas, bem como outros aspectos biológicos, como variabilidade da amostra ou heterogeneidade do tumor. Apesar disso, alguns biomarcadores têm sido implementados com sucesso na prática clínica, esses testes têm algumas limitações e sua sensibilidade e especificidade devem ser refinadas, especialmente para pacientes com tumores em estágios iniciais. Portanto, pode-se afirmar que a proteômica é uma ferramenta para desenvolvimento de novas ferramentas que poderão reduzir significativamente a mortalidade por câncer de ovário [29].

1.4 - Abordagens proteômicas para o estudo do câncer

A proteômica é definida como a caracterização de todo o conteúdo proteico de uma célula, tecido ou organismo sob condições determinadas, incluindo interações proteicas,

modificações e localização. O proteoma é muito dinâmico, com variantes de *splice*, modificações pós-traducionais (PTMs), muitas vezes com múltiplas modificações, que modulam sua função ou atividade. A combinação de medicina de precisão e proteômica aprimora a oncologia com maiores oportunidades para entender os complexos mecanismos de carcinogênese e alvos terapêuticos em nível molecular, revelando potenciais novos biomarcadores para detecção e acompanhamento, permitindo avaliação de tratamentos incluindo eficácia e toxicidade. [35,36].

As proteínas podem ser estudadas como entidades íntegras, abordagem denominada proteômica *top-down*, que tem a vantagem de, em princípio, reportar todas as mudanças que ocorrem na mesma molécula, permitindo a identificação precisa das formas proteicas. No entanto, a proteômica *bottom-up*, na qual peptídeos são gerados pela digestão enzimática das proteínas, é experimentalmente mais tratável e é o fluxo de trabalho de proteômica mais comum. Três abordagens principais são usadas em proteômica *bottom-up*: i) Proteômica de descoberta (ou *shotgun*) por análise dependente de dados (DDA), que visa obter uma cobertura completa do proteoma; ii) Proteômica dirigida, baseada em hipótese, usando de monitoramento de reação selecionadas (SRM), visando aquisição reprodutível, sensível e simplificada de subconjuntos de peptídeos conhecidos de interesse; iii) fragmentação múltipla de todos os peptídeos que eluem da coluna cromatográfica por análise independente de dados (DIA), visando gerar mapas de fragmentos de íons abrangentes para uma amostra [37]

A espectrometria de massa continua sendo a principal plataforma utilizada para análise proteômica. Nos últimos anos, houve muitos avanços na instrumentação, com melhorias na precisão, velocidade e resolução. Instrumentos MS, como Q-TOF, TOF/TOF e Orbitrap, foram desenvolvidos, permitindo uma exploração mais profunda de proteomas em análises cada vez mais curtas. Em particular, as técnicas de análise quantitativa

sensíveis amadureceram, como a marcação metabólica, que é uma alternativa aos métodos químicos para estudos *in vitro*. Ao adicionar aminoácidos marcados isotopicamente à cultura de células, as alterações do proteoma em diferentes estados podem ser detectadas, como alterações no nível de proteína e renovação de proteínas durante a diferenciação celular. A marcação de isótopos estáveis de aminoácidos em culturas de células (SILAC) torna possível a percepção de diferenças na abundância de proteínas entre amostras [38].

As análises não dirigidas demonstram limitação para etapas posteriores na busca por biomarcadores, seja pelas características dos peptídeos detectados, baixa reprodutibilidade ou resultados para proteínas de baixa abundância, devido à natureza estocástica da seleção de íons precursores baseada em abundância, que não pode garantir que os mesmos peptídeos sejam consistentemente detectados em todas as análises. Essa limitação indica que abordagens proteômicas quantitativas globais não direcionadas não são adequadas para a verificação e validação de candidatos a biomarcadores [34].

Em situações em que a marcação da amostra pode não ser desejável, como escassez de amostra, os métodos sem marcação e de monitoramento de reação oferecem alternativas adequadas. O fluxo de trabalho das abordagens com monitoramento de reação, por exemplo, monitoramento de reações múltiplas (MRM), pode ser considerada uma das estratégias quantitativas mais precisas, sendo mais adequada para análises direcionadas e ocasiões de disponibilidade de padrões internos. As estratégias de monitoramento de reação são frequentemente adaptadas a condições experimentais únicas, dessa forma, informações quantitativas absolutas e relativas podem ser determinadas sem marcação, eliminando problemas de amostra e complexidade espectral [39].

O desenvolvimento observado nas análises proteômicas também se deve ao acesso público de recursos de bioinformática para o processamento e interpretação de grandes volumes de dados gerados nos projetos. Protocolos de proteômica robustos, incluindo espectrômetros de massa acessíveis, estimulam aquisição de proteomas completos e precisos, em analogia ao que a introdução do sequenciamento de nova geração forneceu ao campo da genômica [37].

As abordagens não dirigidas são frequentemente empregadas para análises de câncer de ovário, pois proporcionam uma visão global do proteoma para seleção de vias para análises direcionadas adicionais [1]. Desta forma, os experimentos efetuados neste trabalho buscam demonstrar a utilidade das abordagens proteômicas quantitativas no avanço do conhecimento sobre o câncer de ovário.

2 – Objetivo

2.1 - Objetivo Geral

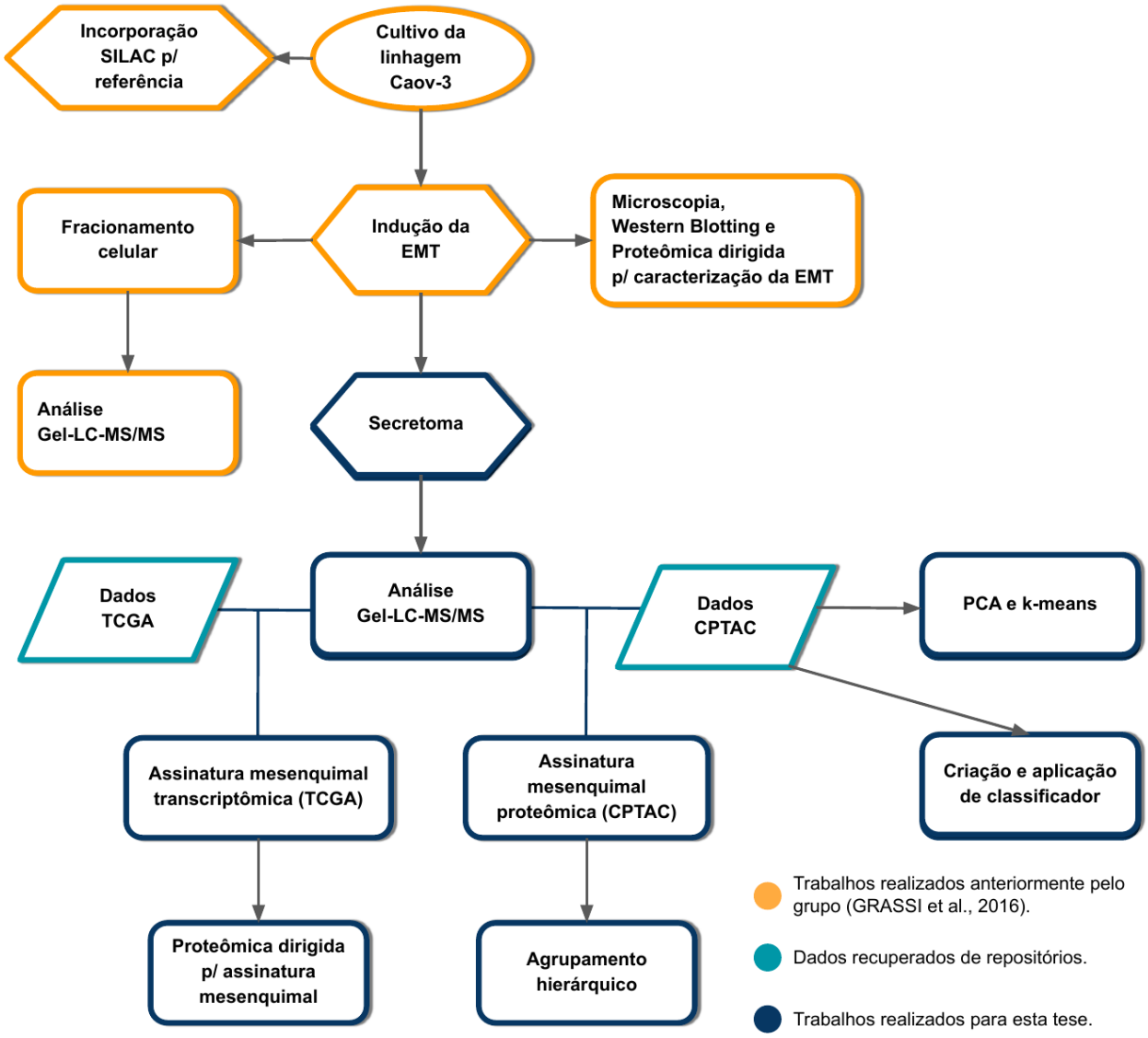
Estudar o modelo de transição epitélio mesenquimal para linhagens celulares representativas do câncer de ovário de origem epitelial (Caov-3), com foco no secretoma para avaliar sua contribuição como modelo de progressão tumoral para proposição de assinaturas do processo de progressão tumoral e metástase em tumores de ovário.

2.2 - Objetivos específicos

- Induzir o modelo de transição epitélio-mesenquimal e realizar sua avaliação no secretoma da linhagem de adenocarcinoma de ovário Caov-3.
- Realizar análise proteômica qualitativa e quantitativa do modelo, através de marcação isotópica.
- Identificar potenciais alvos superexpressos pela indução da EMT no secretoma, suas origens, interações e envolvimento em processos celulares.
- Integrar dados externos aos obtidos nas análises do secretoma do modelo e propor uma assinatura molecular para estudo.
- Desenvolver métodos de proteômica dirigida para monitoramento da assinatura molecular proposta.
- Refinar as assinaturas moleculares propostas através de ferramentas de bioinformática e baseando-se em dados gerados por outros grupos a fim de obter assinaturas relevantes ao subtipo mesenquimal do câncer de ovário.

3 – Métodos

Figura 3: **Metodologias utilizadas para obtenção das assinaturas mesenquimais.**



3.1 - Cultivo de células de câncer de ovário e indução da EMT

A cultura de células provenientes de adenocarcinoma de ovário humano Caov-3 (ATCC® HTB-75TM), foi mantida em garrafas de 75 cm^2 (Greiner Bio-One) com meio DMEM (Dulbecco's Modified Eagle Medium), suplementado com 10% de soro fetal bovino (FBS), 100 U/mL penicilina e 100 µg/mL estreptomicina (Life Technologies, Carlsbad, CA, USA).

As células foram mantidas em 37 °C em incubadora com atmosfera úmida e 5% de CO^2. O crescimento celular foi monitorado em microscópio e as células foram sub cultivadas a cada 3 dias ou dada a confluência de 70% da garrafa. No subcultivo o meio de cultura foi removido e a monocamada formada pelas células lavada com 5 mL de PBS. Em seguida, as células foram incubadas com 5 mL de solução de tripsina : PBS (1:1) a 37 °C, até seu desprendimento. Após isso, foram adicionados 5 mL de meio DMEM com FBS (10%), para conter a ação lítica da tripsina, e esta solução foi centrifugada a 200 x g por 5 min a 4 °C. Em seguida o sedimento celular foi ressuspenso em meio fresco com FBS e uma alíquota adicionada à garrafa de cultivo.

A cultura referência utilizou a abordagem SILAC com o aminoácido lisina com isótopos de carbono-13 ($^{13}C_6$-L-lysine-SILAC Advanced DMEM, Thermo Fisher Scientific) (0.1 mg/mL), suplementado da mesma forma, adicionalmente com 2 mM L-glutamina e 4.5 mg/mL D-glicose. Para incorporação da referência ($^{13}C_6$-L-lysine) as células foram expandidas por cinco passagens até atingirem uma taxa de ao menos 95% de proteínas

marcadas isotopicamente, verificada por espectrometria de massas, tendo como referência a razão m/z entre peptídeos da proteína Actina.

Para a indução da EMT 1,5 × 10^6 células foram semeadas em placas de 100 mm (Corning) em meio DMEM, suplementado com 10% de FBS, então lavadas em solução de tampão fosfato e seu meio substituído por DMEM, sem FBS, Após 24 h DMEM sem FBS contendo fator de crescimento epidermal (EGF) (10 ng/mL) (#236-EG-200, R&D Systems), foi adicionado e as células foram mantidas por 96 h, sendo o meio trocado em intervalos de 24 h. A viabilidade celular foi monitorada pelo ensaio MTT (3-(4,5-dimetiltiazol-2yl)-2,5-di- fenil brometo de tetrazolina).

3.1.1 - Fracionamento subcelular

Frações enriquecidas com proteínas da membrana, proteínas nucleares e proteínas do citosol da linhagem Caov-3 foram obtidas por lise celular e centrifugação diferencial.

As células (1 x 10^6)foram lavadas 2 vezes com tampão PBS gelado e lisadas com o auxílio de raspador de células, em 100 µL de tampão hipotônico, composto por HEPES (0,05 mol/L), NaCl (0,01 mol/L), $MgCl_2$ (0,005 mol/L), EDTA (0,0001 mol/L), NaF (0,001 mol/L), $Na_4P_2O_7.10H_2O$ (0,001 mol/L), Na_3VO_4 (0,001 mol/L) e coquetel inibidor de proteases 5% (P8340; Sigma) (v/v). Para a ruptura celular física, todo o volume coletado foi aspirado por seringa de 1 mL (20 vezes) e 10% do volume total foi separado para

quantificação das proteínas, pelo método colorimétrico de Bradford (Quick StartTM Bradford Protein Assay Kit, Bio-Rad).

Para homogeneização foram misturadas quantidades iguais de proteína entre as amostras, o conteúdo foi homogeneizado durante 30 segundos com o auxílio de homogeneizador de tecidos (Homomix D-130 - Biosystems, PR, Brasil) e centrifugado a 500 × g durante 20 min, a 4 °C.

O material foi transferido para microtubos para nova centrifugação a 16000 × g por 30 min a 4 °C. Já o pellet obtido (fração nuclear) foi ressuspenso em 100 µL de solução de extração de proteínas nucleares, contendo uréia 8M, CHAPS 2% (v/v), Na_3VO_4 0,001 mol/L e coquetel inibidor de proteases 5% (v/v), e armazenado a -80 °C.

O pellet proveniente desta centrifugação, corresponde à fração enriquecida em proteínas da membrana citoplasmática e o sobrenadante correspondente à fração citoplasmática. O pellet foi ressuspenso em 50 µL de tampão de extração de proteínas de membrana, contendo MES (0,025 mol/L), NaCl (0,15 mol/L), Triton X-100 2% (v/v), Na_3VO_4 (0,001 mol/L) e coquetel inibidor de proteases 5% (v/v), e incubado por 60 min, a 4 °C. Após centrifugação a 16000 × g por 20 min, à 4 °C, o sobrenadante correspondente à fração de membrana foi coletado.

O pellet nuclear foi descongelado e submetido à extração em 3 ciclos de banho com ultrassom e gelo, a 45 W, durante 5 min, agitado por 10 segundos e resfriado em banho de gelo por 5 min. Então as amostras foram centrifugadas a 20000 × g por 30 min, a 4 °C. O sobrenadante obtido foi utilizado como fração enriquecida em proteínas nucleares.

3.1.2 - Obtenção e preparação da amostra de secretoma de células Caov-3

Após cada período de 24 h de incubação das células Caov-3 em EGF, um volume equivalente a 5 mL dos meios de cultura foi reservado e mantido congelado para análise posterior, totalizando 20 mL no período de 96 h.

As frações de sobrenadante das culturas foram descongeladas e combinadas, adicionados inibidores de protease (0.5% v/v) (P8340; Sigma) e fosfatase (1 mM Na_3VO_4).

Os sobrenadantes condicionados foram centrifugados a 300 x g por 5 min., filtrados em membrana de poro 0,22 µm para remoção de células intactas e fragmentos insolúveis e concentrados em filtro molecular de 3 kDa (Vivaspin 2, GE Lifesciences).

Após concentração as amostras foram ressuspensas em 200 µL de uréia 8 M / 0.15 M Tris-HCl em pH = 8.8 e quantificadas em proteínas totais pelo método de Bradford (Bio-Rad).

O mesmo processo foi conduzido em células com marcação isotópica, como controle, que foram cultivadas sem EGF. Para análise, a mistura de frações independentes com a mesma quantidade de proteínas totais (12,5 µg) foi preparada, formando os pares: células Caov-3 incubadas com EGF + células Caov-3 SILAC e células Caov-3 sem EGF + células Caov-3 SILAC.

3.1.3 - Preparo de amostras para Análise Gel-LC-MS/MS do secretoma

Amostras referentes aos pares de secretomas de células induzidas e não induzidas à EMT e seus respectivos controles, compostos com isótopos ($^{13}C_6$-L-lisina) (células Caov-3 incubadas com EGF + células Caov-3 SILAC e células Caov-3 sem EGF + células Caov-3 SILAC) com 25 µg de proteínas totais, foram diluídas em 30 µL de tampão Laemmli, contendo ditiotreitol na concentração de 25 mM, fervidas por 5 min. para redução das pontes dissulfeto e alquiladas com iodoacetamida (135 mM) em temperatura ambiente por 30 minutos.

As amostras foram separadas por eletroforese utilizando gel comercial 10% SDS-PAGE (Bio-Rad, Hercules, CA, USA) e as raias correspondentes a cada experimento foram cortadas e divididas em 5 partes de aproximadamente 1,2 cm cada (Figura suplementar 1). Os fragmentos do gel foram lavados 3 vezes utilizando uma solução de bicarbonato de amônio 100 mM: acetonitrila (1:1) e digeridos por tripsina [16].

Os peptídeos trípticos foram extraídos por etapas, primeiro utilizado 200 µL de solução de 50% de acetonitrila em água com 0,1% de ácido fórmico e depois com 70% de acetonitrila e após isso secos em SpeedVac (Thermo Scientific). As misturas de peptídeos passaram por etapa de purificação em fase sólida ZipTip (Supelco, Sigma-Aldrich) de acordo com instruções do fabricante, e foram eluídas em solução de 52.5% acetonitrila / água mais 0.1% de ácido fórmico, secas novamente e ressuspensas em 15 µL de solução

água / acetonitrila 5% mais 0.1% de ácido fórmico para injeção no sistema cromatográfico.

3.2 - Análise Gel-LC-MS/MS do secretoma

As amostras foram analisadas em sistema nano cromatográfico (EASY-nLC II, Thermo Scientific) acoplado ao espectrômetro de massas LTQ Orbitrap Velos mass spectrometer (Thermo Scientific), disponível na Central Analítica do Instituto de Química de São Carlos sob coordenação do Prof. Emanuel Carrilho,utilizando coluna com resina C18 (Magic C18) de 25 cm com 75 µm de diâmetro interno (New Objective). O método cromatográfico utilizado teve fluxo de 300 nL/min por 90 min, variando em um gradiente linear de 5 até 40% de acetonitrila mais 0.1% de ácido fórmico.

A detecção do espectrômetro ocorreu em modo DDA (*data dependent analysis*) com seleção dos 10 íons mais abundantes, sendo duplamente ou triplamente carregados. O processamento dos dados adquiridos para identificação de peptídeos e atribuição de proteínas se deu pelo mecanismo de busca X!Tandem[1'], com Peptide Prophet[2'] e Protein Prophet[3'] para validação e proteoma referência do banco de dados Uniprot.org (Swiss Prot + TrEMBL) contendo 74782 entradas de proteínas em março de 2020.

Os parâmetros da busca incluíram peptídeos trípticos com até dois eventos de miss-cleavage, tolerância de massa de 0,5 Da, modificação fixa em cisteínas devido à alquilação, modificações variáveis pela oxidação em metioninas e modificação de lisina pela marcação isotópica. Apenas peptídeos com índice maior que 0,9 e menos de 20 ppm

de erro de massa foram considerados para quantificação de proteínas pelo algoritmo Q3, representando uma taxa de falsa descoberta de aproximadamente 1%.

3.3 - Validação de marcadores da EMT por Western blotting

As células Caov-3 coletadas após cultivo e separação do sobrenadante foram lavadas com PBS e lisadas com tampão de lise (Tris-HCl 20 mM, pH = 7.5, NaCl 150 mM, Na_2EDTA 1 mM, EGTA 1 mM, 1% Triton X-100, pirofosfato de sódio 2,5 mM, β-glicerofosfato 1 mM, Na_3VO_4 1 mM e Leupeptina 1 µg/mL), em 3 ciclos de banho ultrassônico por 5 min (mantido em gelo), esta solução foi centrifugada à 20000 × g por 30 min a 4 °C e o sobrenadante coletado.

As amostras tiveram as proteínas quantificadas pelo método de Bradford, submetidas à separação por SDS-PAGE e transferidas para membranas de PVDF (GE Lifesciences). As membranas foram bloqueadas com tampão de lavagem (5% de leite em pó sem gordura, Tris-HCl 25 mM, pH = 7.5, NaCl 0,5 M e 0,1% de detergente Tween-20) e incubadas com cada um dos anticorpos primários seguindo as orientações do fabricante, foram eles: Rabbit anti-N-Cadherin (#13116), rabbit anti-E-Cadherin (#3195) (rabbit anti-β-Catenin (#8480), rabbit anti-Vimentin (#5741), rabbit anti-SNAI1 (#3879), rabbit anti-SLUG (#9585) e rabbit anti-GAPDH (#2118) (Cell Signaling). Após 1 h as membranas foram incubadas com os anticorpos secundários (horseradish peroxidase-conjugated goat anti-rabbit IgG (#7074, Cell Signaling) [16].

O complexo anticorpo-proteína foi detectado usando ECL Western Blotting Detection Reagents (GE Lifesciences) com câmera CCD (Image Quant LAS 4000 mini). A análise densitométrica foi realizada usando o *software* ImageJ e as bandas foram normalizadas para a proteína GAPDH.

3.4 - Preparo de amostras de pacientes

Dez amostras de tecido tumoral de pacientes foram obtidas durante a cirurgia, diagnosticadas previamente com câncer de ovário por exame anatomopatológico, estas amostras foram classificadas de acordo com o tipo de tumor e armazenadas em freezer -80 °C até seu processamento para análise.

Em estudo coordenadado pelo Professor Dr. Francisco J. C. dos Reis, as pacientes foram recrutadas no Hospital das Clínicas de Ribeirão Preto e os procedimentos foram aprovados pelo conselho de ética da instituição (CONEP: 1778/2010).

As amostras coletadas foram descongeladas e lavadas com PBS, cortadas em pedaços de aproximadamente 1 mm^3, adicionados em solução para homogeneização (uréia 8 M, Tris-HCl 0.15 M, pH = 8.8, detergente β-octilglicosídeo 0.5% e 5% de inibidor de proteases (P8340; Sigma) e inibidor de fosfatase Na_3VO_4, 1 mM) na proporção de 0,5 mL para 100 mg de tecido.

Para homogeneização das amostras foi utilizada sonda de ultrassom (Sonic Vibra Cell, VCX 500 – VCX 750, USA), em 3 ciclos de 3 min, com pulso de 1 s e pausa de 2 s com intensidade 30. As soluções de tecido homogeneizado foram centrifugadas 20000 ×

g por 30 min a 4 °C e o sobrenadante foi quantificado para proteínas pelo ensaio BCA (Pierce Biotechnology).

3.5 - Preparo das amostras para análise proteômica dirigida

Alíquotas de volume equivalente a 100 µg de proteína, de cada amostra de paciente, foram reduzidas com 5 µL de solução de DTT (10 mg/mL) por 30 min a 37 °C e alquilados com 12.5 µL de solução de IAA (20 mg/mL) por mais 30 min em temperatura ambiente, ao abrigo da luz. Após este processo as amostras foram diluídas com 740 µL de solução Tris-HCl 0.2 M, pH = 8.0 e submetidas à digestão por tripsina, com 4 µg (1:25 enzima/substrato) (V5111, Promega) a 37 °C *overnight*.

Os peptídeos trípticos foram extraídos com colunas de fase sólida OASIS HLB cartridges (1 cc/10 mg) (Waters), previamente condicionadas com acetonitrila e equilibradas com solução 5% de acetonitrila, a eluição foi feita com 70% de acetonitrila e secagem em SpeedVac (Thermo Scientific).

3.6 – Métodos de proteômica dirigida

3.6.1 - Método de proteômica dirigida para avaliação de marcadores epiteliais e mesenquimais utilizando peptídeos sintéticos

A abordagem proteômica dirigida foi empregada considerando a hipótese de alteração fenotípica da linhagem de células Caov-3, submetidas a tratamento com o fator de crescimento EGF, já observada visualmente durante o tratamento e manutenção da

cultura. Esta hipótese foi testada, desta vez, através do monitoramento de alguns marcadores característicos dos fenótipos epiteliais e mesenquimais.

Primeiramente foi buscado no banco de dados do repositório PeptideAtlas, pelos peptídeos mais observados de cada um dos marcadores escolhidos, Foi então solicitada a síntese destes peptídeos (Thermo Scientific) (tabela 2) para padronização e então aplicação do método, com leituras nas amostras de linhagens.

A padronização do método para cromatografia e espectrometria de massas contou com a determinação de janela ótima de tempo de monitoramento do sinal de cada peptídeo, fragmentos mais prováveis / estáveis e também sua energia ótima de colisão.

Os peptídeos escolhidos para representar os marcadores epiteliais e mesenquimais foram testados com a injeção de padrões de peptídeos sintéticos adquiridos pelo laboratório, e então foram determinados os respectivos tempos de retenção. Ao final este método refinado contou com janelas de aquisição de 2 min para cada peptídeo.

A etapa seguinte foi a otimização da energia de colisão concomitante com a escolha de fragmentos mais representativos, apresentando o melhor perfil cromatográfico e coeluição. Para isso, foram considerados todos os fragmentos compreendendo o intervalo de 400-1200 m/z combinados com variações da energia de colisão, para mais e menos, com intervalos de 2 unidades, partindo da sugestão do *software* Skyline[4'].

A separação cromatográfica foi feita utilizando coluna C_{18} (partícula de tamanho 1.8 µm, poro de 100 Å , 1 mm × 150 mm, Waters) em sistema cromatográfico (I-class,

Waters), com gradiente linear e fluxo de 100 µL/ min, 5–30% da fase B, sendo (100% Acetonitrila + 0.1% ácido fórmico) e fase A (95% Água, 5% acetonitrila +0,1% ácido fórmico) por 25 min.

3.6.2 - Método de proteômica dirigida para painel de proteínas relevantes à assinatura mesenquimal

Para avaliar a presença de proteínas relevantes ao fenótipo mesenquimal em amostras de pacientes, desenvolvemos um painel levando em consideração a combinação de proteínas identificadas no secretoma da linhagem Caov-3 induzida à EMT e a assinatura transcriptômica referente ao subtipo mesenquimal, produzida pelo consórcio TCGA. Então apenas as proteínas constantes como superexpressas na linhagem celular Caov-3, por um critério arbitrário e abrangente de 1,3 vezes, coincidentes com os genes mais expressos no subtipo mesenquimal foram selecionados para o desenvolvimento do painel (Tabela 4).

A motivação para obtenção das proteínas deste painel foi verificar a hipótese de contribuição do fenótipo mesenquimal obtido na linhagem Caov-3, modelo reproduzido em laboratório, na compreensão do surgimento de diferentes tipos de tumores de ovário em pacientes.

Obtidas as proteínas alvo, buscou-se a seleção de peptídeos representativos, recorrendo primeiramente ao banco de dados do repositório PeptideAtlas e complementarmente ao algoritmo PREGO acessório ao *software* Skyline. As sequências

dos 5 peptídeos mais observados de cada proteína foram recuperadas e inseridas no *software* Skyline.

Em seguida, prosseguimos com a determinação do tempo de retenção, realizada com uma estimativa com base na calculadora *Sequence Specific Retention Calculator* (SSRCalc), algoritmo disponibilizado no *software* Skyline que leva em conta o tipo de coluna, material, porosidade e sequência de aminoácidos. O algoritmo foi alimentado com resultados de experimentos anteriores. Foram utilizados 420 peptídeos sintéticos diversos como referência, os quais tiveram seus tempos de retenção verificados empiricamente. A janela de abrangência para detecção dos peptídeos de interesse teve coeficiente r = 0,9484 em 7,1 min.

Neste contexto, a determinação do tempo de retenção foi realizada utilizando-se pools das amostras de lisados de tumores de ovário que seriam analisadas posteriormente. Foram selecionadas transições compreendidas no intervalo de 400 a 1200 m/z, priorizando os valores maiores que o precursor. Em um processo contínuo de análise das corridas cromatográficas e refinamento do método, foram selecionados dois peptídeos representativos, para cada uma das proteínas selecionadas. Apenas os peptídeos com maior sinal, melhor comportamento na cromatografia (anteriormente descrita), adequação à janela de monitoramento e de fragmentos coerentes com os critérios adotados foram considerados para a análise final.

Por fim, a detecção dos peptídeos deste painel de assinatura mesenquimal contou com 18 proteínas e foi programado com tempo de monitoramento otimizado para 1 minuto e 4 transições por peptídeo, totalizando 2 peptídeos por proteína analisada.

3.7 - Comparação entre modelo de progressão tumoral e dados do CPTAC

A avaliação de proteínas relevantes ao fenótipo mesenquimal foi realizada a partir de um painel de candidatos, obtido pela combinação de proteínas identificadas no secretoma da linhagem Caov-3 induzida à EMT e, desta vez, com a assinatura proteômica referente ao subtipo mesenquimal produzida pelo consórcio CPTAC. Os dados disponibilizados pelo consórcio CPTAC, na plataforma PDC (Proteomic data commons) provém de análises realizadas por duas instituições, Johns Hopkins University e Pacific Northwest National Laboratory (JHU e PNNL, respectivamente). Através de um critério arbitrário e abrangente de razão EGF/Ctrl > 1,3, mesmo conjunto de proteínas superexpressas utilizado anteriormente, e coincidentes com as proteínas mais expressas no subtipo mesenquimal, foram selecionadas, resultando em um painel com 23 proteínas de interesse (Figura 16).

3.7.1 - Agrupamento hierárquico

Para a nova assinatura proteômica, foi recuperado do *Proteomic Data Common* (PDC) um novo conjunto de dados sem informação de subtipos, apenas do estágio do

tumor. Com a ajuda do *software* Perseus[5'] v.1.6, foi processado um agrupamento hierárquico para verificação da relevância de classificação através dessa assinatura.

O algoritmo de agrupamento hierárquico decompõe o conjunto de dados hierarquicamente, o que é conveniente para agrupamento subsequente. A maioria dos mapas de calor utiliza um algoritmo de agrupamento hierárquico aglomerativo para agrupar dados e exibir essas informações usando um dendrograma. Uma hierarquia aglomerativa sobre n objetos trata primeiro cada objeto como um grupo, a cada passo, os dois grupos mais próximos são mesclados até que todos os n objetos estejam em um grupo, assim o processo de agrupamento é finalizado. Esse método tem ampla aplicabilidade, e a relação entre os grupos facilmente detectável, apesar da maior complexidade [40].

3.8 - Análise de componentes principais (PCA) e k-means

Para visualização de similaridades dos perfis de expressão entre amostras foi utilizado o *software* QLUcore v.3.8, que gera uma projeção tridimensional dos dados referentes à identificação de proteínas adquirida pelo portal *Proteomic Data Commons* (PDC). Essencialmente, a PCA capta informações essenciais de dados com muitas variáveis e projeta a variância dessas variáveis em menos dimensões, porém, mais significativas e de fácil visualização.

O algoritmo k-means, como os outros citados anteriormente, também possui a função de agrupar, sendo que neste caso foi aplicado para determinar um número pré-determinado de grupos, tal que o erro quadrado entre o centro de um grupo e os dados

deste grupo é minimizado. O centro de um grupo é definido pela média dos elementos que o compõe.

O k-means atribui cada elemento a um único grupo de acordo com a similaridade dos valores das variáveis em cada amostra. Nesta aplicação, o número de grupos foi de cinco, correspondendo aos subtipos moleculares identificados pelo CPTAC.

3.8.1 – Construção de classificador com dados do CPTAC (análise supervisionada)

O classificador foi construído através do *software* QLUcore v.3.8 com o algoritmo Random trees utilizando dados do CPTAC com suas respectivas informações de subtipo molecular para em seguida ser aplicado ao novo conjunto de dados sem informação de subtipo molecular.

O algoritmo Random trees estabelece regras para tomada de decisão e são criadas estruturas onde as variáveis são verificadas paralelamente. Caso a variável seja favorável à classificação que corresponde, o fluxo de verificação continua. Caso contrário, é selecionada outra variável. A denominação *random* se dá pelo caráter aleatório de seleção das variáveis para cada modelo; Para isso, são usados dados de treinamento, neste caso, o conjunto de dados com informação sobre subtipos. Escolhidas as variáveis determinantes, o classificador pode ser aplicado em dados sem informação, neste contexto, ausência dos subtipos moleculares.

4 – Resultados

4.1 - Considerações iniciais

Neste trabalho, avançamos nos estudos publicados anteriormente pelo grupo[16], relativos às alterações ocasionadas pela EMT em células de câncer de ovário (linhagem Caov-3). Desta vez, focamos no secretoma, avaliando a contribuição das proteínas identificadas para a assinatura mesenquimal, para testá-la em amostras de tumores. Abordamos aplicações proteômicas para identificação, seleção e monitoramento de proteínas relevantes ao processo de progressão tumoral em modelo de transição epitélio-mesenquimal. Para este fim foram utilizadas diretamente diferentes abordagens proteômicas, destacando: i) proteômica "*shotgun*"; ii) proteômica dirigida e iii) marcação isotópica. Somado a isso, incluímos no estudo uma análise de dados proteômicos externos.

A abordagem proteômica *shotgun* foi empregada juntamente com marcação isotópica para identificar e mensurar proteínas diferencialmente expressadas no secretoma de células Caov-3 induzidas ao processo de EMT. Com esta identificação inicial foi possível selecionar alvos e aplicar a abordagem proteômica dirigida, objetivando quantificar relativamente a expressão de proteínas em amostras de diferentes tipos de tumor. Além desta aplicação, um método de proteômica dirigida também foi desenvolvido para atestar a presença de biomarcadores clássicos da EMT na linhagem celular Caov-3 utilizando padrões analíticos, trazendo mais confiança na análise e quantificação.

Os dois métodos de proteômica dirigida desenvolvidos, sendo o primeiro, para avaliação de marcadores da EMT e o segundo, para proteínas relevantes da assinatura mesenquimal, foram concebidos de formas distintas. Para o painel da EMT contamos com peptídeos sintéticos, assim o ajuste de parâmetros, tanto cromatográficos, como ligados

ao espectrômetro puderam ser otimizados para maior sensibilidade. Já para o painel de proteínas selecionadas da assinatura mesenquimal foram utilizados peptídeos com as sequências mais observadas no repositório SRMAtlas, estes tiveram sua identificação testada em pools de amostras dos tumores e os mais responsivos integraram o método final, sendo ajustado apenas sua janela de monitoramento e melhores fragmentos.

Essenciais para obtenção das assinaturas foram os dados dos consórcios TCGA[30] e CPTAC[31,42]. O TCGA sequenciou centenas de tumores, destacando sua complexidade e heterogeneidade em diferentes tipos de câncer, permitiu a identificação de novos perfis genômicos específicos que podem ser relevantes no campo clínico. Por sua vez, o CPTAC também fez avanços significativos na simplificação da pesquisa de biomarcadores de proteínas, desenvolvendo protocolos de proteômica quantitativa precisos e reprodutíveis para avaliar candidatos a biomarcadores usando análise proteômica. A recuperação destes dados, disponibilizados de forma simples e acessível, possibilitou a priorização de alvos de uma fonte relevante e representativa, dificilmente obtida localmente.

Outro componente inserido neste trabalho para obtenção de alvos foi a análise de dados públicos, para comparação e extensão das possibilidades que a integração deste conhecimento traz. Para este fim foram utilizados modelos distintos, com dois conjuntos de dados diferentes obtidos do CPTAC e avaliada a separação de grupos entre graus de tumor e semelhança com a assinatura mesenquimal. Apresentamos a seguir um fluxograma que ilustra as estratégias experimentais empregadas em nosso estudo.

4.2- Painel de biomarcadores da EMT

A identificação e quantificação de peptídeos representativos de proteínas características do processo da EMT serviu como avaliação efetiva de sua ocorrência por outra técnica, dado que o processo já havia sido observado morfologicamente e fenotípicamente.

A aplicação da proteômica dirigida para determinação da EMT é indicada, pois se enquadra na premissa da técnica, ser dirigida por uma hipótese e contar com conhecimento prévio, já que este processo conta com marcadores bem conhecidos e alvos estabelecidos. Residindo o desafio no desenvolvimento e aplicação de métodos adequados.

Complementarmente à busca em repositórios e uso de ferramentas bioinformáticas optamos pela verificação empírica dos parâmetros para construção do método, propiciada pela aquisição de padrões dos peptídeos representativos de cada marcador escolhido. Os peptídeos sintéticos adquiridos (tabela 2), cuja sequência constava entre os mais observados no banco de dados do PeptideAtlas, foram testados para seu tempo ótimo de monitoramento e energia ótima de colisão.

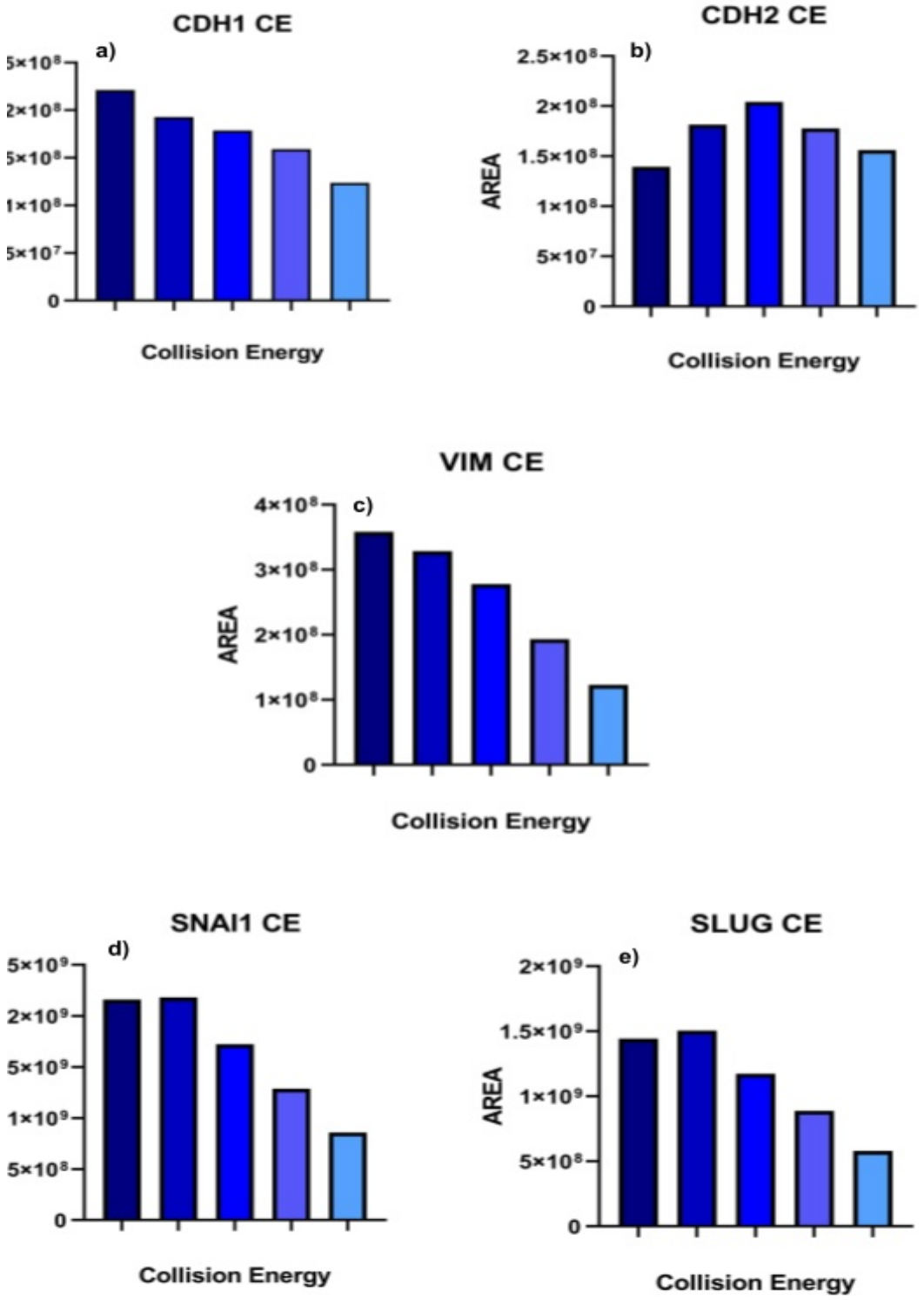
A primeira etapa na análise dos peptídeos deste painel foi a determinação de seu tempo de retenção em nosso método cromatográfico, utilizando como auxílio a calculadora de previsão, como já descrito anteriormente. Nesta etapa inicial todos os peptídeos apresentaram comportamento condizente com a previsão, mostrando picos dentro da janela estabelecida, restando apenas ajuste da janela de monitoramento para o método final.

Tabela 2: Otimização de energia de colisão para painel de biomarcadores clássicos da EMT.

Protein Gene	Peptide Sequence	Predicted time window (min)		Collision energy range
CDH1	NTGVISVVTTGLDR	8,92	16,92	21,5 - 29,5
CDH2	GPFPQELVR	7,28	15,28	14,3 - 22,3
VIM	ILLAELEQLK	12,74	20,74	16,6 - 24,6
SNAI1	EYLSLGALK	7,66	15,66	13,3 - 21,3
SNAI2	EYVSLGALK	5,53	13,53	13,1 - 21,1

Após a obtenção das janelas de monitoramento, iniciamos a otimização da energia de colisão dos peptídeos, onde para cada peptídeo foram selecionados fragmentos do tipo y de m/z no intervalo 400-1200 para monitorar. Estes peptídeos foram submetidos a diferentes energias de colisão, tendo como base a sugestão do *software* Skyline. Em seguida, foram determinados quatro outros valores para o teste, variando duas unidades para mais e para menos, totalizando cinco testes de energia de colisão para cada peptídeo (figura 4). Para o método final foram selecionados os fragmentos associados a energia de colisão que proporcionaram maior sinal na aquisição. A visualização dos dados dos testes realizados através do *software* Skyline possibilitou a otimização de energia de colisão individual para cada fragmento.

Figura 4: **Otimização de energia de colisão de peptídeos representativos de biomarcadores clássicos da EMT.** A otimização de energia de colisão ocorreu com testes, partindo do referencial sugerido pelo *software* Skyline foram testadas variações para valores maiores e menores com intervalo de duas unidades. Cada cor representa uma energia de colisão, sendo um gradiente das energias mais baixas (escuras) até mais altas (claras).



A otimização de energia dos peptídeos referentes às proteínas CDH2, SNAI1 e SLUG apresentou seu máximo de intensidade dentro do intervalo de energias testadas, já os peptídeos referentes à CDH1 e VIM mostraram tendência ao aumento de sua intensidade com menores energias, dada a sugestão do *software* Skyline.

O método final contou com os seguintes parâmetros otimizados:

Tabela 3: Parâmetros utilizados no monitoramento de biomarcadores clássicos da EMT.

Protein Gene	Peptide Sequence	Precursor m/z	Product m/z	Used collision energy	Measured retention time (min)
CDH1	NTGVISVVTTGLDR	716,39373	1060,5997 947,515636 761,415194	23,46 21,46 21,46	13,8
CDH2	GPFPQELVR	521,787641	888,493778 741,425364 644,3726	18,25 18,25 22,25	11,15
VIM	ILLAELEQLK	585,360639	943,545873 830,461809 759,424696	16,61 18,61 18,61	17,18
SNAI1	EYLSLGALK	497,2844	701,455602 588,371538 501,339509	15,34 15,34 13,34	10,61
SNAI2	EYVSLGALK	490,276575	687,439952 588,371538 501,339509	15,09 15,09 13,09	10,61

4.3 - Avaliação do modelo de Indução da EMT

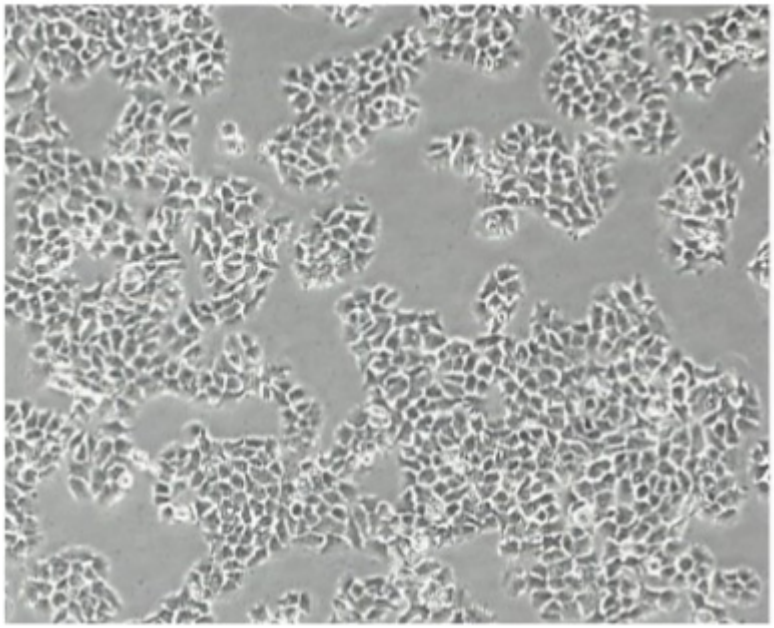
O modelo de indução da EMT utilizado foi desenvolvido e avaliado pelo grupo em outros trabalhos. A linhagem celular, tratamento, fator de crescimento e sua concentração foram escolhidos em particular por mostrar maior diferenciação morfológica nos experimentos realizados [16].

A linhagem de células de adenocarcinoma de ovário Caov-3 foi tratada com fator de crescimento EGF na concentração de 10 ng/mL por 96 h com troca diária do meio, sendo este reservado para as análises posteriores e as células utilizadas para avaliação morfológica e fenotípica para características da EMT.

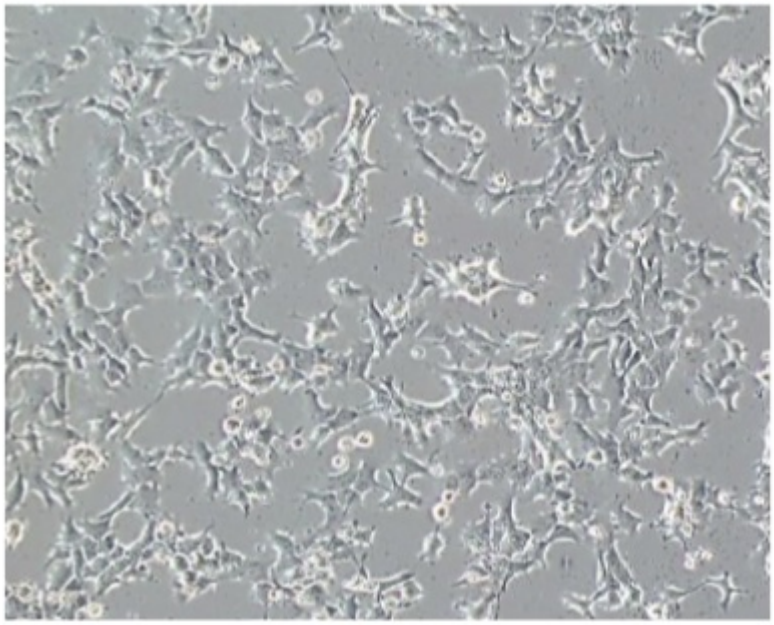
Ao final das etapas de cultivo e indução da EMT as células foram avaliadas para verificação do resultado do tratamento, pela observação de eventuais alterações condizentes com o processo, assim imagens de microscopia no conjunto de células imediatamente pré e pós tratamento foram obtidas e comparadas (figura 5). Foi observado que o tratamento gerou mudanças como pode ser visto na figura a seguir.

Figura 5: **Imagens representativas das alterações morfológicas de células da linhagem Caov-3 tratadas com fator de crescimento EGF.** a) Células da linhagem Caov-3 cultivadas sem tratamento. b) Células da linhagem Caov-3 tratadas com fator de crescimento EGF 10 ng/mL por 96 h. Imagens obtidas por microscopia de contraste de fases com aumento de 40 x.

a)



b)



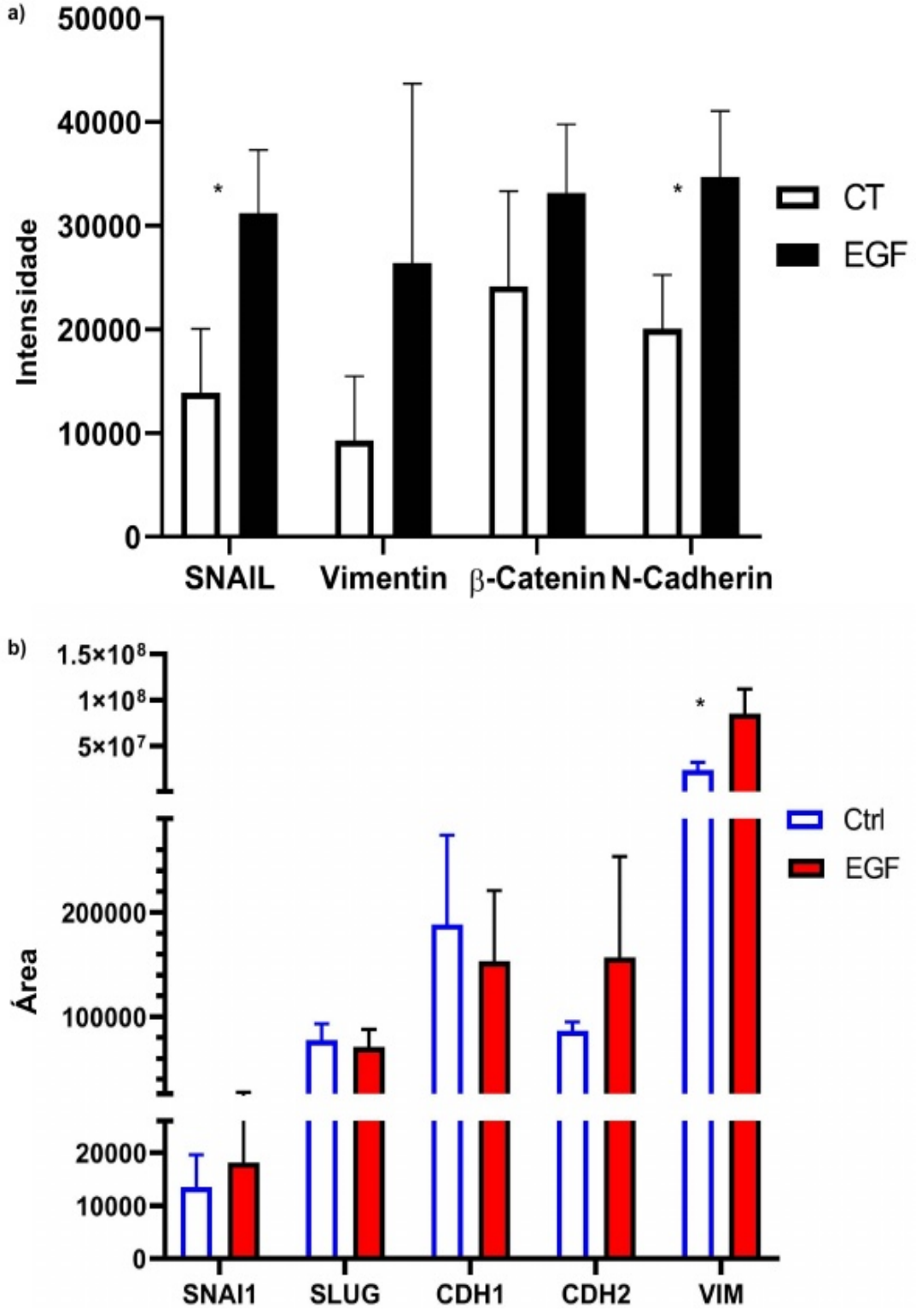
As imagens de microscopia mostram as células antes e depois do tratamento de 96 h. Podem ser observadas diversas alterações. No campo pré tratamento é possível observar a formação de grupos de aglomerados celulares assemelhados a sincícios, o que acontece em menor proporção e intensidade no campo pós tratamento, uma visível redução do contato celular. Já as células individualmente se tornaram alongadas, com maior área e multi-fusiformes, algumas parecidas com fibroblastos. Estas observações morfológicas estão de acordo com a EMT pela obtenção de células com características mais próximas à mesenquimais.

Então, buscou-se a confirmação do processo por alterações fenotípicas no nível de expressão proteica, para isso marcadores moleculares clássicos indicativos da EMT foram quantificados, e com este objetivo foram utilizadas duas técnicas, Western blotting e Espectrometria de massas pela técnica de MRM utilizando o painel desenvolvido para este fim, cobrindo algumas proteínas relevantes (figura 6).

O resultado obtido por Western blotting (Figura suplementar 1) demonstrou aumento significativo nas intensidades de bandas referentes aos marcadores N-Caderina e SNAI1, ambos associados ao fenótipo mesenquimal.

Para avaliação da reprodutibilidade de resultados e confirmação da indução efetiva do processo da EMT, o painel de biomarcadores da EMT desenvolvido com padrões de peptídeos sintéticos representativos das proteínas foi empregado na análise das amostras de extrato total da cultura de células. Este método permitiu a quantificação relativa simultânea dos marcadores em um experimento otimizado.

Figura 6: **Quantificação relativa por Western blotting e LC-MRM de biomarcadores clássicos da EMT em cultura de células Caov-3 tratadas com fator de crescimento. a) Western blotting. b) MRM.** O método otimizado para avaliação de marcadores da EMT foi aplicado em triplicata em amostras provenientes de extrato total de células. * p<0,1 t-test student.



Utilizando espectrometria de massas através da abordagem proteômica dirigida em amostras de extrato total das células Caov-3 com triplicata de injeção, tivemos a expressão diferencial dos marcadores.

Nesta análise do modelo de progressão tumoral foram observados aumentos expressivos na quantidade de peptídeos representativos das proteínas consideradas marcadores mesenquimais, N-Caderina e Vimentina, sendo o aumento desta última significativo estatisticamente, SNAI1 apresenta tendência ao aumento. Já a proteína E-caderina, marcador epitelial, apresentou redução de seus níveis, também observado sensivelmente em SLUG. Baseados nestes resultados, concluímos que o modelo por nós utilizado representa bem a EMT e, portanto, pode ser utilizado para auxiliar na proposição de assinaturas moleculares para avaliação de amostras clínicas.

4.4 - Identificação de proteínas no secretoma de células Caov-3 induzidas à EMT

O experimento de identificação das proteínas do secretoma foi realizado nas culturas induzidas e não induzidas à EMT da linhagem Caov-3. Além disso foi utilizado um controle de células cultivadas em meio contendo isótopos pesados de lisina($^{13}C_6$). A amostra de cada secretoma (referência x tratado) foi composta pelo meio recolhido diariamente da respectiva cultura de células (controle x EGF) com adição de mesma massa de proteínas de controle, obtida da linhagem com incorporação de isótopos. A incorporação foi verificada por espectrometria de massas e foram determinadas cinco passagens necessárias.

A combinação dos meios de cultura que compuseram os secretomas (células Caov-3 incubadas com EGF + células Caov-3 SILAC x células Caov-3 sem EGF + células

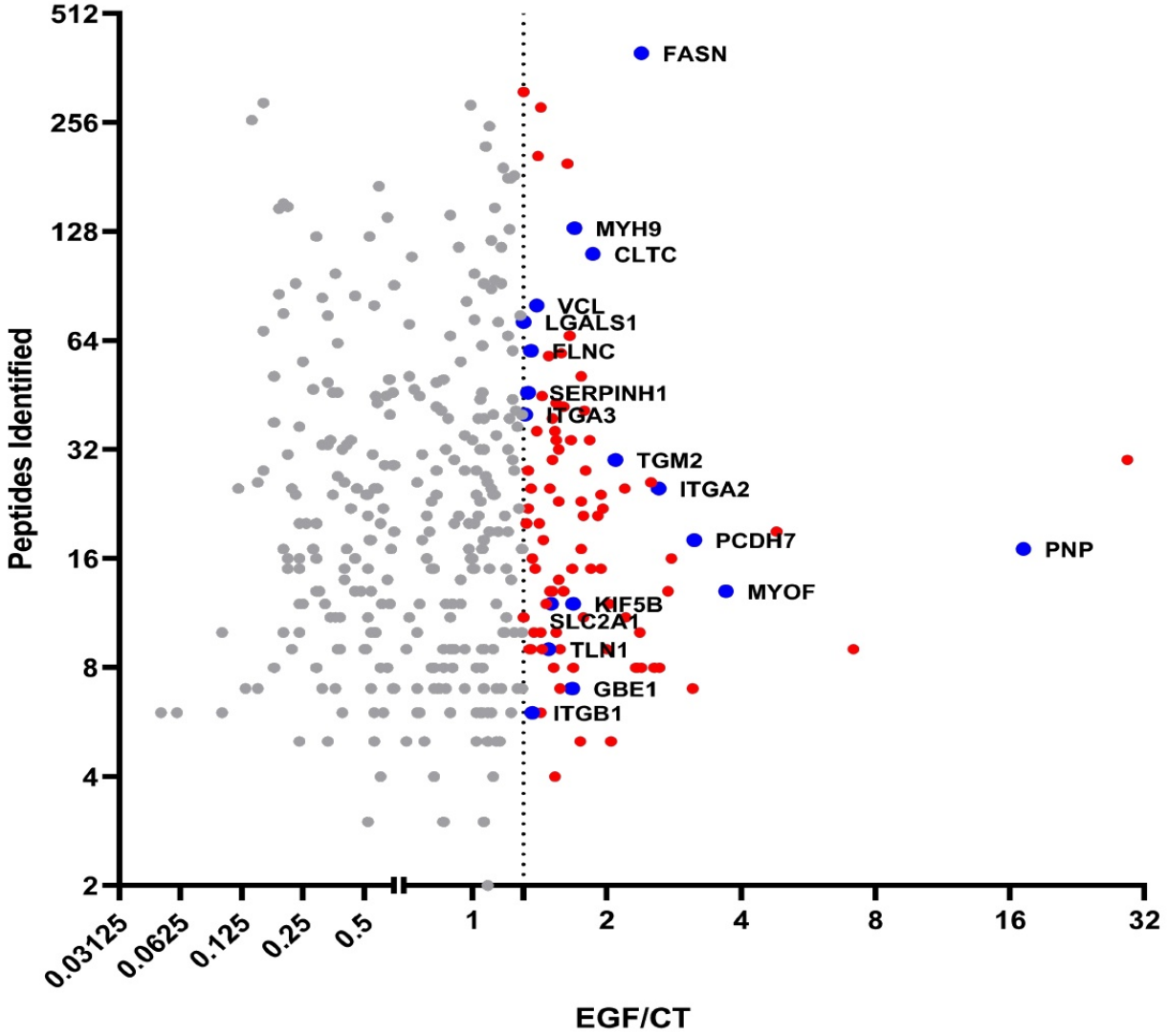
Caov-3 SILAC) foram submetidas à eletroforese em gel de poliacrilamida para fornecer amostras menos complexas para o sistema LC-MS.

O gel resultante da eletroforese foi dividido em 5 fragmentos por faixa (Figura suplementar 1), processado e as proteínas extraídas submetidas à análise, após análise os dados de identificação de cada fração do gel foram combinados.

As proteínas identificadas em cada secretoma com ao menos dois peptídeos únicos e taxa de falsa descoberta (FDR) < 1% foram consideradas para etapas posteriores de análise.

.

Figura 7: **Proteínas identificadas por LC-MS/MS no secretoma de células Caov-3.** Gráfico de dispersão mostrando a identificação diferencial de proteínas usando a razão EGF/CT vs.número de peptídeos identificados por proteína (apenas >2 peptídeos únicos). Proteínas em cinza (com razão EGF/CT < 1,3) não foram consideradas para os estudos a seguir. Proteínas com razão EGF/CT > 1,3, em vermelho. Proteínas consideradas para o ensaio MRM em destacadas em azul (LANFREDI et al. 2021).



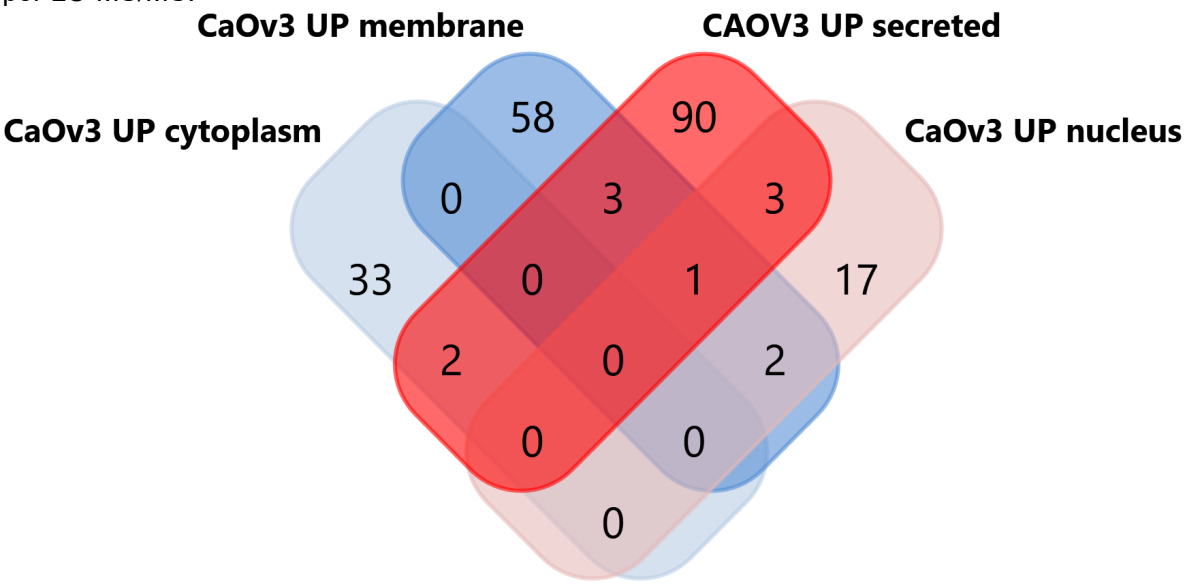
O gráfico da Figura 7 demonstra a distribuição de proteínas identificadas em razão de identificação no tratamento com EGF por identificação no controle, normalizada pela razão isotópica, contra número de peptídeos únicos identificados. A razão isotópica mostra a quantificação relativa de cada proteína ao controle utilizado, sendo a denominação *heavy* (H) sempre relativo a cultura com incorporação de isótopos de maior massa, neste caso utilizada como referência, já o número de peptídeos únicos indica a cobertura na identificação de cada proteína.

No total, obtivemos correspondência de identificação positiva para aproximadamente 10.000 peptídeos em cada experimento, correspondendo a 2.847 proteínas identificadas (representadas por um nome de gene), reduzindo para 697, utilizando os critérios mencionados e 99 superexpressas com critério de razão > 1,3.

A razão utilizada provém da combinação de todas as identificações, portanto uma “razão da razão”, gerada da combinação de identificação de conjunto controle e conjunto tratado.

Devido ao critério de corte utilizado para considerar as proteínas identificadas diferencialmente expressas consistindo em valor relativamente “baixo” (1,3), pelo eventual efeito de diluição, conduzimos uma análise em subfrações celulares para comparar numericamente apenas o conjunto de proteínas superexpressas (figura 8).

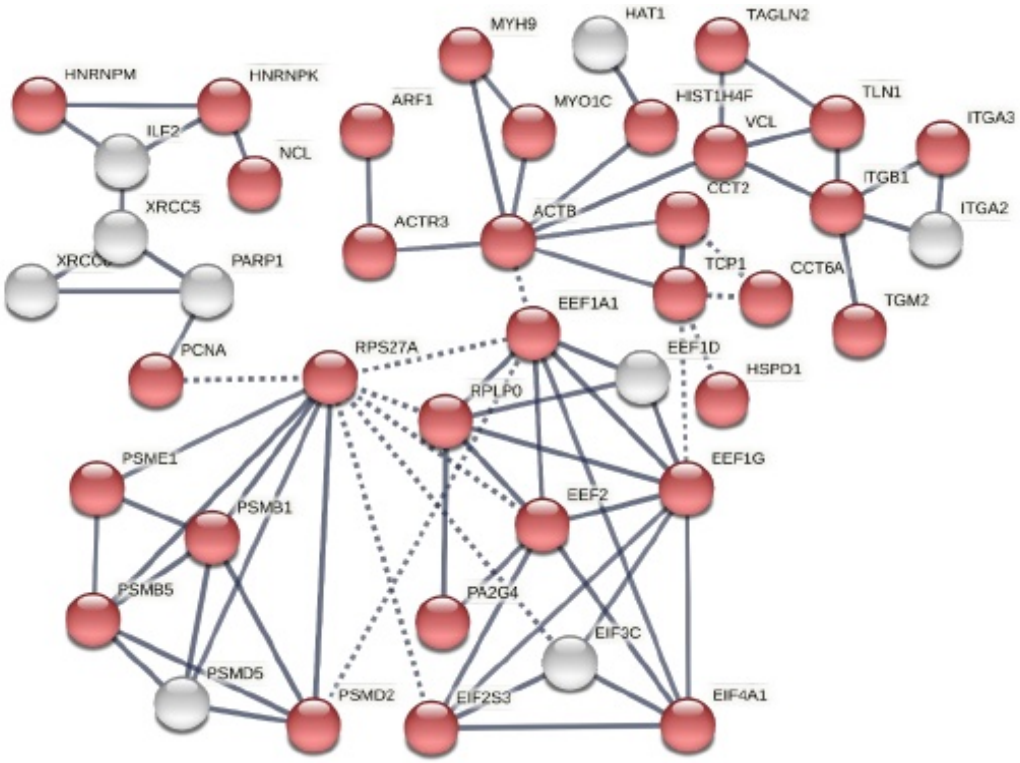
Figura 8: **Diagrama de Venn com proteínas superexpressas.** Identificadas em cada subfração celular obtida de cultura de células da linhagem Caov-3 e secretoma, quantificadas por LC-MS/MS.



No diagrama de Venn observamos maior número de proteínas identificadas no subconjunto de proteínas superexpressas no secretoma, porém com baixa sobreposição nas subfrações celulares, indicando que apesar de incluir mais proteínas devido o critério adotado, isso praticamente não resultou em abrangência de proteínas enriquecidas em outras subfrações.

Após a identificação, analisamos as proteínas superexpressas utilizando a ferramenta *open-source* String Database para obtenção de uma rede de interação associando as proteínas, baseada em experimentos e dados da plataforma, com critério restritivo de confiança 0,7, apenas com proteínas conectadas (figura 9).

Figura 9: **Rede de interação obtida pelo *software* String contendo proteínas superexpressas no secretoma obtido de células Caov-3.** Proteínas agrupadas por proximidade agrupadas por k-mean, considerando índice de confiança de 0,9. Com destaque em vermelho proteínas identificadas como componentes de vesículas extracelulares.



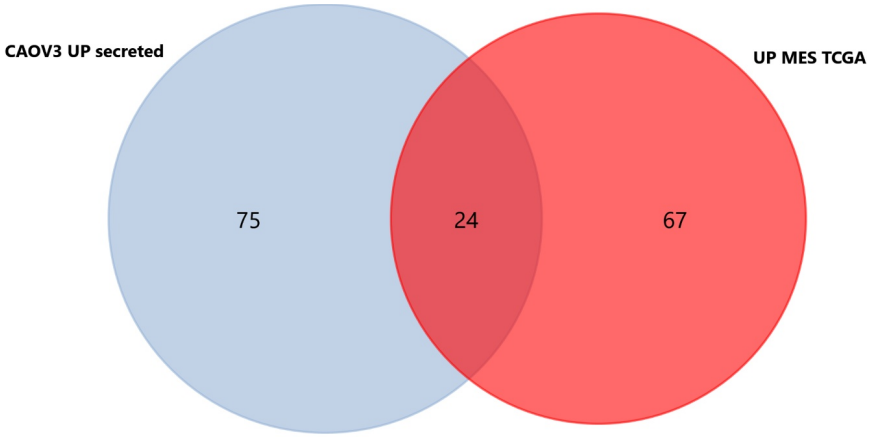
As redes obtidas pela interação entre proteínas selecionadas por sua superexpressão no processo induzido da EMT, em células Caov-3, contempla diversas classes de proteínas. Pode ser observada a rede de proteínas, destacada na figura, formada pelo agrupamento de suas correlações experimentais, em bancos de dados e na literatura. Esta rede está associada aos processos biológicos de deubiquitinação de proteínas, transporte mediado por vesículas, tradução e síntese de proteínas, localização e transporte de proteínas[41] (Figura suplementar 2). Um nó importante que une as sub-redes é a proteína ribossômica S27a-Ubiquitina-40S (RPS27A), que é fundida a uma cópia do polipeptídeo ubiquitina. Outro destaque observado são proteínas identificadas em vesículas, sendo a maioria das proteínas incluídas na análise.

4.5 - Construção de uma assinatura molecular proteo-genômica do subtipo mesenquimal para o Câncer de Ovário

A análise de proteínas superexpressas na cultura de células Caov-3 serviu de base para a seleção de biomarcadores relevantes à determinação da assinatura mesenquimal. Primeiramente, recuperamos os dados disponibilizados do consórcio TCGA referentes às proteínas secretadas identificadas no modelo. Pela integração de dados proteômicos próprios e dados externos de transcriptômica foi obtido o primeiro painel de alvos para avaliação por proteômica dirigida em amostras de pacientes, considerando a expressão média aumentada no subtipo mesenquimal.

O painel de proteínas relevantes da assinatura mesenquimal foi feito com a contribuição de dois conjuntos de dados (figura 10), as proteínas superexpressas no secretoma e os genes mais expressos no subtipo mesenquimal, produzido pelo consórcio TCGA.

Figura 10: **Diagrama de sobreposição de proteínas superexpressas no secretoma com transcritos mais expressos no subtipo mesenquimal (TCGA).** Apenas transcritos com maior média de expressão e correspondente identificação em nossos dados foram selecionados.



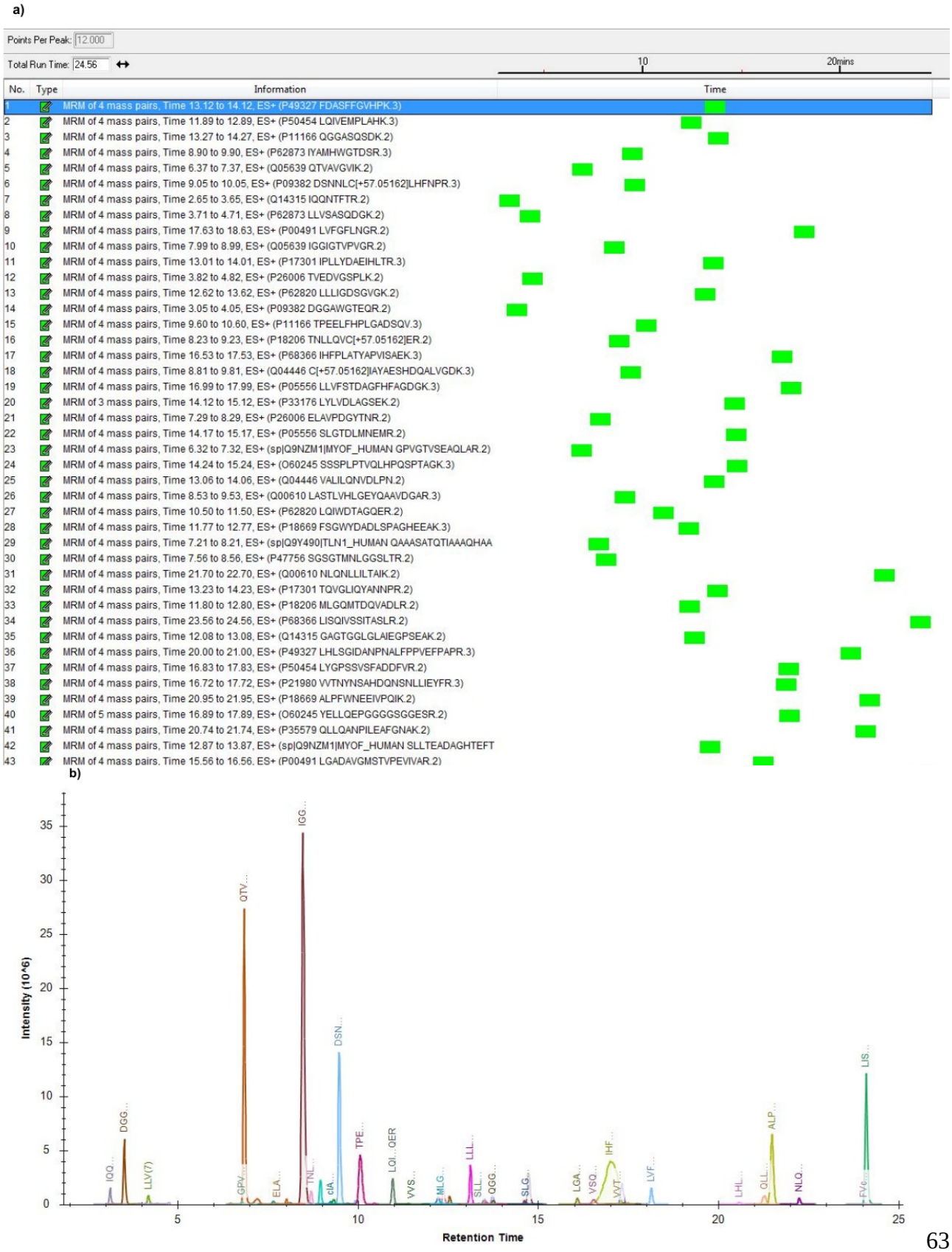
As sequências de peptídeos de cada proteína constante no painel foram recuperadas do repositório PeptideAtlas e testada em pool de amostras de tumor a serem

avaliadas posteriormente. Os peptídeos representativos que obtiveram sinal foram mantidos no método, os outros foram descartados. Neste passo para proteínas sem peptídeos representativos remanescentes foi acessado o algoritmo PREGO para sugestão de mais peptídeos.

A determinação do tempo de retenção ocorreu como descrito anteriormente, com estimativa de tempo de retenção calculada e pool das amostras para testes.

O desenvolvimento deste painel resultou no método cromatográfico otimizado em termos de janela de monitoramento e fragmentos ilustrado abaixo na figura 11, diferentemente do painel de marcadores da EMT, este painel não contou com otimização de energia de colisão, pela ausência de padrões para teste.

Figura 11: **Método de proteômica dirigida para painel de proteínas da assinatura mesenquimal.** a) Representação do método de proteômica dirigida utilizado em todas as amostras de pacientes com janelas de monitoramento otimizadas em verde. b) Exemplo de cromatograma obtido das amostras de pacientes, mostrando todos os picos cromatográficos obtidos.



O método final para monitoramento contou com os seguintes parâmetros:

Tabela 4: Painel de peptídeos para o monitoramento de proteínas relevantes à assinatura mesenquimal. Continua ►

Protein Gene	Peptide Sequence	Precursor Mz	Precursor Charge(+)	Product Mz	Average Measured Retention Time (min.)
PNP	LVFGFLNGR	512,1111	2	909,4941 810,4257 663,3573 606,3358	18,1
PNP	LGADAVGMSTVPEVIVAR	893,5484	2	1070,6204 983,5884 783,4723 557,377	16,1
MYOF	GPVGTVSEAQLAR	643,2207	2	1031,548 873,4789 774,4104 558,3358	6,8
MYOF	SLLTEADAGHTEFTDEVYQNESR	871,9027	3	895,4268 796,3584 633,2951 505,2365	13,3
PCDH7	YELLQEPGGGGSGGESR	847,3817	2	917,4071 820,3544 706,3115 649,29	17,3
PCDH7	SSSPLPTVQLHPQSPTAGK	645,0576	3	922,4741 785,4152 688,3624 560,3039	14,7
ITGA2	IPLLYDAEIHLTR	518,9442	3	839,4734 768,4363 639,3937 526,3096	13,5
FASN	FDASFFGVHPK	418,1399	3	918,4832 831,4512 684,3828 537,3144	13,6
FASN	LHLSGIDANPNALFPPVEFPAPR	825,2778	3	1156,615 1009,5465 912,4938 587,33	20,6
TGM2	VVSGMVNCNDDQGVLLGR	967,5988	2	1086,5538 857,4839 742,457 614,3984	11,5
TGM2	VVTNYNSAHDQNSNLLIEYFR	833,5705	3	1154,6204 840,4614 727,3774 614,2933	17,2
CLTC	NLQNLLILTAIK	677,8368	2	998,6608 771,5339 658,4498 545,3657	22,2
CLTC	LASTLVHLGEYQAAVDGAR	658,0698	3	1136,5331 1079,5116 950,469 787,4057	9
MYH9	QLLQANPILEAFGNAK	864,4981	2	1059,5833 849,4465 736,3624 607,3198	21,3
MYH9	SMEAEMIQLQEELAAAER	1025,6552	2	1129,5848 1016,5007 630,357 517,2729	22,2
KIF5B	FVCSPDEVMDTIDEGK	922,0177	2	1007,4714 777,3625 662,3355 561,2879	24

KIF5B	LYLVDLAGSEK	604,6978	2	931,5095 818,4254 719,357	14,6
GBE1	VALILQNVDLPN	655,2732	2	1138,6467 1025,5626 557,293 458,2245	13,5
SLC2A1	QGGASQSDK	439,444	2	749,3424 635,2995 564,2624 477,2304	13,7
SLC2A1	TPEELFHPLGADSQV	547,5976	3	1070,5265 786,3992 689,3464 576,2624	10,1
TLN1	QAAASATQTIAAAQHAASTPK	666,063	3	1052,5483 981,5112 910,4741 711,3784	7,6
TLN1	VSQMAQYFEPLTLAAVGAASK	1092,2621	2	1098,6517 1001,599 787,4672 674,3832	16,5
ITGB1	SLGTDLMNEMR	634,2335	2	1066,4656 908,3964 793,3695 680,2854	14,7
ITGB1	LLVFSTDAGFHFAGDGK	594,9983	3	1121,501 1006,4741 935,437 731,3471	17,5
VCL	MLGQMTDQVADLR	739,8647	2	1048,5092 917,4687 816,421 701,3941	12,3
VCL	TNLLQVCER	567,1563	2	1031,5302 804,4032 691,3192 464,1922	8,7
FLNC	IQQNTFTR	504,5612	2	766,3842 638,3257 524,2827 423,235	3,1
FLNC	GAGTGGLGLAIEGPSEAK	793,3755	2	901,4625 830,4254 717,3414 531,2773	12,5
SERPINH1	LYGPSSVSFADDFVR	830,9094	2	1055,5156 956,4472 869,4152 722,3468	17,3
SERPINH1	LQIVEMPLAHK	427,1979	3	1037,5812 924,4971 825,4287 565,3457	12,4
ITGA3	ELAVPDGYTNR	618,1689	2	992,4796 822,3741 725,3213 553,2729	7,7
LGALS1	DSNNLCLHFNPR	496,5475	3	943,4567 783,426 670,342 533,2831	9,5
LGALS1	DGGAWGTEQR	539,0492	2	847,4057 776,3686 590,2893 533,2678	3,5

4.6 - Análise de amostras clínicas para a assinatura proteômica do subtipo mesenquimal

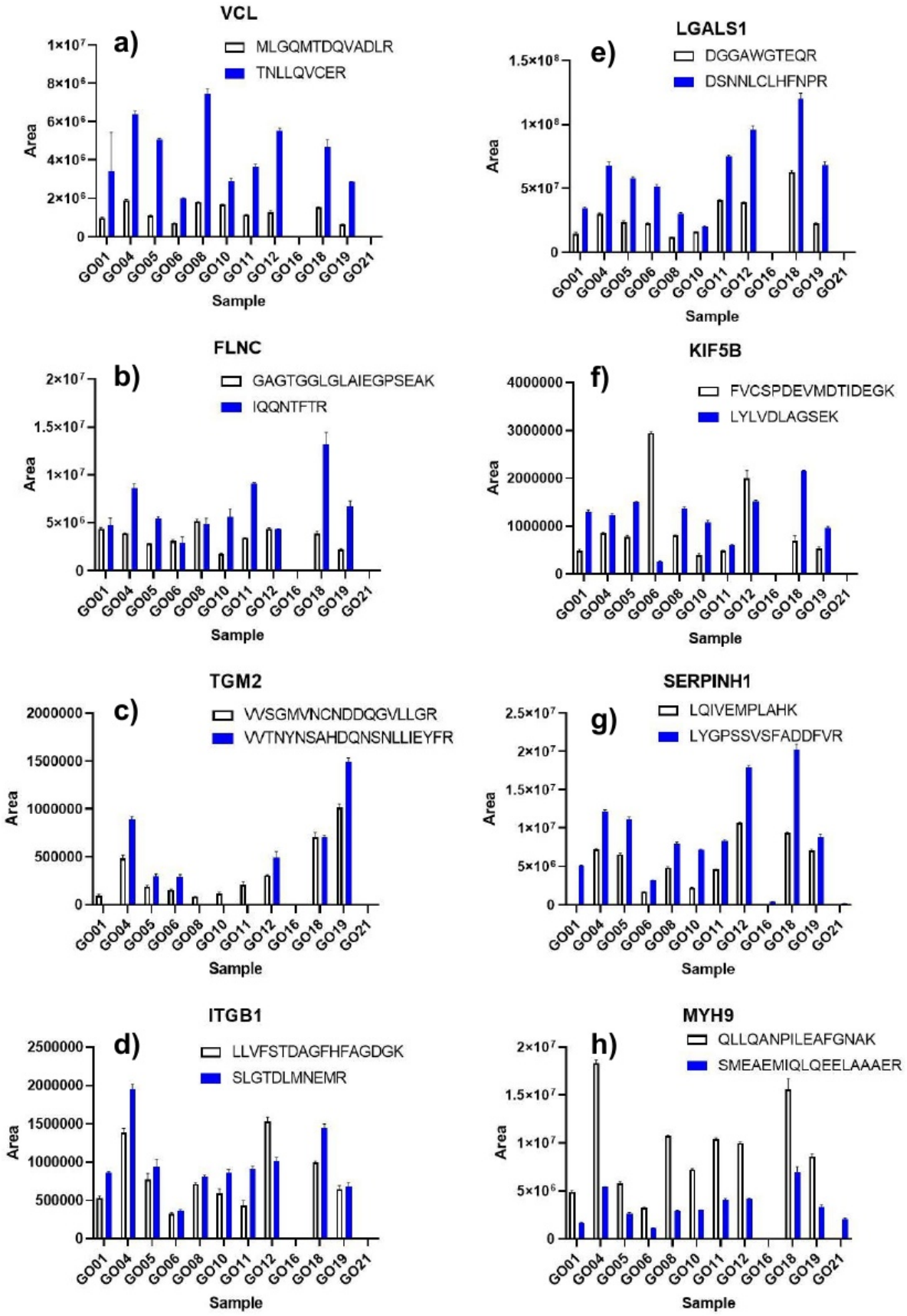
As 10 amostras utilizadas para avaliação do painel são provenientes de cirurgia em pacientes diagnosticadas com câncer de ovário, detalhadas na tabela 5. Estas amostras passaram por avaliação anatomopatológica para determinação de tipos e estágios, porém sem observação de assinatura molecular.

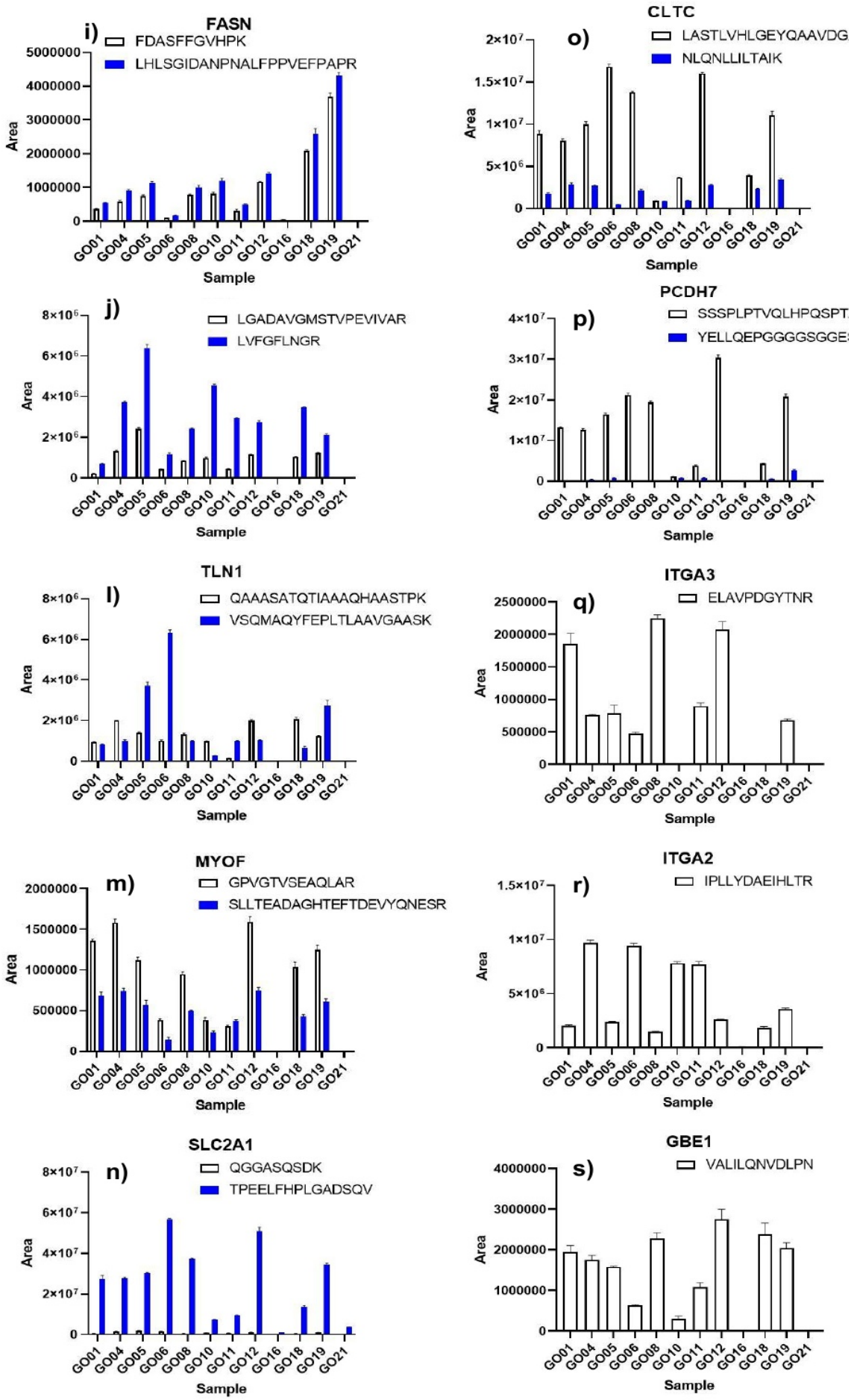
Tabela 5: Identificação de amostras de tumores de pacientes submetidos à análise por MRM.

Case	Age (Years)	Histology	Stage	Type
GO 01	58	High-Grade Serous Ovarian Cancer (HGSOC)	IIA	II
GO 05	64	High-Grade Serous Ovarian Cancer (HGSOC)	IIIB	II
GO 06	45	Mucinous	IV	I
GO 10	56	High-Grade Serous Ovarian Cancer (HGSOC)	IIIC	II
GO 11	63	High-Grade Serous Ovarian Cancer (HGSOC)	IIIC	II
GO 12	56	Metastatic Endometrial Carcinoma (MEC)	IIIA	NA
GO 16	41	Low Grade Serous (LGS)	IIIA	I
GO 18	64	High-Grade Serous Ovarian Cancer (HGSOC)	IIIC	II
GO 19	71	High-Grade Serous Ovarian Cancer (HGSOC)	IIIB	II
GO 21	56	Clear Cell (CC)	IA	I

NA = not applicable

Figura 12: **Resultados individuais por proteína do método de proteômica dirigida desenvolvido para assinatura mesenquimal.** Cada gráfico representa os resultados de uma proteína, cada barra representa a área de pico cromatográfico, obtida do peptídeo monitorado. Continua ►

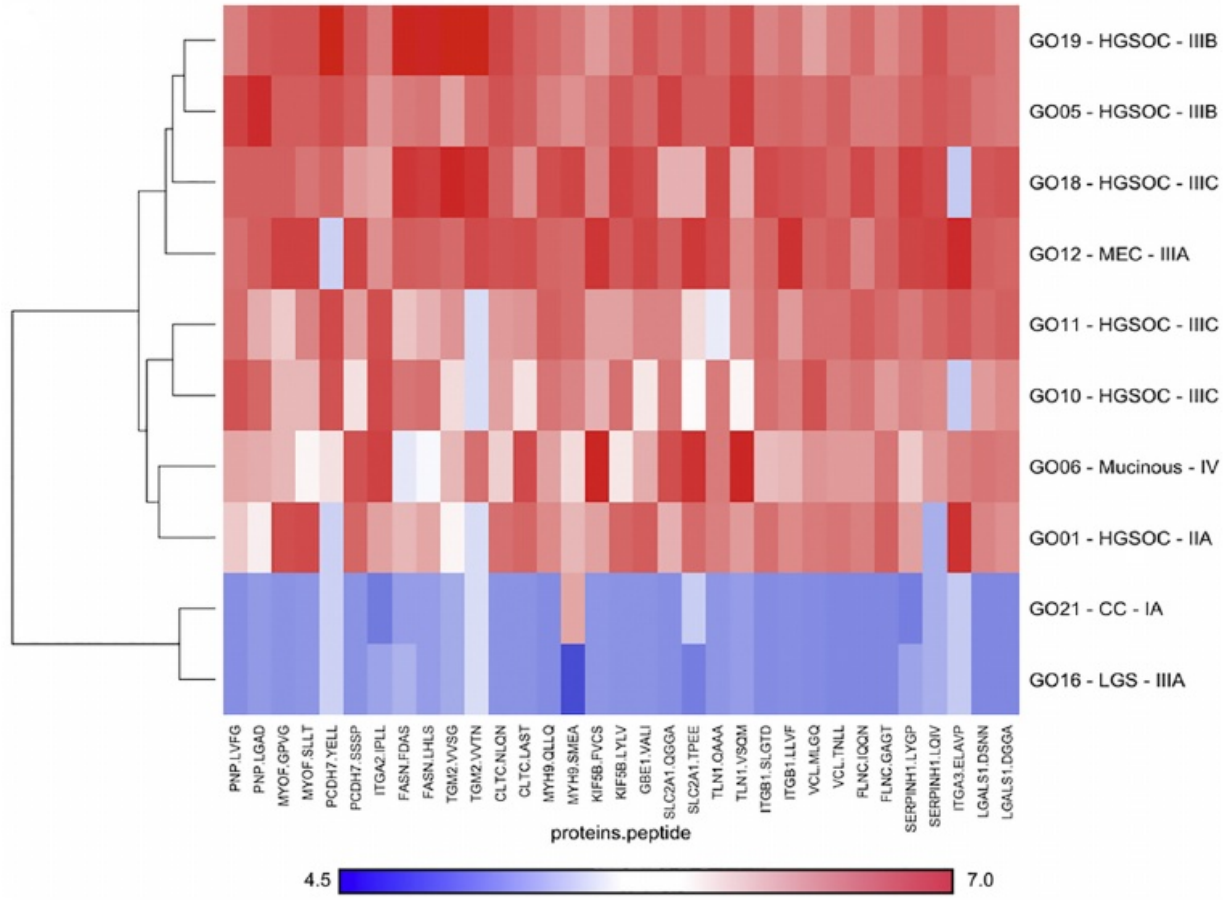



i)
FASN
FDASFFGVHPK
LHLSGIDANPNALFPPVEFPAPR
Area
Sample
j)
LGADAVGMSTVPEVIVAR
LVFGFLNGR
Area
Sample
l)
TLN1
QAAASATQTIAAAQHAASTPK
VSQMAQYFEPLTLAAVGAASK
Area
Sample
m)
MYOF
GPVGTVSEAQLAR
SLLTEADAGHTEFTDEVYQNESR
Area
Sample
n)
SLC2A1
QGGASQSDK
TPEELFHPLGADSQV
Area
Sample
o)
CLTC
LASTLVHLGEYQAAVDG
NLQNLLILTAIK
Area
Sample
p)
PCDH7
SSSPLPTVQLHPQSPT
YELLQEPGGGGSGGE
Area
Sample
q)
ITGA3
ELAVPDGYTNR
Area
Sample
r)
ITGA2
IPLLYDAEIHLTR
Area
Sample
s)
GBE1
VALILQNVDLPN
Area
Sample
GO01
GO04
GO05
GO06
GO08
GO10
GO11
GO12
GO16
GO18
GO19
GO21

A quantificação individual de peptídeos pela abordagem MRM (Figura 12) contou com 2 peptídeos para a maioria das proteínas identificadas, sendo exceção ITGA2, ITGA3 e GBE1 contando com apenas um peptídeo, já SLC2A1 e PCDH7, apesar de contarem com 2 peptídeos em seu monitoramento, mostraram sinal apenas para um deles na maioria das amostras. O sinal de cada par de peptídeos é concordante e proporcional entre a maioria das amostras, excetuando as amostras 6, 12 e 18 na identificação de peptídeos referentes às proteínas TLN1, ITGB1,KIF5B E FLNC, porém apresentando resultados consistentes nas triplicatas.

Um dendograma sumarizando os resultados de MRM foi montado com agrupamento hierárquico para melhor observação do possível agrupamento obtido pelo monitoramento do painel de proteínas.

Figura 13: **Avaliação da assinatura molecular de expressão gênica de proteínas em amostras tumorais de pacientes.** Análise de agrupamento das intensidades de peptídeos por MRM obtidas para amostras de pacientes individuais. As intensidades representam a média decorridas triplicadas. Um limite arbitrário de detecção igual a 4,5 (escala log 10) foi considerado para todos os peptídeos avaliados. O dendrograma apresentado à esquerda do *heatmap,* indica a segregação de amostras de pacientes com base no monitoramento MRM(LANFREDI et al. 2021).



O dendograma obtido pela avaliação do *heatmap* mostra uma clara divisão entre os tipos de amostras, menos responsivas, GO16 e GO21, únicas de seus tipos, e o restante, mais responsivas em diferentes níveis para cada peptídeo. Há um agrupamento de amostras do mesmo tipo de tumor (HGSC) em diferentes pontos, com separação de graus, IIA de IIIC, incluindo uma das amostras de câncer de endométrio, também de alto grau. A amostra de câncer de ovário mucinoso, também única de seu tipo, porém de mais elevado grau entre todas as amostras, se junta às de tipo (HGSC).

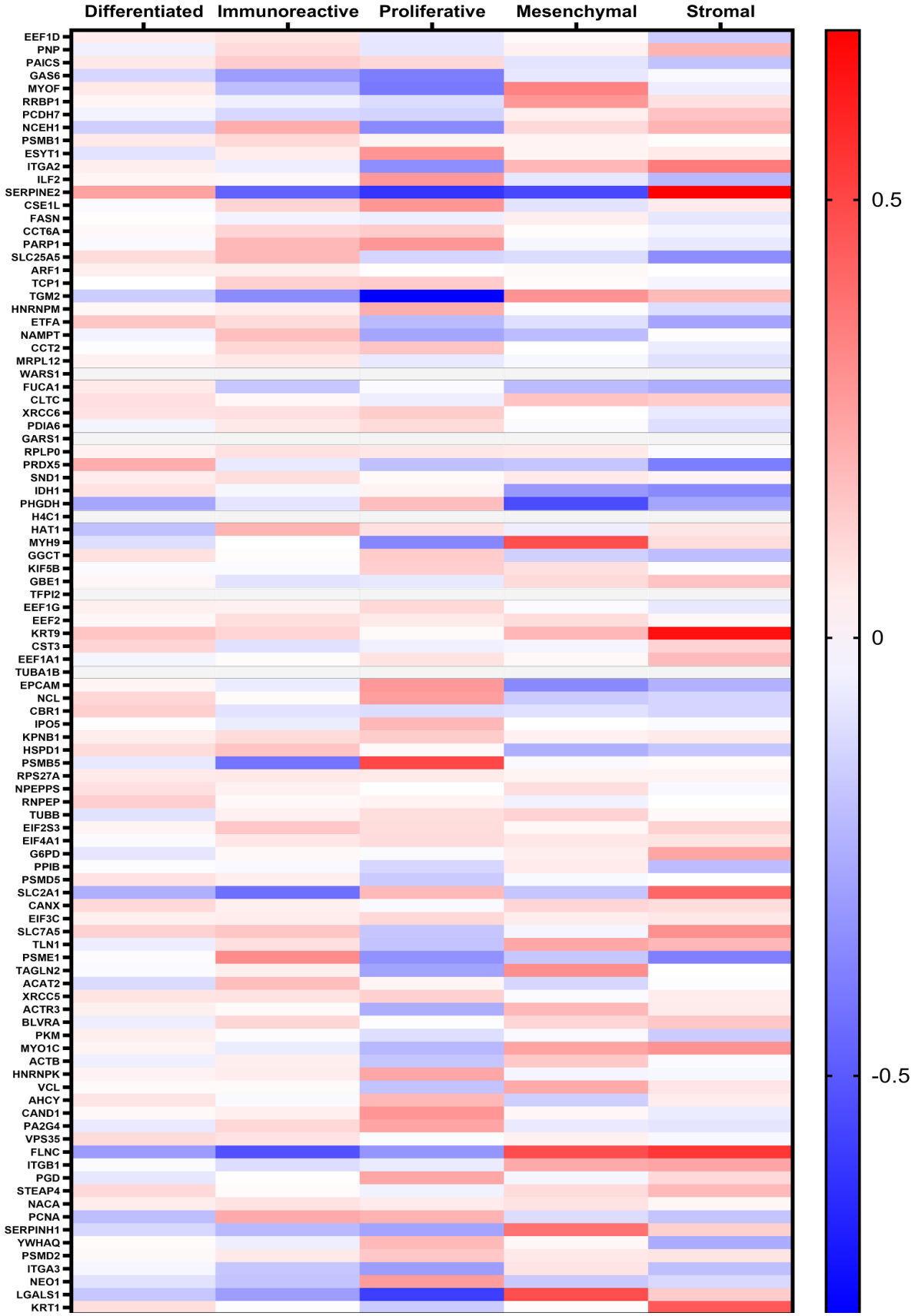
Considerando os resultados obtidos para as amostras de pacientes analisadas, fica claro o potencial da combinação de dados para refinamento das assinaturas moleculares para classificação de subtipos de câncer. Também é evidente que o número de pacientes a serem analisados precisa ser estendido para um resultado estatístico mais evidente. Assim, partimos para a integração mais detalhada dessas assinaturas iniciais com dados de outras iniciativas proteômicas focadas no câncer de ovário.

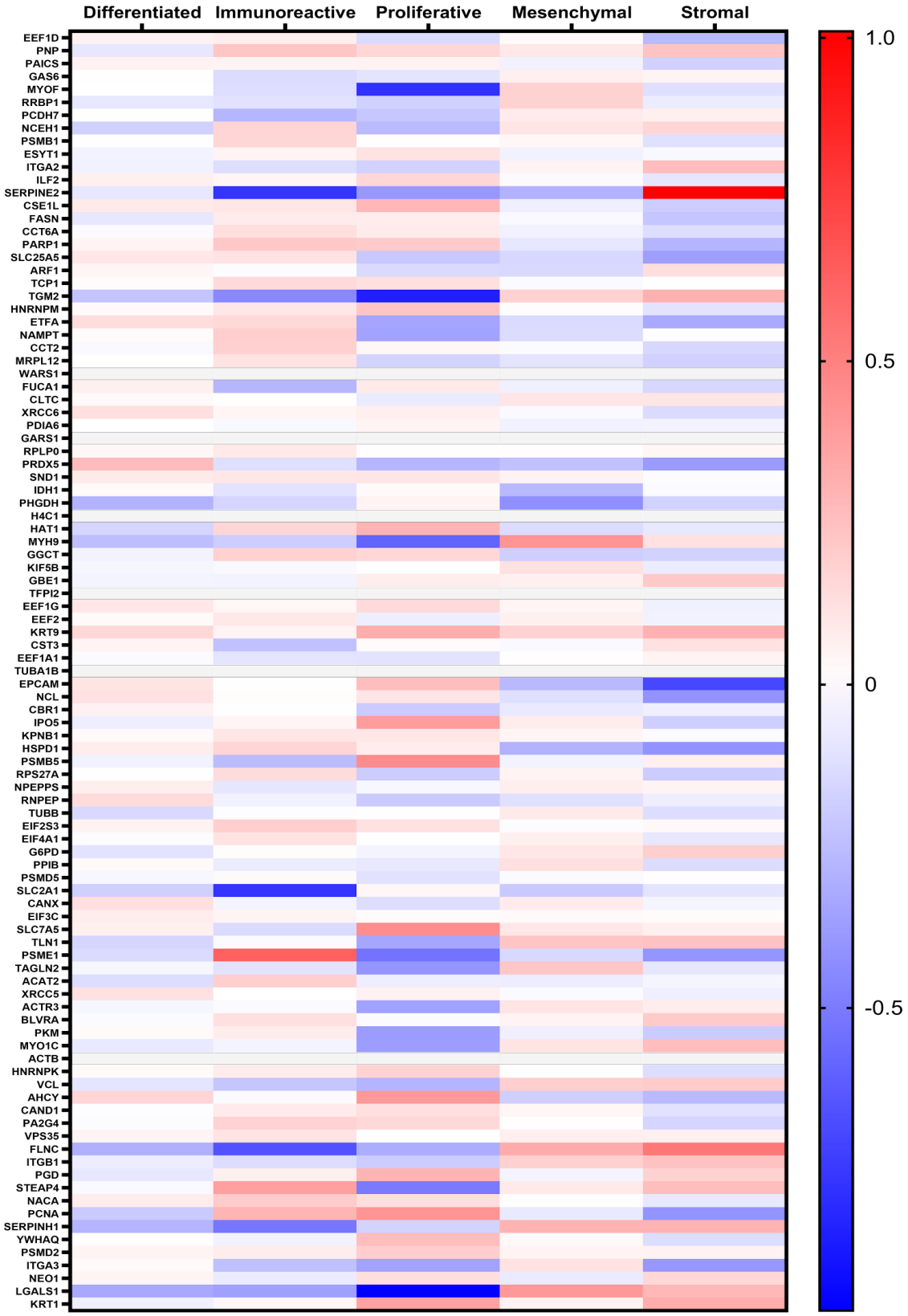
4.7 - Evolução da assinatura proteômica do subtipo mesenquimal para o câncer de ovário

Os alvos obtidos a partir das proteínas superexpressas na cultura de células Caov-3 serviram como guia na seleção de alvos para esta análise, onde, desta vez, utilizamos a assinatura do subtipo mesenquimal dos dados proteômicos obtidos pelo CPTAC. Para esta análise a referência de assinatura é a integração dos dados proteômicos próprios de cultura de células e externos, também proteômicos do consórcio CPTAC, por sua vez obtidos de tumores de pacientes diagnosticados com câncer de ovário, mais especificamente do tipo seroso de alto grau.

A seleção de alvos para monitoramento foi feita considerando apenas as proteínas aumentadas no subtipo mesenquimal, então fizemos o levantamento da expressão média de nossa lista de 99 proteínas superexpressas no secretoma da cultura de células Caov-3, no conjunto de dados proteômicos de cada um dos subtipos obtida pelo CPTAC, sendo eles, diferenciado, imunorreativo, mesenquimal e estromal (Figura 14).

Figura 14: **Identificação de proteínas por assinaturas proteômicas de subtipos moleculares.** Expressão proteica organizada selecionando apenas as proteínas superexpressas no secretoma da linhagem Caov-3, para comparação dos cinco subtipos moleculares dos tumores HGSC obtidos pelo CPTAC. Sendo a primeira do centro JHU e a segunda do centro PNNL. Continua ►

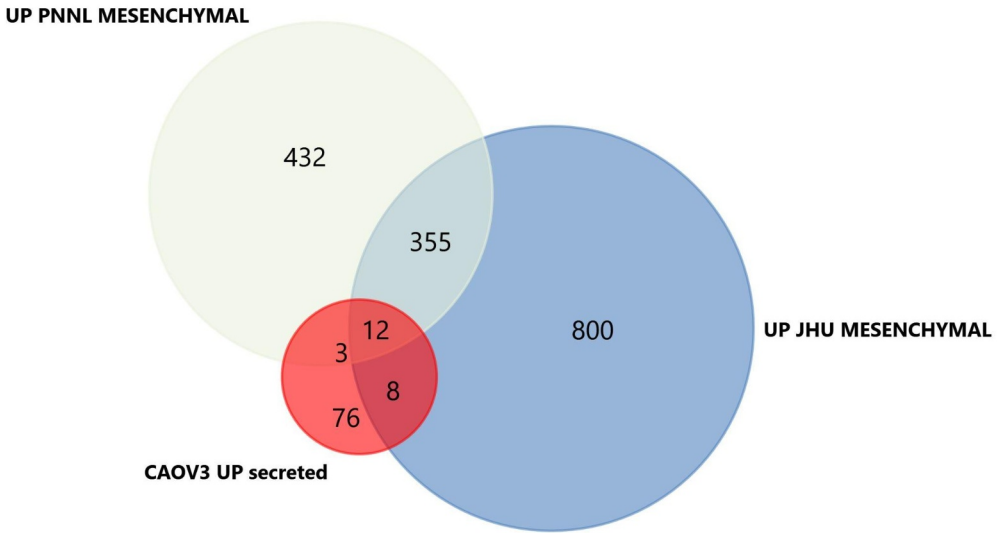



Differentiated
Immunoreactive
Proliferative
Mesenchymal
Stromal
EEF1D
PNP
PAICS
GAS6
MYOF
RRBP1
PCDH7
NCEH1
PSMB1
ESYT1
ITGA2
ILF2
SERPINE2
CSE1L
FASN
CCT6A
PARP1
SLC25A5
ARF1
TCP1
TGM2
HNRNPM
ETFA
NAMPT
CCT2
MRPL12
WARS1
FUCA1
CLTC
XRCC6
PDIA6
GARS1
RPLP0
PRDX5
SND1
IDH1
PHGDH
H4C1
HAT1
MYH9
GGCT
KIF5B
GBE1
TFPI2
EEF1G
EEF2
KRT9
CST3
EEF1A1
TUBA1B
EPCAM
NCL
CBR1
IPO5
KPNB1
HSPD1
PSMB5
RPS27A
NPEPPS
RNPEP
TUBB
EIF2S3
EIF4A1
G6PD
PPIB
PSMD5
SLC2A1
CANX
EIF3C
SLC7A5
TLN1
PSME1
TAGLN2
ACAT2
XRCC5
ACTR3
BLVRA
PKM
MYO1C
ACTB
HNRNPK
VCL
AHCY
CAND1
PA2G4
VPS35
FLNC
ITGB1
PGD
STEAP4
NACA
PCNA
SERPINH1
YWHAQ
PSMD2
ITGA3
NEO1
LGALS1
KRT1
1.0
0.5
0
-0.5

Apesar das diferentes intensidades, é possível observar concordância nos dados das duas instituições para expressão das proteínas em cada subtipo. No entanto, não é possível observar uma correlação clara entre o conjunto das proteínas selecionadas e sua expressão ligada a um único subtipo específico, porém várias delas estão aumentadas no subtipo mesenquimal, e estas foram utilizadas para as análises posteriores, compondo a assinatura mesenquimal.

Os novos alvos da assinatura mesenquimal foram os que apresentaram maior expressão média no subtipo mesenquimal, assim, verificamos a sobreposição dos conjuntos de proteínas reguladas em cada conjunto de dados com as proteínas identificadas no secretoma do modelo de cultura celular.

Figura 15: **Diagrama de sobreposição de proteínas superexpressas no secretoma com proteínas mais expressas no subtipo mesenquimal (CPTAC)**. Em vermelho proteínas identificadas no secretoma, Em cinza e azul proteínas identificadas nos centros referência, PNNL e JHU respectivamente. Apenas proteínas com maior média de expressão no subtipo mesenquimal foram selecionadas.

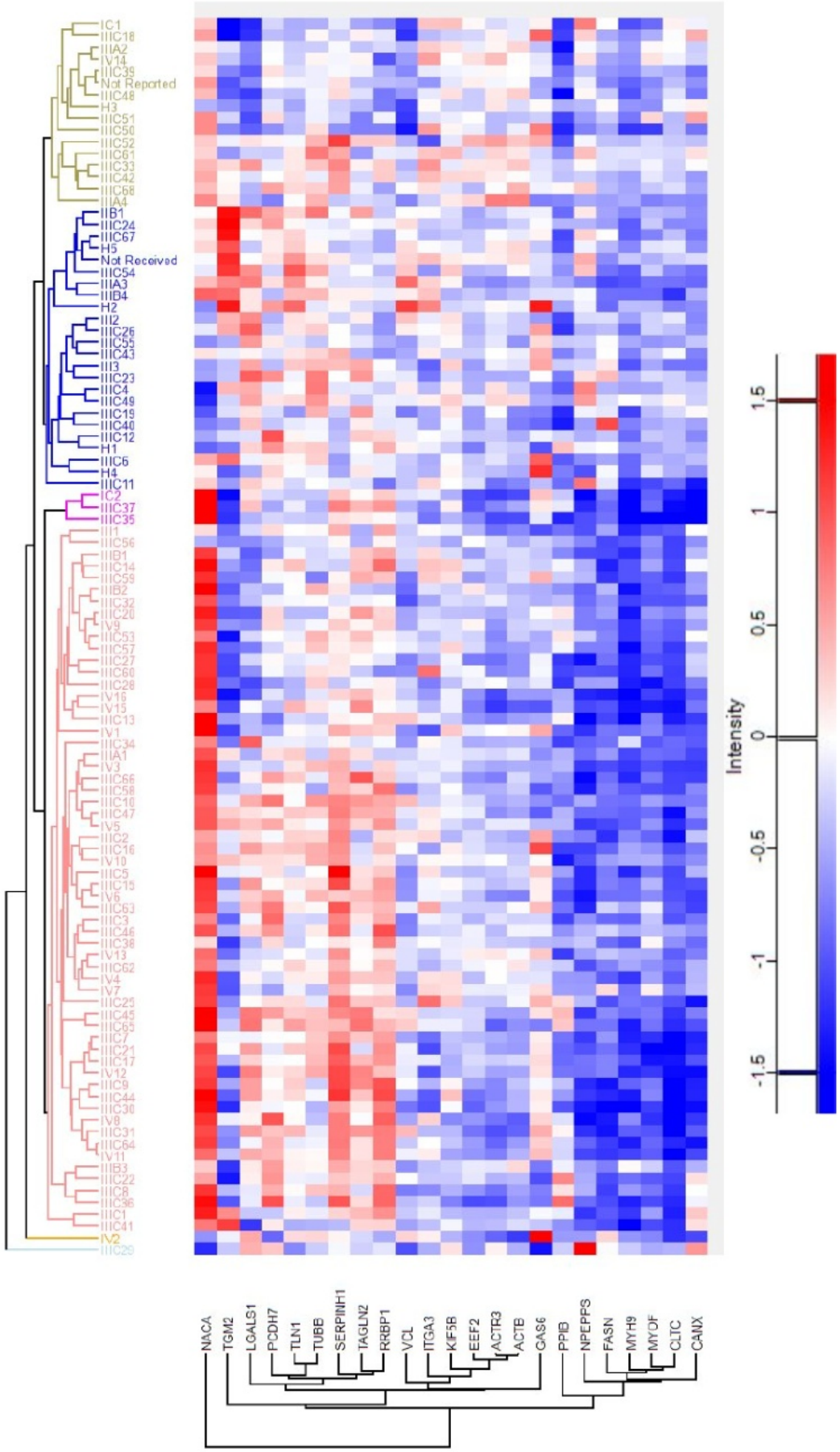


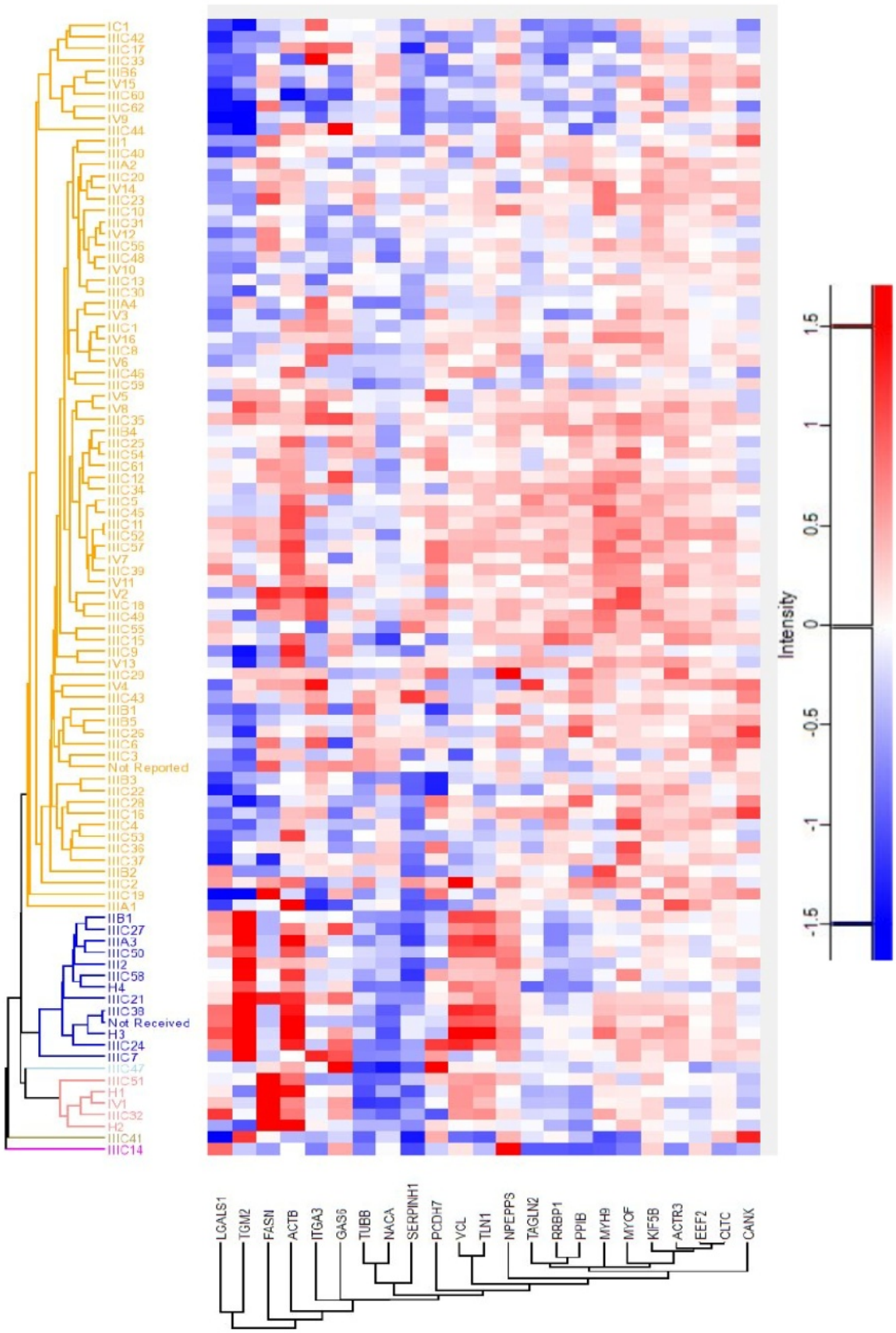
O diagrama de Venn da figura 15 apresenta sobreposição de 23 proteínas entre o conjunto de dados completo proveniente do CPTAC e o nosso modelo, com a maioria

delas compartilhadas entre os dados das duas instituições, sendo estes concordantes, como observado anteriormente.

A avaliação da capacidade de estratificação deste painel, resultante da integração de dados proteômicos, foi testada com um segundo conjunto de dados obtidos do consócio CPTAC. Este outro conjunto de dados não acompanha informação de subtipo, apenas de grau e estágio dos tumores, desta forma a assinatura mesenquimal foi avaliada em sua capacidade de agrupar estes parâmetros na figura 16.

Figura 16: **Avaliação da assinatura molecular proteômica referente ao subtipo mesenquimal em amostras tumorais de pacientes com câncer de ovário.** Obtidas de dados do CPTAC. No dendograma vertical estágios dos casos avaliados. Continua ►




IC1
IIIC42
IIIC17
IIIC33
IIIB6
IV15
IIIC60
IIIC62
IV9
IIIC44
III1
IIIC40
IIIA2
IIIC20
IV14
IIIC23
IIIC10
IIIC31
IV12
IIIC56
IIIC48
IV10
IIIC13
IIIC30
IIIA4
IV3
IIIC1
IV16
IIIC8
IV6
IIIC46
IIIC59
IV5
IV8
IIIC35
IIIB4
IIIC25
IIIC54
IIIC61
IIIC12
IIIC34
IIIC5
IIIC45
IIIC11
IIIC52
IIIC57
IV7
IIIC39
IV11
IV2
IIIC18
IIIC49
IIIC55
IIIC15
IIIC9
IV13
IIIC29
IV4
IIIC43
IIIB1
IIIB5
IIIC26
IIIC6
IIIC3
Not Reported
IIIB3
IIIC22
IIIC28
IIIC16
IIIC4
IIIC53
IIIC36
IIIC37
IIIB2
IIIC2
IIIC19
IIIA1
IIB1
IIIC27
IIIA3
IIIC50
III2
IIIC58
H4
IIIC21
IIIC38
Not Received
H3
IIIC24
IIIC7
IIIC47
IIIC51
H1
IV1
IIIC32
H2
IIIC41
IIIC14
LGALS1
TGM2
FASN
ACTB
ITGA3
GAS6
TUBB
NACA
SERPINH1
PCDH7
VCL
TLN1
NPEPPS
TAGLN2
RRBP1
PPIB
MYH9
MYOF
KIF5B
ACTR3
EEF2
CLTC
CANX
Intensity
1.5
1
0.5
0
-0.5
-1
-1.5

A aplicação da assinatura proteômica mesenquimal desenvolvida mostrou em geral agrupamento uniforme entre amostras de mais alto grau e estágio, enquanto amostras de tecidos saudáveis e de menor grau se separaram sem uma distinção definida de perfis de expressão, contando com apenas algumas exceções.

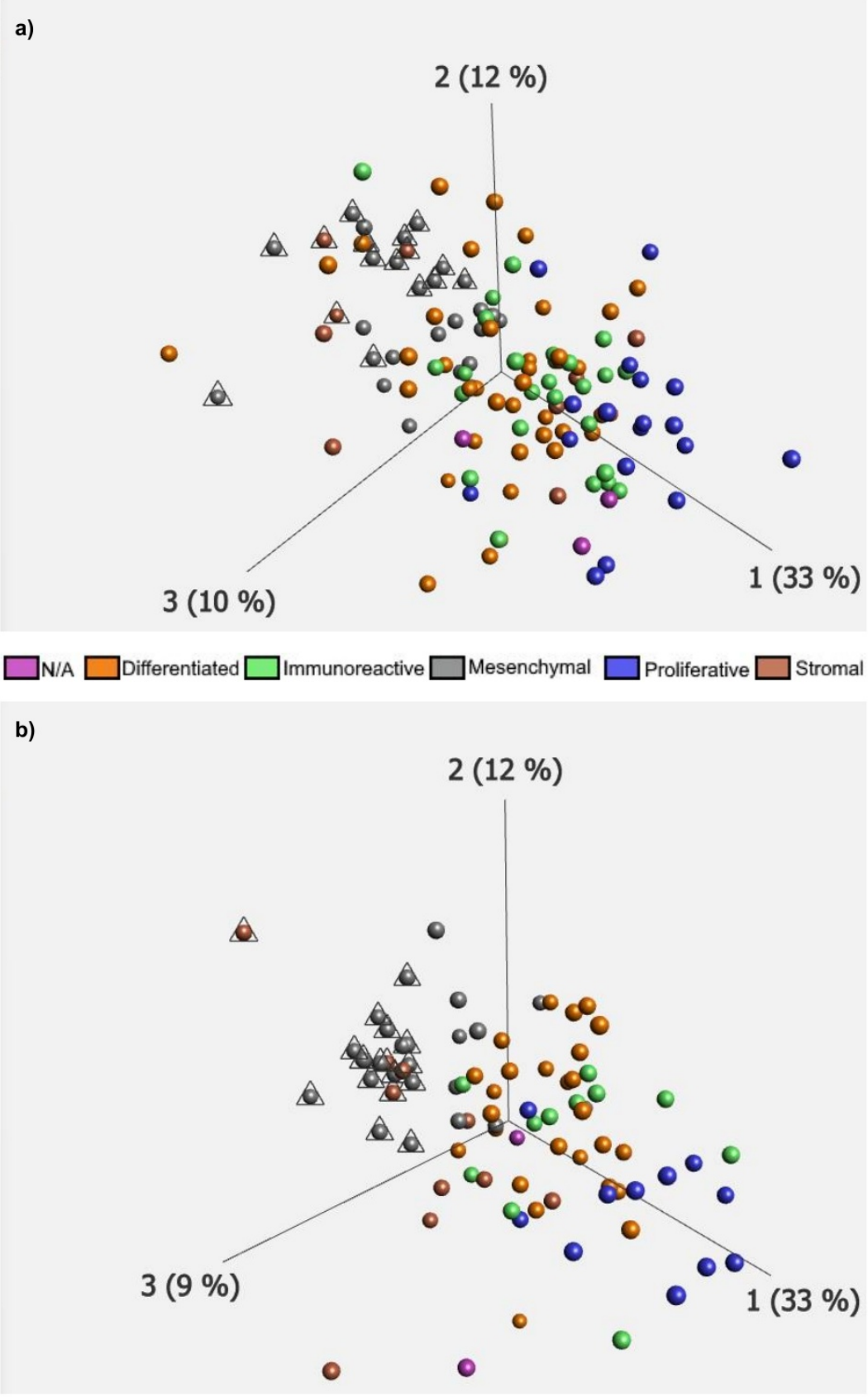
O agrupamento foi limitado a seis grupos, pois a separação em mais grupos, gera somente outros grupos de amostra única, devido à homogeneidade dos dados de expressão.

4.8 - Análise por agrupamento e classificadores das assinaturas moleculares

Após a obtenção de assinaturas relevantes na identificação do fenótipo mesenquimal, baseada na identificação de proteínas reguladas do processo da EMT, iniciamos a busca por alvos sem este viés, ou seja, contando apenas com a expressão obtida de todas as proteínas identificadas. Para tal utilizamos os dois conjuntos de dados mencionados anteriormente, sendo o primeiro com informação de subtipos moleculares, e o segundo apenas com informação de estágio.

A primeira abordagem utilizada foi a projeção em PCA com agrupamento por k-means, realizada nos dois conjuntos de dados (JHU e PNNL). A projeção é realizada com a contribuição de todas as proteínas identificadas na amostra, sem seleção. A formação de grupos pelo algoritmo k-means cria grupos baseados na proximidade destes pontos, resultante da similaridade de sua expressão.

Figura 17: **Projeção de PCA das amostras dos diferentes subtipos moleculares do câncer de ovário de alto grau.** a)JHU. b)PNNL. Cada esfera representa um caso avaliado, coloridos por subtipo molecular e posicionados de acordo com diferentes níveis de proteínas. Em destaque(triângulo) agrupamento por algoritmo k-means coincidente com subtipo mesenquimal.



A distribuição das amostras, em pontos na projeção em PCA tridimensional apresentou regiões em sua maioria indistintas em relação aos subtipos, com maior concentração em pólos opostos dos subtipos mesenquimal e proliferativo, ainda assim sem clara definição. O agrupamento dimensionado para cinco grupos, devido aos cinco subtipos, mostrou um dos grupos, destacado com triângulos, altamente coincidente com o subtipo mesenquimal, fato observável nos dois conjuntos de dados

Em seguida, outra abordagem utilizada para obtenção de assinaturas relevantes ao subtipo mesenquimal foi o treinamento e aplicação de classificadores. Nesta abordagem, um primeiro conjunto de dados, com classificação em subtipos foi usado para o treinamento do classificador, então este classificador (compreendendo 547 variáveis, no caso as proteínas identificadas) foi aplicado em um segundo conjunto de dados, sem classificação em subtipos.

A projeção resultante mostrou a probabilidade, segundo o algoritmo de classificação Random trees, de cada amostra pertencer ao subtipo mesenquimal. Pode ser notada uma concentração das maiores probabilidades em uma região, com destaque em triângulos para as amostras de maior estágio coincidindo parcialmente.

Figura 18: **Projeção em PCA de amostras do CPTAC classificadas por algoritmo Random trees.** Com maior probabilidade do subtipo mesenquimal em verde e menor em vermelho, destacando (triângulo) subgrupo de amostras de mais alto estágio do câncer de ovário. Modelo e teste apenas com dados JHU.

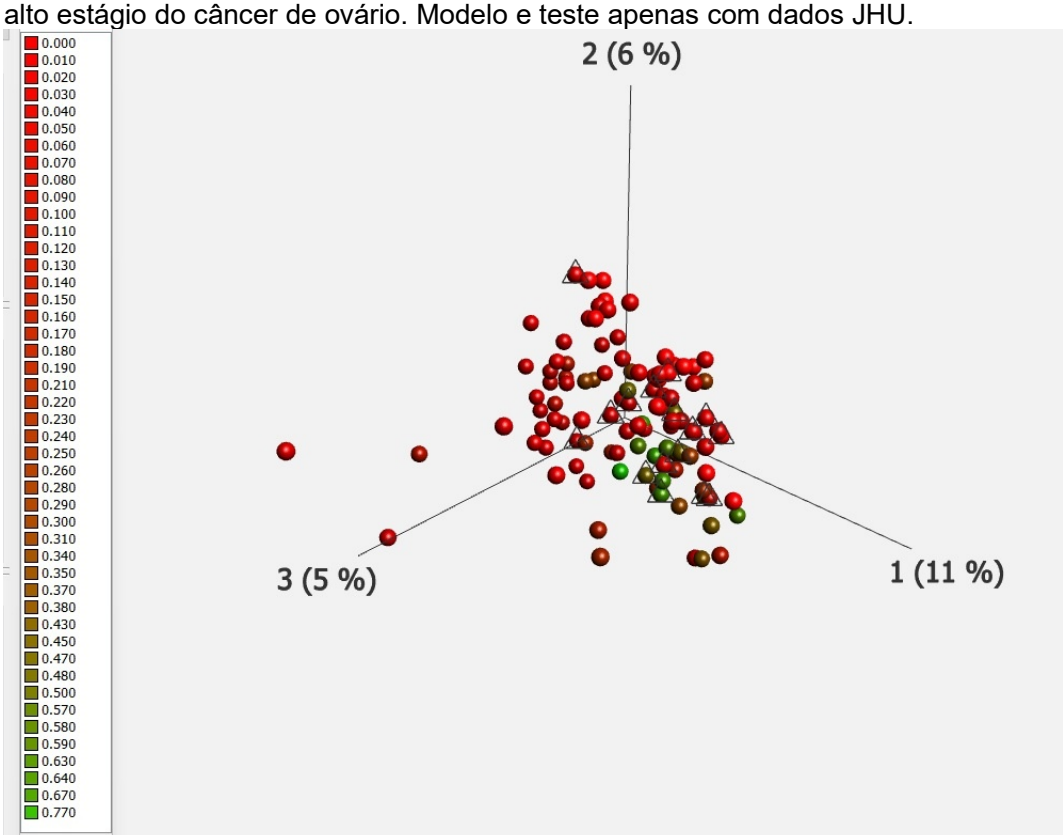


Tabela 6: Variáveis(proteínas) utilizadas pelo algoritmo Random trees na composição do modelo probabilístico. Continua ►

Protein Gene					
A1BG	CHD4	FZD1	LUZP1	POU2F1	SNRPB2
ABR	CHD7	GABPA	LXN	PPM1F	SNRPC
ACIN1	CILP	GALNT1	LYN	PPP1R18	SNRPD2
ACSL1	CLTB	GBP1	LYZ	PPP1R9B	SNRPD3
ACTA1	CNN1	GBP2	MAGOH	PRKG1	SNRPF
ACTG2	CNN2	GC	MAP1A	PROM2	SNRPG
ACTN1	CNN3	GFPT2	MAPRE3	PROS1	SNX18
ACTR2	CNPY4	GIMAP1	MARCKS	PRPF40A	SNX9
ACTR3	CNRIP1	GIMAP4	MARCKSL1	PSMB10	SOD2
ADNP	COL11A1	GIT2	MATR3	PSMB5	SON
AEBP1	COL12A1	GMFG	MBD3	PSMB6	SORBS1
AFM	COL1A1	GNAI2	MICAL1	PSMB8	SORBS3
AGT	COL1A2	GNB4	MMP11	PSMB9	SPAG9
AHSG	COL3A1	GPX8	MMP14	PSME1	SPARC
AIF1	COL5A1	GSN	MMP2	PSME2	SPATS2L
ALB	COL5A2	GTF3C4	MNDA	PTGIS	SRPX2
ALDH1A3	COL6A1	GYG1	MPP1	PTK7	SRRT
ALOX5	COL6A2	HABP2	MPRIP	PTPRC	SRSF1
ALOX5AP	COL6A3	HCK	MRC2	PYCARD	SRSF11
AMBP	COL8A1	HCLS1	MRVI1	RAB11FIP5	SRSF3
AMPD3	COLEC12	HDLBP	MSH2	RAB31	SRSF5
ANKRD44	COMMD3-BMI1	HEPH	MSH6	RAB32	SRSF9
ANTXR1	COPG1	HK3	MSN	RAC2	SSRP1
ANTXR2	COPZ2	HLA-DRA	MVP	RAI14	STK10
ANXA2	CORO1A	HNRNPA1	MXRA5	RAP2B	STOM
AOC3	CORO1C	HNRNPC	MYH10	RARRES2	SUDS3
APBB1IP	CP	HNRNPH3	MYH9	RBBP5	SULF1
APCS	CPSF6	HNRNPL	MYL12A	RBM14	SYNE1
APOA1	CPZ	HNRNPLL	MYL6	RBM15	SYNPO
APOA2	CRTAP	HNRNPM	MYL9	RBM39	TAGLN
APOB	CSRP2	HNRNPR	MYLK	RBMX	TAGLN2
APOF	CTHRC1	HNRNPU	MYO1D	RCN3	TAP1
APOH	CTSK	HOPX	MYO1E	RCOR2	TAP2
APOL1	CYBA	HPX	MYO1F	RCSD1	TARDBP
ARHGAP18	CYBB	HRG	MYO1G	RENBP	TBC1D2B
ARHGAP45	DAB2	HSPG2	MYO5A	RFTN1	TBXAS1
ARHGDIB	DAPK3	HTRA1	NAB2	RHOG	TFCP2
ARID2	DCN	IARS2	NCF1	RILPL1	TGFB1
ARPC1B	DDX17	ICE1	NCF2	RNASE4	TGFB1I1
ARPC2	DEK	IFT81	NCKAP1L	RPRD2	TGFBI
ARPC3	DHX9	IGHV1-69-2	NEXN	RSU1	THBS1
ARPC4	DOCK10	IGLL1	NNMT	S100A10	THBS2
ARPC5	DOCK2	IGLL5	NPEPL1	S100A16	THBS3
ARSB	DOCK8	IK	NR2F2	S100A8	THRAP3
ASH2L	DPF2	IKBIP	NRF1	S100A9	THY1
ASPN	DPYD	IL16	NRP2	SAA2-SAA4	TIMP2

ATP6V1B2	DPYSL3	ILF2	NT5E	SAP18	TLN1
AZGP1	DRG1	ILK	NUDT21	SARNP	TLR2
B2M	DUSP3	INPP5D	NXF1	SASH3	TMPO
BAZ1B	ECM1	IQGAP1	OLFML3	SEC23A	TMTC3
BGN	EFEMP2	ITGA1	ORC2	SEC24D	TNC
BIN2	EHMT2	ITGA11	ORM1	SELENOP	TNPO3
BRD8	ELMO1	ITGA5	ORM2	SEPT11	TPM1
BST1	EMILIN1	ITGAL	OSTF1	SEPT2	TPM2
BTK	ENG	ITGAM	P3H1	SEPT5	TPM4
BUB3	ENPP1	ITGAV	P3H3	SEPT7	TPR
BUD31	ENTPD1	ITGAX	P3H4	SERPINA1	TPSAB1
C17orf62	EP400	ITGB2	P4HA2	SERPINA6	TPX2
C1QB	EPB41L3	ITGB5	PAF1	SERPINA7	TRA2B
C1QC	ERH	ITIH1	PALLD	SERPINC1	TRADD
C1R	EVL	ITIH2	PAPSS2	SERPIND1	TRIM21
C1S	F12	ITIH3	PARP4	SERPINF1	TRIM33
C2	F13A1	ITIH4	PARVA	SERPINF2	TRMT112
C3	FAM114A1	KANK2	PBRM1	SERPINH1	TRRAP
C4BPA	FAM129A	KCNAB2	PCOLCE	SF1	TSPO
C4BPB	FAP	KDELC1	PDGFRB	SF3A1	TYMP
C5	FBLIM1	KDM1A	PDLIM2	SF3A3	UBN1
C6	FBLN1	KHDRBS1	PDLIM3	SFPQ	UBR5
C7	FBLN2	KLKB1	PDLIM4	SFRP4	VASP
C8A	FBN1	KNG1	PDLIM5	SFSWAP	VAV1
C8B	FCER1G	KYNU	PDLIM7	SFXN3	VCAM1
C9	FCGR1A	L3HYPDH	PECAM1	SGCD	VCAN
CALB2	FCGR2B	LACTB	PHC1	SGCE	VIM
CALD1	FCGR3A	LAMA4	PHF10	SH3D19	VIRMA
CALHM5	FCN2	LAMB1	PHLDB1	SH3GL1	WARS
CARD16	FERMT2	LAP3	PICALM	SH3GLB1	WAS
CASP1	FERMT3	LCP2	PLAU	SH3PXD2B	WBP11
CASP4	FHL2	LEO1	PLCG1	SIN3A	WDR5
CAVIN3	FHL3	LGALS1	PLEC	SIRPA	XPO1
CCDC80	FILIP1L	LIG3	PLEK	SKAP2	XRCC1
CD14	FKBP11	LMCD1	PLEKHO2	SMARCA4	YEATS4
CD44	FKBP7	LMF2	PLG	SMARCA5	ZC3H15
CD74	FLNA	LMNA	PLOD2	SMARCB1	ZMYM3
CD84	FLNC	LMOD1	PLS3	SMARCC1	ZMYND8
CDC42EP1	FMNL1	LOX	PLXDC2	SMARCC2	ZNF207
CDC73	FN1	LPP	PLXNC1	SMARCD1	ZNF638
CDH11	FNBP1	LPXN	PNN	SMARCE1	ZYX
CDK2AP1	FNDC1	LRG1	POLR2A	SNRNP200	
CFB	FNDC3B	LRP1	POLR2B	SNRNP70	
CFH	FSCN1	LRRC15	POLR2C	SNRPA	
CFHR5	FUS	LSP1	POLR2G	SNRPA1	
CFI	FYB1	LUM	POSTN	SNRPB	

5 -Discussão

5.1 - Identificação de proteínas no secretoma de células Caov-3

O câncer de ovário, como em outras neoplasias malignas, tem a metástase como característica marcante de sua agressividade, assim a compreensão do desenvolvimento da metástase, desde o início, juntamente de suas implicações é relevante para compreensão e proposição de meios para sua detecção em estágios iniciais, prevenção e terapias. Portanto o processo da transição epitélio-mesenquimal que marca o início da diferenciação celular e posterior aquisição de características que favorecem sua mobilidade, capacidade de invasão, diferenciação e proliferação, é de interesse fundamental.

O modelo de EMT desenvolvido anteriormente e abordado neste trabalho se dá pela manutenção de células de uma linhagem de adenocarcinoma de ovário Caov-3, já comprometida em algumas de suas capacidades para manutenção da homeostase, como exemplo, amplificação de receptores (NOTCH) e deficiência na supressão tumoral pela expressão defectiva de mRNA (TP53, STK11), características observadas regularmente em tumores de pacientes [43]. Ainda estimuladas com fator de crescimento EGF (10 ng/mL), uma concentração superior a fisiológica [44], com o objetivo de exacerbar o estímulo.

A sinalização do receptor do fator de crescimento epidérmico (EGFR) regula vários eventos de desenvolvimento no ovário. A ligação do EGF ao EGFR inicia várias vias de sinalização, incluindo vias de fosfatidilinositol-3-quinase (PI3K)/AKT e proteína quinase ativada por mitógeno (MAPK)/ERK. A desregulação da sinalização de EGFR promove a progressão do câncer de ovário e está associada a um fenótipo agressivo e mau prognóstico dos tumores ovarianos [45].

A estratégia quantitativa usada para comparação dos efeitos obtidos pelo tratamento na cultura celular com a marcação metabólica por assimilação de (C-Lys$^{13}C_6$) identificou proteínas superexpressas. Porém como são necessárias algumas passagens o tratamento direto destas células(marcadas) com fator de crescimento EGF comprometeu sua viabilidade e eventualmente suas características morfológicas e fenotípicas, interessantes ao estudo. Então este controle quantitativo foi feito de forma paralela e não sequencial, ou seja, as células tratadas com fator de crescimento não passaram por marcação isotópica, apenas tiveram sua identificação de proteínas comparada, assim como a cultura controle (sem tratamento por EGF), também apenas com sua identificação de proteínas comparada.

Outra característica da identificação de proteínas superexpressas neste trabalho foi a determinação do valor de corte para consideração da superexpressão, relativamente baixo, justificado por dois fatores, primeiro, principalmente, a diluição do sobrenadante de cultura a cada substituição do meio, onde este foi somado a cada troca e não separado para análises individuais, o que demandaria muito mais cultivos simultâneos para obtenção de material suficiente. Também pela metodologia adotada para identificação de proteínas e sua quantificação, com a digestão tríptica temos, idealmente, uma população de peptídeos enriquecida com aminoácidos lisina e arginina em C-terminal, contudo dado o caráter quantitativo, citado anteriormente, apenas peptídeos com lisina contaram para a quantificação.

As adaptações na estratégia de identificação proteômica citadas enriqueceram a confiança nos dados pela realização das mesmas identificações em diferentes experimentos, também pela adoção de critérios qualitativos e quantitativos, como as características dos peptídeos identificados e cobertura das proteínas. Esta identificação rendeu a determinação de 99 proteínas superexpressas no secretoma das células Caov-3

durante o processo de EMT. O processo da EMT em si foi confirmado morfologicamente pela observação e fenotipicamente de forma parcial pela expressão de marcadores clássicos da EMT, de forma significativa apenas SNAI1 e N-Caderina [16], porém isto indica, de fato a ocorrência do caráter de uma transição, onde há coexistência de marcadores dos dois fenótipos, ainda afetada pelas diferentes populações celulares que podem estar em etapas diversas do processo.

As proteínas superexpressas identificadas no secretoma possuem baixa coincidência de proteínas superexpressas em subfrações celulares, indicando a conservação do secretoma durante sua obtenção, questão recorrente em análises do tipo, por vezes identificadas com produto de lise e contaminação por soro, sendo esta última desconsiderada pela marcação isotópica metabólica.

As proteínas secretadas demonstram grande relevância no papel da comunicação célula-célula e sinalização, que resultam no comportamento destas e portanto em suas condições em diferentes estágios das doenças. Logo, modelos desses ensaios são um recurso rico na pesquisa para identificar potenciais biomarcadores no sangue, uma fonte preferencial de testes. Embora os fluidos biológicos estejam disponíveis, eles apresentam obstáculos à análise devido à grande variedade de espécies e concentrações de proteínas, e os potenciais biomarcadores geralmente estão presentes em baixas concentrações[46,47].

A cultura de células simula apenas parcialmente as condições encontradas *in vivo*, os testes de descoberta em estágio inicial são aplicáveis porque têm o potencial de indicar a composição das secreções para se ter noção de vias e processos biológicos alterados. De fato, as células cancerosas interagem com seu ambiente para adquirir um fenótipo maligno caracterizado por crescimento descontrolado, agressividade, evasão de resposta imune e similares. Nesses casos, seu sucesso depende de uma série de ações

desencadeadas por proteínas secretadas, tornando-as alvos valiosos como biomarcadores e ação de drogas. Portanto, estudos com meios condicionados se mostraram um método simples e direto para a obtenção desses proteomas [48].

A importância do papel do microambiente tumoral foi verificada por Pearce et al. 2018, através de análises derivadas de genômica e proteômica, utilizando amostras de pacientes HGSC com grande abrangência de estados patológicos, variando de lesões localizadas a grande comprometimento metastático, para determinação de um perfil de progressão da doença centrado em moléculas comuns a este meio, como citocinas e componentes da matriz extracelular[49].

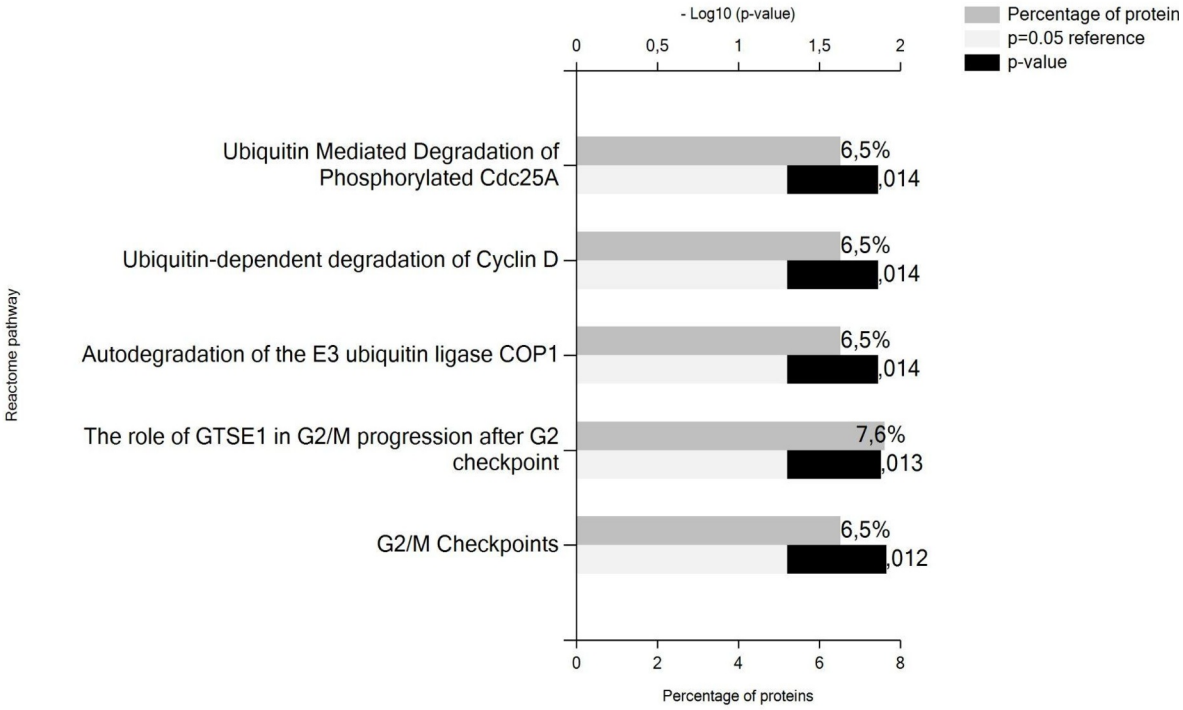
Notavelmente a maioria das proteínas identificadas são comumente encontradas e relatadas como transportadas por vesículas extracelulares. As vesículas extracelulares são alvos emergentes em análises proteômicas e apresentam grande relevância como carreadoras de proteínas, mRNAs, entre outras biomoléculas ligadas a sinalização e/ou regulação, potencialmente capazes de modular o microambiente tumoral e induzir processos intracelulares como a própria EMT, sendo que a função de cada molécula depende ainda de sua localização celular, justificando um aprofundamento particular no estudo destas vesículas.

O ponto de conexão de toda a rede de interação das proteínas (nó) é a proteína ribossomal RPS27A (ribosomal protein S27a ubiquitin-40S), expressa com uma extensão do polipeptídeo Ubiquitina em C-terminal. Há referências desta proteína em diversas funções reguladoras importantes, como maturação do ribossomo, reparo do DNA, regulação celular, ativação de quinases e sinalização celular. RPS27A também regula o receptor EGFR [50,51].

Dentre as proteínas identificadas, um conjunto relevante vinculado à diferentes processos se destaca pela complementação de achados na análise de subfrações

celulares, onde foi identificado o bloqueio do ciclo celular na fase G1 após indução da EMT, identificado por Grassi et al. 2016, justamente observamos entre as proteínas superexpressas no secretoma PSMB1, PSMB5, PSMD2, PSMD5, PSME1 e RPS27A, componentes do proteassoma responsáveis, entre outras, pelas funções identificadas na figura abaixo.

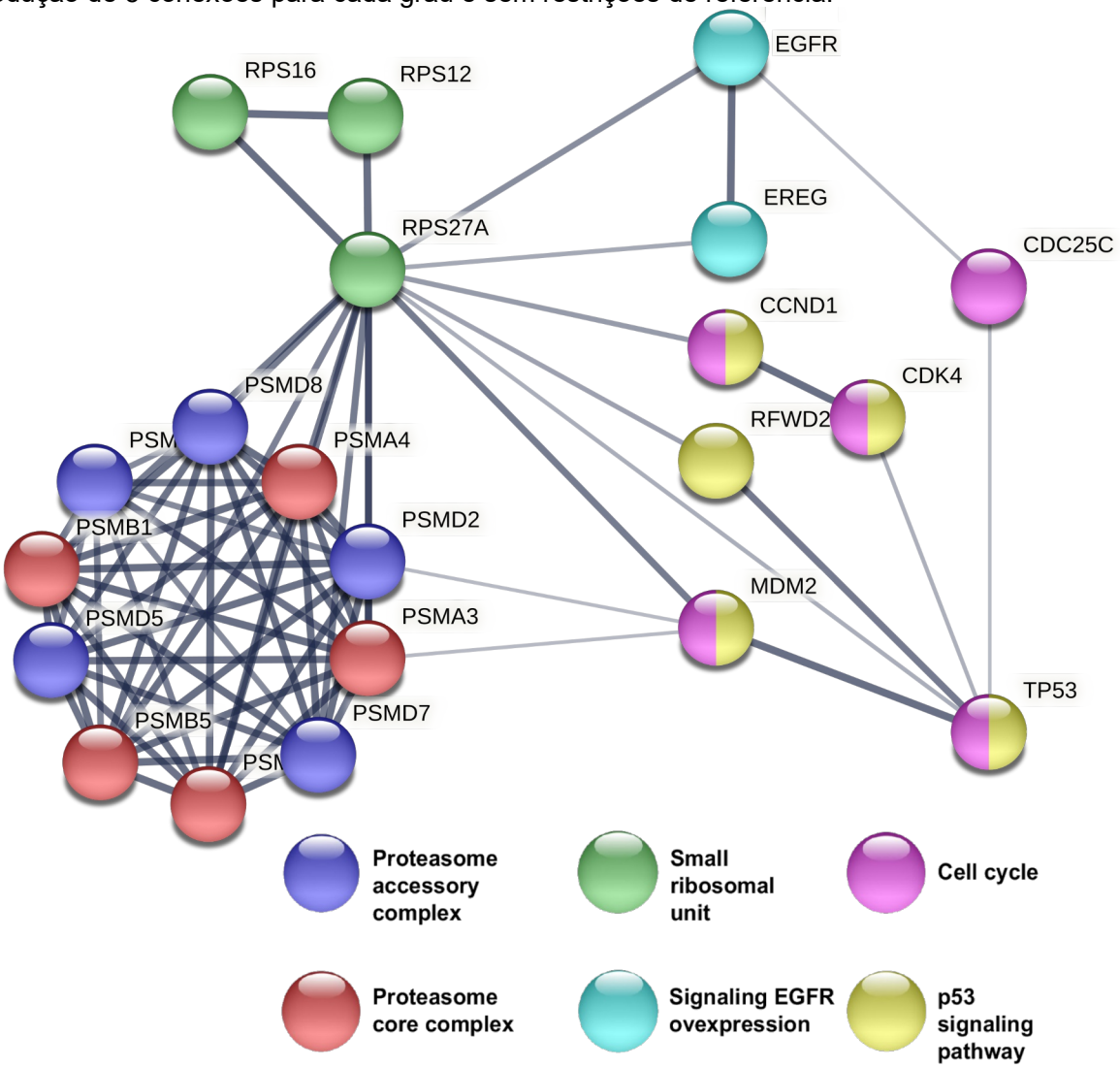
Figura 19: **Associação de proteínas selecionadas identificadas no secretoma com funções do reactoma.** Obtida pelo *software* Funrich, mostrando a relevância de cada função no conjunto total das proteínas identificadas.



Estas funções associadas à degradação de CDC25A, CCND1, COP1, estão diretamente ligadas aos processos essenciais do ciclo celular e reparação de danos no DNA, respectivamente: i) CDC25A controla ciclinas promotoras do processo mitótico (CDK1, CDK2, CCNE1); ii) CCND1 em complexo com CDK4 para controle da mitose e regulação de RB1 para controle da transição G1/S; COP1 mediador da ubiquitinação e degradação de TP53, ponto de controle da transição G1/S, que, como mencionado anteriormente, é defectivo nesta linhagem. A seguir uma representação das interações foi

montada com proteínas identificadas e seus respectivos alvos para ilustrar a disposição do controle celular mencionada.

Figura 20: **Rede de interações montada com proteínas selecionadas identificadas no secretoma.** Apenas proteínas selecionadas foram introduzidas na aplicação STRING-DB, com introdução de 5 conexões para cada grau e sem restrições de referência.

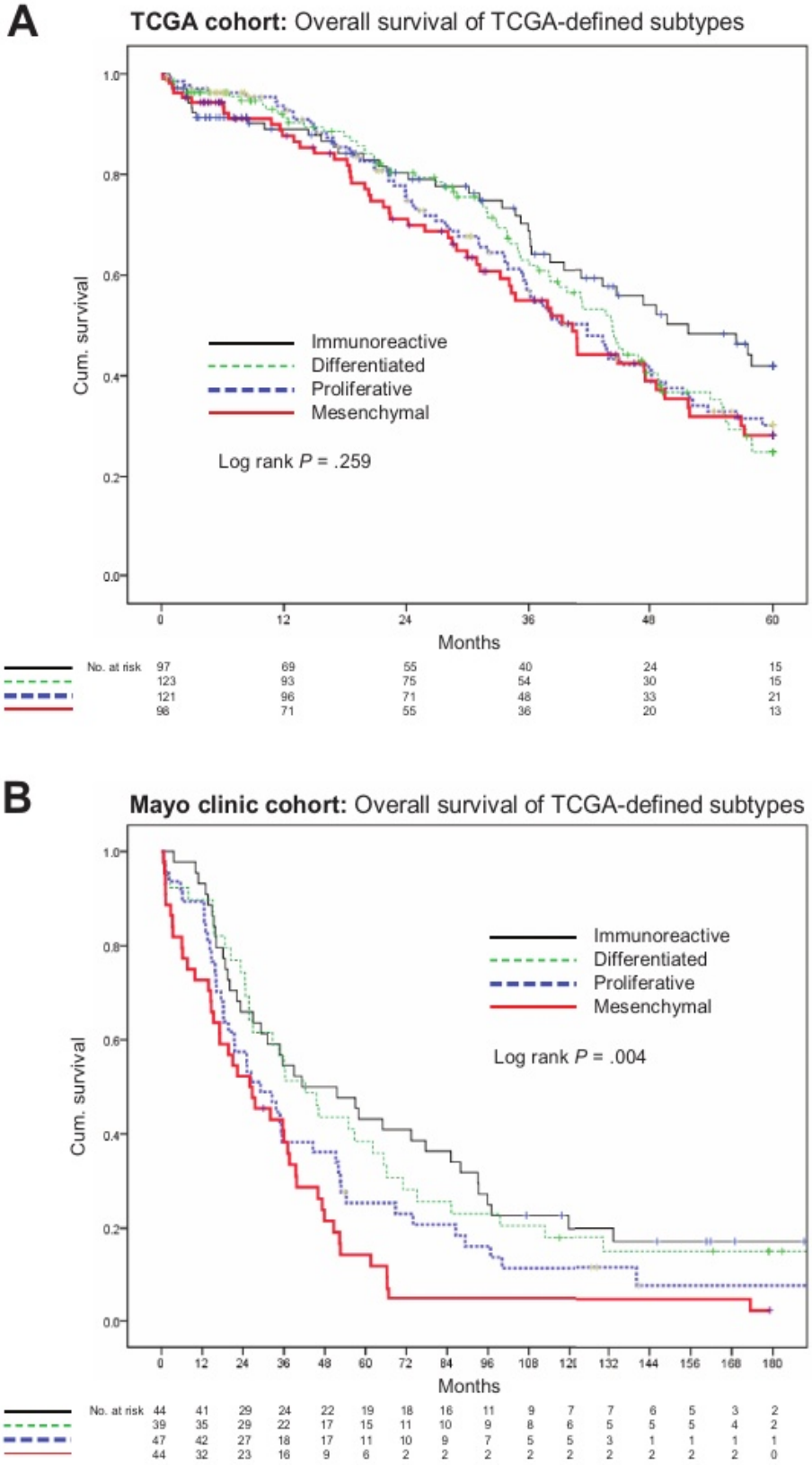


5.2 - Painéis de assinatura mesenquimal partindo de proteínas superexpressas no secretoma

A realização da indução da EMT objetiva a representação de uma etapa crítica para a progressão tumoral, a aquisição de capacidades essenciais para metástase, neste desenvolvimento as células adquiriram um fenótipo com características mais próximas de células mesenquimais, então, neste trabalho, as proteínas identificadas no secretoma do modelo foram aplicadas na comparação com dados obtidos por consórcios, TCGA e CPTAC. com o objetivo comum de identificar perfis em tumores de pacientes. Este paralelo entre proteínas identificadas no secretoma das células “proto-mesenquimais” e identificações ômicas feitas por consórcios em tumores de pacientes se valeu do subtipo mesenquimal identificado em ambas expressões, por, entendermos, representar maior propensão à metástase, e consequentemente, risco aos pacientes.

Além disso, o conhecimento de subtipos moleculares de tumores em câncer de ovário pode ser relevante para o prognóstico, apesar do estudo original do TCGA não notar significância estatística na sobrevida. Outro estudo acompanhando pacientes da Mayo Clinic, com aplicação das mesmas assinaturas moleculares notou significância em um acompanhamento mais longo [52].

Figura 21: **Curvas de sobrevivência por subtipo molecular.** a) Curvas obtidas pelo estudo original do TCGA com 489 pacientes, mostrando baixa correlação entre sobrevivência e subtipos de tumores HGSC. b) Curvas obtidas pelo estudo da Mayo Clinic aplicando mesma classificação de subtipos determinada pelo TCGA no acompanhamento de 174 pacientes, mostrando relevância e correlação na classificação de subtipos de tumores HGSC (KONECKY, G. E. et al. 2014).



O primeiro painel emergente da combinação de dados próprios e externos (TCGA), passou pelo processo de desenvolvimento de métodos para proteômica dirigida que conta com diversas etapas para composição de um método suficientemente sensível para avaliação das amostras, tais como, seleção de peptídeos, escolha de fragmentos e otimização de parâmetros cromatográficos, resultando em robustez para emprego em grande número de amostras. O método contou com cobertura da maioria das proteínas propostas e excetuando alguns casos foi capaz de monitorar com reprodutibilidade dois peptídeos únicos de cada uma das 18 proteínas.

As amostras de tumores de ovário utilizadas foram de diversas histologias, diferentes tipos e estágios, sendo a maior parte de HGSC, refletindo a realidade da distribuição de tumores diagnosticados, onde este tipo predomina. As amostras GO16 e GO21 se mostraram pouco responsivas ao monitoramento de todos os peptídeos selecionados,. Então foram submetidas novamente ao processo de preparo e monitoramento, ainda assim, não houve alteração dos resultados, coincidente com sua exclusividade histológica entre as amostras, o que indica a capacidade do painel de segregar tumores de histologias distintas da HGSC. Além de agrupar os mais altos estágios, como no caso das amostras GO06, GO12, de tumor mucinoso e de endométrio respectivamente, logo não HGSC, porém de estágio muito avançado. Apoiando a relevância das proteínas selecionadas.

Apesar da avaliação limitada, produzida com baixo número de amostras, ser incapaz de validar estas proteínas (Figura 16) como alvos para diagnóstico ou prognóstico, sua capacidade de segregar tumores conforme suas propriedades justifica a aplicação em grupos maiores de amostras com variedade de tipos tumorais.

A divulgação dos dados do consórcio CPTAC, primeiro em 2017 com análises proteômicas conduzidas nas mesmas amostras já classificadas pelo TCGA com subtipos tumorais com base em perfis moleculares, e mais tarde em 2021 com a disponibilização de dados de um segundo conjunto de amostras, desta vez, sem subtipos tumorais, porém ambas com abordagem quantitativa, possibilitou novas estratégias de desenvolvimento de assinaturas e seu teste.

O primeiro conjunto de dados do CPTAC, com as mesmas amostras do TCGA, confirmou a presença de subtipos moleculares no padrão de expressão de proteínas, desta vez cinco, adicionando o perfil “Estromal”. As assinaturas de cada perfil foram avaliadas usando as proteínas identificadas superexpressas na análise do secretoma, revelando algumas proteínas interessantes (tabela 7) para a composição da nova assinatura mesenquimal.

O segundo painel emergente da combinação de dados próprios e externos, desta vez, o primeiro conjunto de dados do CPTAC, classificados por subtipo de tumor utilizando as proteínas superexpressas no secretoma foi construído com a seleção das proteínas mais expressas no subtipo mesenquimal, contando com algumas proteínas já observadas no painel anterior, reforçando a expressão dos alvos anteriores também no nível proteico. Este painel foi aplicado no segundo conjunto de dados do CPTAC, contendo amostras saudáveis e apenas a informação de estágio dos tumores.

Os resultados desta nova assinatura neste conjunto de amostras são interessantes, pois possibilitam a avaliação em um grande conjunto de amostras dificilmente acessível por iniciativas individuais. Integra amostras de tecido saudável e conta com vários estágios de tumores, porém agrega apenas amostras de tipo único de tumor, HGSC, não refletindo a real distribuição de tumores diagnosticados, o que pode

comprometer o desempenho dos marcadores selecionados. Ainda assim, observamos a estratificação de amostras saudáveis do grupo mais numeroso de amostras de mais elevado estágio.

Na tabela seguinte constam as principais proteínas usadas em cada painel e seu papel em diversos tipos de tumor.

Tabela 7: Principais proteínas utilizadas superexpressas no secretoma que compuseram as assinaturas. Obtidas em uma busca por referências recentes, priorizando artigos abertos, recentes e com análise de amostras de pacientes. Não foram encontradas referências para as proteínas NACA e NPEPPS com estes parâmetros.

Gene de referência	Tipo de tumor	Localização celular (geral)	Referência
Assinatura Proteogenômica (Caov3 – TCGA)			
FLNC	Breast, Head and Neck	cytoplasm [GO:0005737]; cytoskeleton [GO:0005856]; cytosol [GO:0005829]	53, 54
GBE1	Leukemia, Lung	cytoplasm [GO:0005737]; cytosol [GO:0005829]; extracellular exosome [GO:0070062]	55, 56
SLC2A1	Solid tumors, Prostate, Lung, Gastric, Liver,	apical plasma membrane [GO:0016324]; cytosol [GO:0005829]; extracellular exosome [GO:0070062]	57 – 62
Comum assinatura Proteômica e Proteogenômica			
CLTC	Breast, Colorectal, Liver, Prostate	clathrin coat [GO:0030118];cytosol [GO:0005829]; extracellular exosome [GO:0070062]; membrane [GO:0016020]	63 – 66
FASN	Bladder, Head and Neck, Glioma	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]; membrane [GO:0016020]; plasma membrane [GO:0005886]	67 – 69
KIF5B	Pancreas, Lung	cytosol [GO:0005829]; membrane [GO:0016020]; microtubule [GO:0005874]; vesicle [GO:0031982]	70, 71
LGALS1	Glioma, Leukemia, Bladder, Pancreas	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]	72 – 75
MYH9	Glioma, Endometrium, colorectal,	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]; membrane [GO:0016020]; nucleus [GO:0005634];	76 – 78
MYOF	Pancreas, Kidney	extracellular exosome [GO:0070062]; nuclear membrane [GO:0031965]; plasma membrane [GO:0005886]	79, 80
PCDH7	Bladder	plasma membrane [GO:0005886]	81
SERPINH1	Kidney, Lung	endoplasmic reticulum-Golgi intermediate compartment [GO:0005793]; extracellular space [GO:0005615]; membrane raft [GO:0045121]	82, 83
TGM2	Lymphoma, Colorectal, Head and Neck, Gastric, Osteosarcoma	cytosol [GO:0005829]; extracellular exosome [GO:0070062]; nucleus [GO:0005634]; plasma membrane [GO:0005886]	84 – 88
TLN1	Breast, Thyroid	cytosol [GO:0005829]; extracellular exosome [GO:0070062]; plasma membrane [GO:0005886];	89, 90
VCL	Gastric, Bladder	cytosol [GO:0005829]; extracellular exosome [GO:0070062]; plasma membrane [GO:0005886];	91, 92
Assinatura Proteômica (Caov-3 - CPTAC)			
ACTR3	Liver, Lung	cytosol [GO:0005829]; extracellular exosome [GO:0070062]; membrane [GO:0016020]; nucleus [GO:0005634]	93, 94
CANX	Breast, Lung, Prostate, Head and Neck	extracellular exosome [GO:0070062]; membrane [GO:0016020]	95 – 98
EEF2	Gastric, Prostate	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]; membrane [GO:0016020]; nucleus [GO:0005634]	99, 100
GAS6	Various, Ovarian, Bladder	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]	101 – 103
PPIB	Endometrium	cytoplasm [GO:0005737]; extracellular exosome [GO:0070062]; membrane [GO:0016020]; nucleus [GO:0005634]	104
RRBP1	Bladder, Ovarium, Endometrium, Breast, Colorectal	integral component of endoplasmic reticulum membrane [GO:0030176]; membrane [GO:0016020]	105 – 109
TAGLN2	Breast, Bladder	cytosol [GO:0005829]; extracellular exosome [GO:0070062]	110 – 112

Até o momento muitos estudos tentaram obter biomarcadores diagnósticos e prognósticos do câncer de ovário, apesar disso esta estratégia ainda não está definida e, considerando o tratamento precoce, há interesse em encontrar a combinação mais eficaz na triagem e detecção precoce dos tipos mais agressivos [113]. Ademais, a necessidade de obter testes diagnósticos abrangentes para um painel de biomarcadores precisa, além de atender os critérios, estar disponível de forma barata, rápida e confiável enquanto ferramentas [114].

As abordagens proteômicas, hoje, superam a básica identificação de proteínas, com a capacidade desenvolvida é possível, principalmente pela proteômica dirigida, executar análises de vias, síntese e degradação, modificações pós-traducionais, entre outras, de forma quantitativa e integrar esta capacidade de geração de dados com outras abordagens, como transcriptômica e metabolômica, na composição de perfis significativos na pesquisa biomédica, clínica e biotecnológica [115].

5.3 - Análise por agrupamento e classificadores das assinaturas moleculares

As assinaturas abordadas anteriormente tiveram como ponto de partida a identificação de proteínas superexpressas no secretoma de células Caov-3, em um modelo de progressão tumoral, que por sua vez não representa toda a complexidade dos tumores de ovário, com seus diversos tipos de células, microambiente e interações com o organismo, mas mantém algumas das características das células originárias. Esta redução permite a compreensão de fenômenos biológicos que contribuem para o processo de metástase, de forma mais isolada.

Os dados de amostras de 169 pacientes obtidos pelo CPTAC com a identificação de mais de 9 mil proteínas, com abordagem quantitativa, representam uma oportunidade para criação de modelos fundada na realidade dos tumores HGSC.

O primeiro modelo desenvolvido utilizou do conjunto de dados do CPTAC com informações sobre os subtipos, atribuindo as proteínas como variáveis individuais e independentes para classificação de amostras .

A primeira abordagem utilizada foi a projeção dos dados em PCA das amostras, o resultado dessa projeção é uma redução da complexidade dos níveis de todas as proteínas identificadas, para o posicionamento das amostras em um plano tridimensional. Nesta primeira abordagem já é possível visualizar o agrupamento de alguns dos subtipos moleculares em regiões de maior concentração, principalmente o subtipo mesenquimal, porém ainda sem definição. A segunda abordagem utilizada foi a criação de grupos por proximidade dos pontos utilizando o algoritmo k-means, esta proximidade é resultante da equivalência da expressão de cada amostra, resultando na criação de grupos altamente semelhantes. A análise, com adição desta camada, mostrou predominância quase unânime de amostras do subtipo mesenquimal em um dos grupos criados.

O segundo modelo desenvolvido recorreu aos dois conjuntos de dados do CPTAC, o primeiro como conjunto para treinamento, proporcionando informações sobre os subtipos, o segundo como conjunto para teste do classificador criado. Neste modelo apenas os dados referentes às amostras analisadas pelo grupo JHU foram consideradas.

Após a projeção em PCA, como descrito anteriormente, em seguida foi aplicado o classificador baseado no algoritmo Random trees. Com o classificador criado, utilizando apenas algumas das variáveis (547 proteínas), deu-se o teste deste no segundo conjunto de dados, proporcionando como resultado a probabilidade das amostras pertencerem ao

subtipo molecular mesenquimal. Os índices mais altos de probabilidade se restringiram a uma única região da projeção. Adicionalmente destacamos as amostras de estágio IV (mais alto) para verificar esta sobreposição, o que ocorreu parcialmente.

As análises bioinformáticas são essenciais na manipulação da informação gerada pelas tecnologias ômicas, estas permitem, através de métodos estatísticos e machine learning que os pesquisadores partindo de grandes conjuntos quantitativos de dados cheguem a alguns alvos relevantes para biomarcadores, desafio que exige também a comparação constante dos dados, curadoria e padronização de métodos [116]. Complementarmente, a visualização de dados é útil em vários aspectos da pesquisa, pois, se tratando de grandes conjuntos de dados, complexos para a capacidade de síntese humana, estas ferramentas auxiliam o processo, como a PCA, que salienta os aspectos mais relevantes. Mesmo no nível individual, saber onde se enquadra cada caso em um grupo é importante para a tomada de decisão [117].

5.4 – Considerações finais

A simplificação de problemas biológicos de extrema complexidade para modelos mais simples, acessíveis e reprodutíveis é uma das frentes para o estudo de doenças, como o câncer. Importante para aplicação de terapias de precisão, o conhecimento de minúcias dos processos, ainda que não integradas, proporciona informações relevantes que somadas têm o potencial desejado. Neste trabalho procuramos complementar as análises de subfrações celulares conduzidas anteriormente pelo grupo, utilizando o secretoma de células Caov-3 em processo de EMT para obtenção de assinaturas relevantes e análises funcionais.

As abordagens proteômicas se beneficiam da evolução dos equipamentos, metodologias e recursos da espectrometria de massas para a compreensão de várias

questões biológicas, em diferentes escalas. Assim como em outras abordagens ômicas, na proteômica a capacidade de aquisição de dados trouxe a esperança da medicina translacional de aplicação dos conhecimento obtido em sistemas biológicos complexos, porém há desafios, como a reprodutibilidade, complexidade das amostras, diversidade biológica, entre outros. Iniciativas compostas por várias entidades, como TCGA e CPTAC têm unido recursos na criação de metodogias robustas para tecnologias cada vez mais sensíveis a fim de contrapor os desafios e gerar dados representativos da realidade das enfermidades.

O desenvolvimento de métodos de proteômica dirigida guiado pela identificação de proteínas em abordagem *shotgun* é uma estratégia que se beneficia de medidas confiáveis de alvos selecionados em grandes coortes de amostras sem o problema intrínseco da perda de espécies associada a aquisição estocástica da proteômica *shotgun*. Atualmente a proteômica dirigida com benefício das recentes tecnologias da espectrometria de massas (mobilidade iônica, *ion trap*) e suas metodologias é capaz de quantificar painéis centenas de alvos em um único ensaio [118].

Com a disponibilidade de dados gerada pelo acesso das abordagens ômicas, a necessidade de integração dos dados é emergente, pois possibilita a montagem de uma imagem representativa do todo, esforço empreendido pela bioinformática, no desenvolvimento do racional, plataformas e *software*s para aplicação, cujo acesso livre acelera o desenvolvimento científico. Neste trabalho as abordagens bioinformáticas utilizadas representam exercícios iniciais de possibilidades na busca por alvos de interesse, podendo ser estendidas e complementadas.

6 – Conclusões

- O desenvolvimento de um método de proteômica dirigida permitiu a observação complementar de ocorrência da EMT pela quantificação relativa de marcadores clássicos deste processo.

- A abordagem quantitativa aplicada na marcação de proteínas permitiu a identificação de proteínas diferencialmente expressas no tratamento por EGF com alto grau de confiança.

- Foram identificados alvos relevantes na análise do secretoma, respaldados pela literatura e análises de suas interações.

- O método de proteômica dirigida desenvolvido para a assinatura mesenquimal possibilitou a estratificação de amostras de diferentes tipos de tumores de câncer de ovário, de forma condizente com suas características patológicas.

- A integração de dados aliada à técnicas de bioinformática gerou assinaturas passíveis de investigação em conjuntos maiores e mais abrangentes de amostras, pelos potencial de seus resultados.

7 – Figuras suplementares

Figura S1: **Avaliação molecular da indução da EMT.** a) Western blotting mostrando incremento de marcadores da EMT (N-Cadherin, β-Catenin, Vimentin and Snai1) nas células Caov-3 tratadas com EGF. b) Gel SDS-PAGE demonstrativo, mostrando perfil eletroforético dos sobrenadantes da cultura de células Caov-3 e seu padrão de corte para análise por LC-MS/MS. Adaptado de (LANFREDI et al. 2021).

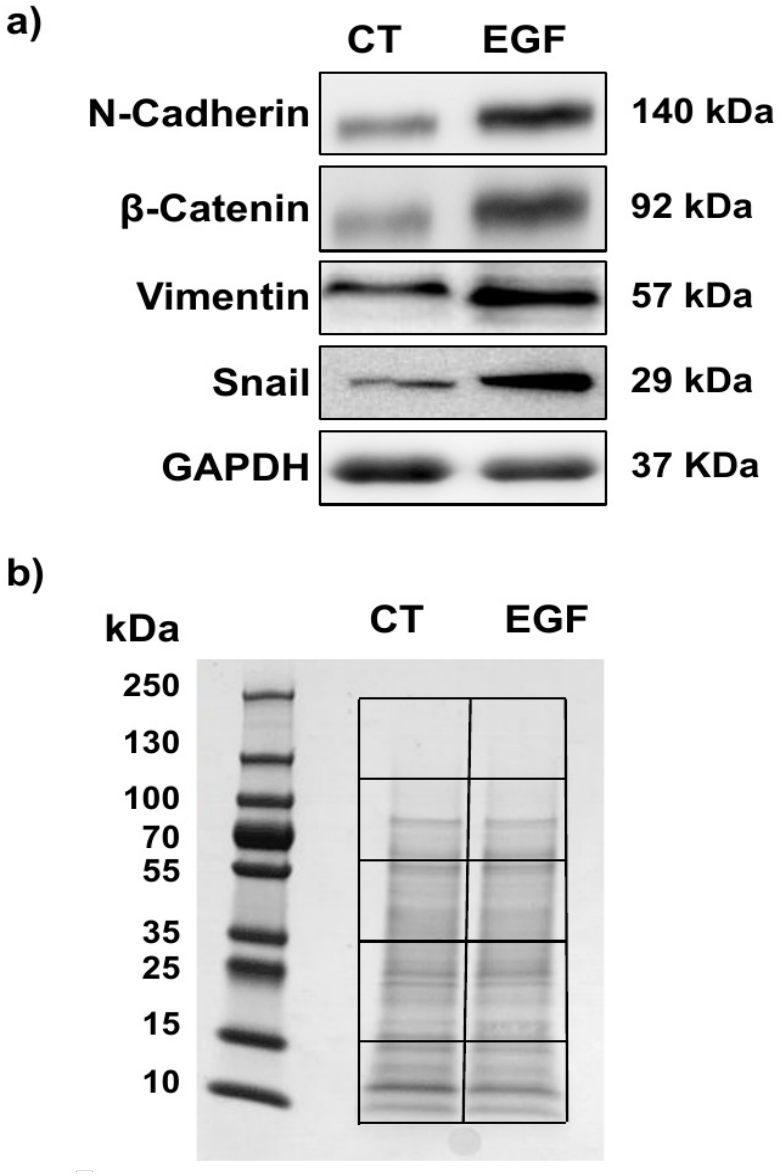
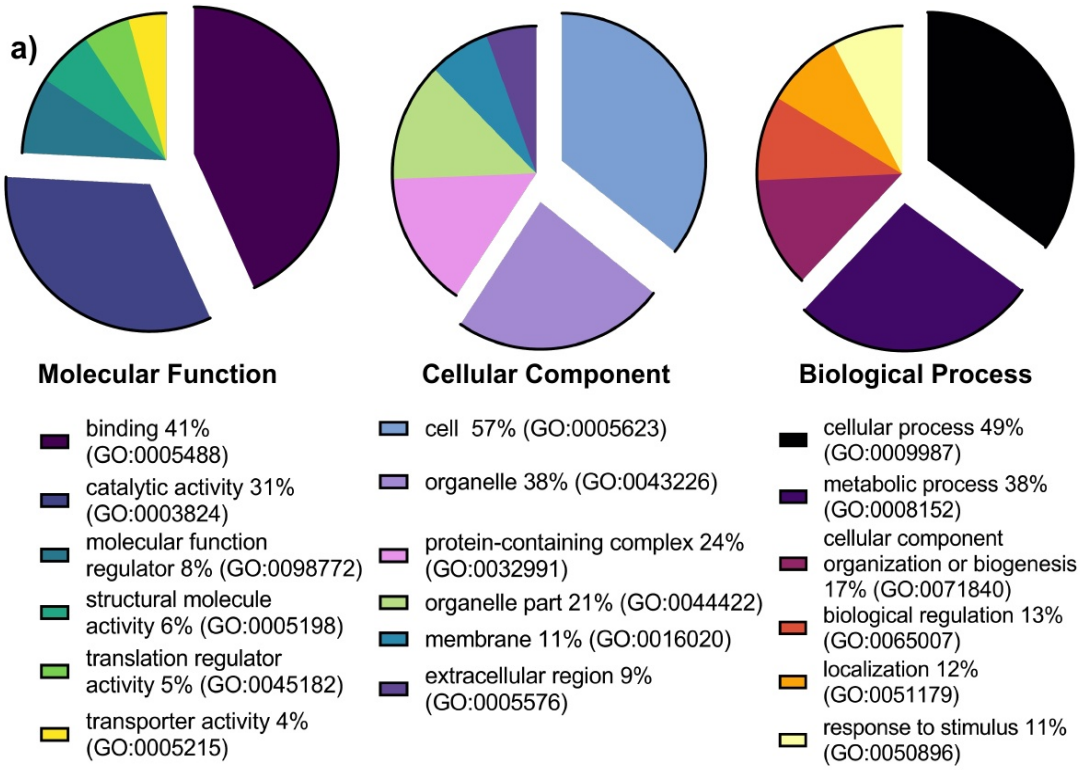
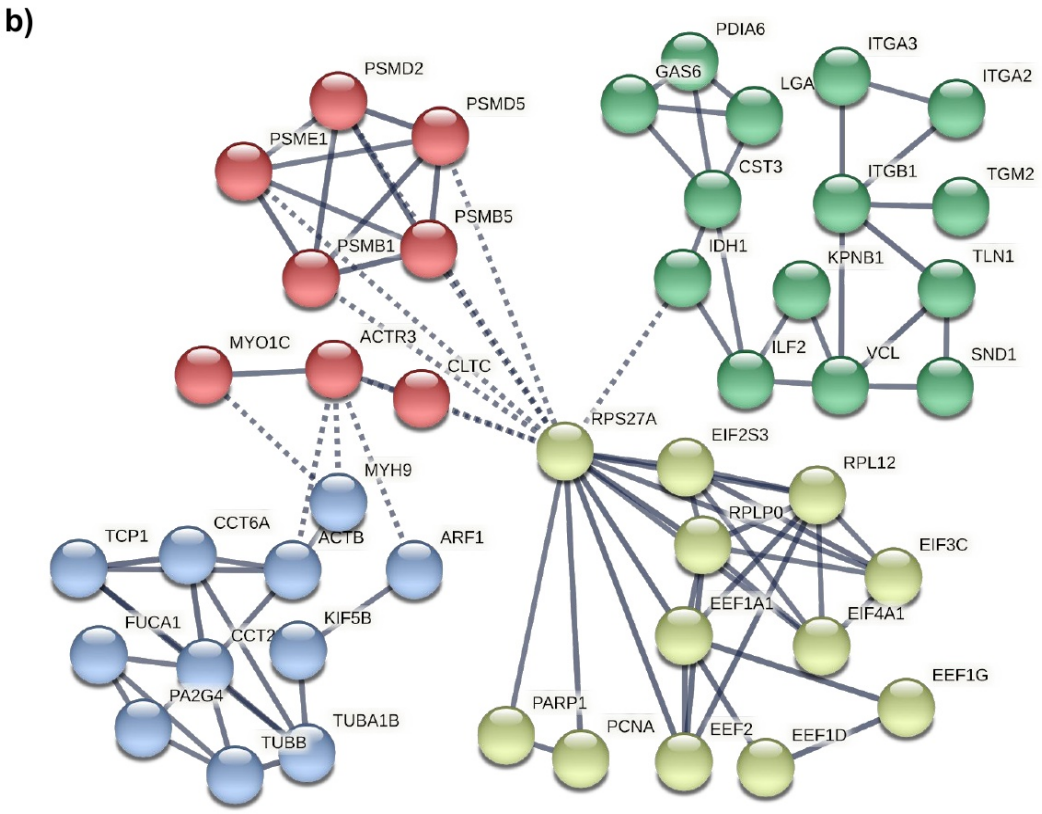


Figura S2: **Análises do conjunto de proteínas superexpressas nas células Caov-3 após indução da EMT.** a) análise de Gene Onthology pelo Panther DB, exibindo categorias majoritárias. b) Análise STRING-DB para proteínas reguladas gerou sub-redes de proteínas agrupadas por k-mean, considerando a maior confiança (0,9) para interações. As diferentes proteínas coloridas correspondem, respectivamente, às seguintes sub-redes funcionalmente relacionadas: (vermelho) deubiquitinação de proteínas, (azul) transporte mediado por vesículas, (amarelo) tradução e síntese de proteínas e (verde) localização e transporte de proteínas. (LANFREDI et al. 2021).





8 – Referências bibliográficas

(1) Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians* **2018**, *68* (6), 394–424. https://doi.org/10.3322/caac.21492.

(2) Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D. M.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods. *Int J Cancer* **2019**, *144* (8), 1941–1953. https://doi.org/10.1002/ijc.31937.

(3) INCA. *Estatísticas de câncer | INCA - Instituto Nacional de Câncer*. https://www.inca.gov.br/numeros-de-cancer (accessed 2022-05-15).

(4) Mhawech-Fauceglia, P.; Afkhami, M.; Pejovic, T. MET/HGF Signaling Pathway in Ovarian Carcinoma: Clinical Implications and Future Direction. *Patholog Res Int* **2012**, *2012*, 960327. https://doi.org/10.1155/2012/960327.

(5) Doubeni, C. A.; Doubeni, A. R.; Myers, A. E. Diagnosis and Management of Ovarian Cancer. *Am Fam Physician* **2016**, *93* (11), 937–944.

(6) Burgess, B. T.; Kolesar, J. M. Defective DNA Repair in Hereditary Ovarian Cancers: Implications for Therapy. *Am J Health Syst Pharm* **2018**, *75* (21), 1697–1707. https://doi.org/10.2146/ajhp180124.

(7) Morgan, R. J.; Armstrong, D. K.; Alvarez, R. D.; Bakkum-Gamez, J. N.; Behbakht, K.; Chen, L.-M.; Copeland, L.; Crispens, M. A.; DeRosa, M.; Dorigo, O.; Gershenson, D. M.; Gray, H. J.; Hakam, A.; Havrilesky, L. J.; Johnston, C.; Lele, S.; Martin, L.; Matulonis, U. A.; O'Malley, D. M.; Penson, R. T.; Percac-Lima, S.; Pineda, M.; Plaxe, S. C.; Powell, M. A.; Ratner, E.; Remmenga, S. W.; Rose, P. G.; Sabbatini, P.; Santoso, J. T.; Werner, T. L.; Burns, J.; Hughes, M. Ovarian Cancer, Version 1.2016, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw* **2016**, *14* (9), 1134–1163. https://doi.org/10.6004/jnccn.2016.0122.

(8) FIGO. *Ovarian Cancer Stages | Staging for Ovarian Cancer*. https://www.cancer.org/cancer/ovarian-cancer/detection-diagnosis-staging/staging.html (accessed 2022-05-15).

(9) Auersperg, N. Ovarian Surface Epithelium as a Source of Ovarian Cancers: Unwarranted Speculation or Evidence-Based Hypothesis? *Gynecol Oncol* **2013**, *130* (1), 246–251. https://doi.org/10.1016/j.ygyno.2013.03.021.

(10) Klotz, D. M.; Wimberger, P. Cells of Origin of Ovarian Cancer: Ovarian Surface Epithelium or Fallopian Tube? *Arch Gynecol Obstet* **2017**, *296* (6), 1055–1062. https://doi.org/10.1007/s00404-017-4529-z.

(11) Hollis, R. L.; Gourley, C. Genetic and Molecular Changes in Ovarian Cancer. *Cancer Biol Med* **2016**, *13* (2), 236–247. https://doi.org/10.20892/j.issn.2095-3941.2016.0024.

(12) Kotsopoulos, J.; Gronwald, J.; Karlan, B.; Rosen, B.; Huzarski, T.; Moller, P.; Lynch, H. T.; Singer, C. F.; Senter, L.; Neuhausen, S. L.; Tung, N.; Eisen, A.; Foulkes, W. D.; Ainsworth, P.; Sun, P.; Lubinski, J.; Narod, S. A.; Hereditary Ovarian Cancer Clinical Study Group. Age-Specific Ovarian Cancer Risks among Women with a BRCA1 or BRCA2 Mutation. *Gynecol Oncol* **2018**, *150* (1), 85–91. https://doi.org/10.1016/j.ygyno.2018.05.011.

(13) Shih, I.-M.; Wang, Y.; Wang, T.-L. The Origin of Ovarian Cancer Species and Precancerous Landscape. *Am J Pathol* **2021**, *191* (1), 26–39. https://doi.org/10.1016/j.ajpath.2020.09.006.
(14) Reinartz, S.; Finkernagel, F.; Adhikary, T.; Rohnalter, V.; Schumann, T.; Schober, Y.; Nockher, W. A.; Nist, A.; Stiewe, T.; Jansen, J. M.; Wagner, U.; Müller-Brüsselbach, S.; Müller, R. A Transcriptome-Based Global Map of Signaling Pathways in the Ovarian Cancer Microenvironment Associated with Clinical Outcome. *Genome Biol* **2016**, *17* (1), 108. https://doi.org/10.1186/s13059-016-0956-6.
(15) Granados, M. L.; Hudson, L. G.; Samudio-Ruiz, S. L. Contributions of the Epidermal Growth Factor Receptor to Acquisition of Platinum Resistance in Ovarian Cancer Cells. *PLoS One* **2015**, *10* (9), e0136893. https://doi.org/10.1371/journal.pone.0136893.
(16) Grassi, M. L.; Palma, C. de S.; Thomé, C. H.; Lanfredi, G. P.; Poersch, A.; Faça, V. M. Proteomic Analysis of Ovarian Cancer Cells during Epithelial-Mesenchymal Transition (EMT) Induced by Epidermal Growth Factor (EGF) Reveals Mechanisms of Cell Cycle Control. *J Proteomics* **2017**, *151*, 2–11. https://doi.org/10.1016/j.jprot.2016.06.009.
(17) Pastushenko, I.; Blanpain, C. EMT Transition States during Tumor Progression and Metastasis. *Trends Cell Biol* **2019**, *29* (3), 212–226. https://doi.org/10.1016/j.tcb.2018.12.001.
(18) Peters, P. N.; Schryver, E. M.; Lengyel, E.; Kenny, H. Modeling the Early Steps of Ovarian Cancer Dissemination in an Organotypic Culture of the Human Peritoneal Cavity. *J Vis Exp* **2015**, No. 106, e53541. https://doi.org/10.3791/53541.
(19) Derynck, R.; Weinberg, R. A. EMT and Cancer: More Than Meets the Eye. *Dev Cell* **2019**, *49* (3), 313–316. https://doi.org/10.1016/j.devcel.2019.04.026.
(20) Lambert, A. W.; Pattabiraman, D. R.; Weinberg, R. A. Emerging Biological Principles of Metastasis. *Cell* **2017**, *168* (4), 670–691. https://doi.org/10.1016/j.cell.2016.11.037.
(21) Stemmler, M. P.; Eccles, R. L.; Brabletz, S.; Brabletz, T. Non-Redundant Functions of EMT Transcription Factors. *Nat Cell Biol* **2019**, *21* (1), 102–112. https://doi.org/10.1038/s41556-018-0196-y.
(22) Nieto, M. A.; Huang, R. Y.-J.; Jackson, R. A.; Thiery, J. P. EMT: 2016. *Cell* **2016**, *166* (1), 21–45. https://doi.org/10.1016/j.cell.2016.06.028.
(23) Ramesh, V.; Brabletz, T.; Ceppi, P. Targeting EMT in Cancer with Repurposed Metabolic Inhibitors. *Trends Cancer* **2020**, *6* (11), 942–950. https://doi.org/10.1016/j.trecan.2020.06.005.
(24) Sanjabi, S.; Oh, S. A.; Li, M. O. Regulation of the Immune Response by TGF-β: From Conception to Autoimmunity and Infection. *Cold Spring Harb Perspect Biol* **2017**, *9* (6), a022236. https://doi.org/10.1101/cshperspect.a022236.
(25) Charkhchi, P.; Cybulski, C.; Gronwald, J.; Wong, F. O.; Narod, S. A.; Akbari, M. R. CA125 and Ovarian Cancer: A Comprehensive Review. *Cancers (Basel)* **2020**, *12* (12), E3730. https://doi.org/10.3390/cancers12123730.
(26) Bradbury, M.; Borràs, E.; Pérez-Benavente, A.; Gil-Moreno, A.; Santamaria, A.; Sabidó, E. Proteomic Studies on the Management of High-Grade Serous Ovarian Cancer Patients: A Mini-Review. *Cancers (Basel)* **2021**, *13* (9), 2067. https://doi.org/10.3390/cancers13092067.
(27) Carvalho, V. P. de; Grassi, M. L.; Palma, C. de S.; Carrara, H. H. A.; Faça, V. M.; Reis, F. J. C. dos; Poersch, A. The Contribution and Perspectives of Proteomics to

Uncover Ovarian Cancer Tumor Markers. *Translational Research* **2019**, *206*, 71–90. https://doi.org/10.1016/j.trsl.2018.11.001.
(28) Ueland, F. R. A Perspective on Ovarian Cancer Biomarkers: Past, Present and Yet-To-Come. *Diagnostics (Basel)* **2017**, *7* (1), E14. https://doi.org/10.3390/diagnostics7010014.
(29) Swiatly, A.; Plewa, S.; Matysiak, J.; Kokot, Z. J. Mass Spectrometry-Based Proteomics Techniques and Their Application in Ovarian Cancer Research. *J Ovarian Res* **2018**, *11*, 88. https://doi.org/10.1186/s13048-018-0460-6.
(30) Bell, D.; Berchuck, A.; Birrer, M.; Chien, J.; Cramer, D. W.; Dao, F.; Dhir, R.; DiSaia, P.; Gabra, H.; Glenn, P.; Godwin, A. K.; Gross, J.; Hartmann, L.; Huang, M.; Huntsman, D. G.; Iacocca, M.; Imielinski, M.; Kalloger, S.; Karlan, B. Y.; Levine, D. A.; Mills, G. B.; Morrison, C.; Mutch, D.; Olvera, N.; Orsulic, S.; Park, K.; Petrelli, N.; Rabeno, B.; Rader, J. S.; Sikic, B. I.; Smith-McCune, K.; Sood, A. K.; Bowtell, D.; Penny, R.; Testa, J. R.; Chang, K.; Creighton, C. J.; Dinh, H. H.; Drummond, J. A.; Fowler, G.; Gunaratne, P.; Hawes, A. C.; Kovar, C. L.; Lewis, L. R.; Morgan, M. B.; Newsham, I. F.; Santibanez, J.; Reid, J. G.; Trevino, L. R.; Wu, Y.-Q.; Wang, M.; Muzny, D. M.; Wheeler, D. A.; Gibbs, R. A.; Getz, G.; Lawrence, M. S.; Cibulskis, K.; Sivachenko, A. Y.; Sougnez, C.; Voet, D.; Wilkinson, J.; Bloom, T.; Ardlie, K.; Fennell, T.; Baldwin, J.; Nichol, R.; Fisher, S.; Gabriel, S.; Lander, E. S.; Ding, L.; Fulton, R. S.; Koboldt, D. C.; McLellan, M. D.; Wylie, T.; Walker, J.; O'Laughlin, M.; Dooling, D. J.; Fulton, L.; Abbott, R.; Dees, N. D.; Zhang, Q.; Kandoth, C.; Wendl, M.; Schierding, W.; Shen, D.; Harris, C. C.; Schmidt, H.; Kalicki, J.; Delehaunty, K. D.; Fronick, C. C.; Demeter, R.; Cook, L.; Wallis, J. W.; Lin, L.; Magrini, V. J.; Hodges, J. S.; Eldred, J. M.; Smith, S. M.; Pohl, C. S.; Vandin, F.; Upfal, E.; Raphael, B. J.; Weinstock, G. M.; Mardis, E. R.; Wilson, R. K.; Meyerson, M.; Winckler, W.; Getz, G.; Verhaak, R. G. W.; Carter, S. L.; Mermel, C. H.; Saksena, G.; Nguyen, H.; Onofrio, R. C.; Lawrence, M. S.; Hubbard, D.; Gupta, S.; Crenshaw, A.; Ramos, A. H.; Ardlie, K.; Chin, L.; Protopopov, A.; Zhang, J.; Kim, T. M.; Perna, I.; Xiao, Y.; Zhang, H.; Ren, G.; Sathiamoorthy, N.; Park, R. W.; Lee, E.; Park, P. J.; Kucherlapati, R.; Absher, D. M.; Waite, L.; Sherlock, G.; Brooks, J. D.; Li, J. Z.; Xu, J.; Myers, R. M.; Laird, P. W.; Cope, L.; Herman, J. G.; Shen, H.; Weisenberger, D. J.; Noushmehr, H.; Pan, F.; Triche Jr, T.; Berman, B. P.; Van Den Berg, D. J.; Buckley, J.; Baylin, S. B.; Spellman, P. T.; Purdom, E.; Neuvial, P.; Bengtsson, H.; Jakkula, L. R.; Durinck, S.; Han, J.; Dorton, S.; Marr, H.; Choi, Y. G.; Wang, V.; Wang, N. J.; Ngai, J.; Conboy, J. G.; Parvin, B.; Feiler, H. S.; Speed, T. P.; Gray, J. W.; Levine, D. A.; Socci, N. D.; Liang, Y.; Taylor, B. S.; Schultz, N.; Borsu, L.; Lash, A. E.; Brennan, C.; Viale, A.; Sander, C.; Ladanyi, M.; Hoadley, K. A.; Meng, S.; Du, Y.; Shi, Y.; Li, L.; Turman, Y. J.; Zang, D.; Helms, E. B.; Balu, S.; Zhou, X.; Wu, J.; Topal, M. D.; Hayes, D. N.; Perou, C. M.; Getz, G.; Voet, D.; Saksena, G.; Zhang, J.; Zhang, H.; Wu, C. J.; Shukla, S.; Cibulskis, K.; Lawrence, M. S.; Sivachenko, A.; Jing, R.; Park, R. W.; Liu, Y.; Park, P. J.; Noble, M.; Chin, L.; Carter, H.; Kim, D.; Samayoa, J.; Karchin, R.; Spellman, P. T.; Purdom, E.; Neuvial, P.; Bengtsson, H.; Durinck, S.; Han, J.; Korkola, J. E.; Heiser, L. M.; Cho, R. J.; Hu, Z.; Parvin, B.; Speed, T. P.; Gray, J. W.; Schultz, N.; Cerami, E.; Taylor, B. S.; Olshen, A.; Reva, B.; Antipin, Y.; Shen, R.; Mankoo, P.; Sheridan, R.; Ciriello, G.; Chang, W. K.; Bernanke, J. A.; Borsu, L.; Levine, D. A.; Ladanyi, M.; Sander, C.; Haussler, D.; Benz, C. C.; Stuart, J. M.; Benz, S. C.; Sanborn, J. Z.; Vaske, C. J.; Zhu, J.; Szeto, C.; Scott, G. K.; Yau, C.; Hoadley, K. A.; Du, Y.; Balu, S.; Hayes, D. N.; Perou, C. M.; Wilkerson, M. D.; Zhang, N.; Akbani, R.; Baggerly, K. A.; Yung, W. K.; Mills, G. B.; Weinstein, J. N.; Penny, R.; Shelton, T.; Grimm, D.; Hatfield, M.; Morris, S.;

Yena, P.; Rhodes, P.; Sherman, M.; Paulauskis, J.; Millis, S.; Kahn, A.; Greene, J. M.; The Cancer Genome Atlas Research Network; (Participants are arranged by area of contribution and then by institution.); Disease working group and tissue source sites; Genome sequencing centres: Baylor College of Medicine; Broad Institute; Washington University in St Louis; Cancer genome characterization centres: Broad Institute/Dana-Farber Cancer Institute; Harvard Medical School; HudsonAlpha Institute/Stanford University; University of Southern California/Johns Hopkins University; Lawrence Berkeley National Laboratory; Memorial Sloan-Kettering Cancer Center; University of North Carolina at Chapel Hill; Genome data analysis centres: Broad Institute; Johns Hopkins University; University of California Santa Cruz/Buck Institute; The University of Texas MD Anderson Cancer Center; Biospecimen core resource; Data coordination centre. Integrated Genomic Analyses of Ovarian Carcinoma. *Nature* **2011**, *474* (7353), 609–615. https://doi.org/10.1038/nature10166.

(31) Zhang, H.; Liu, T.; Zhang, Z.; Payne, S. H.; Zhang, B.; McDermott, J. E.; Zhou, J.-Y.; Petyuk, V. A.; Chen, L.; Ray, D.; Sun, S.; Yang, F.; Chen, L.; Wang, J.; Shah, P.; Cha, S. W.; Aiyetan, P.; Woo, S.; Tian, Y.; Gritsenko, M. A.; Clauss, T. R.; Choi, C.; Monroe, M. E.; Thomas, S.; Nie, S.; Wu, C.; Moore, R. J.; Yu, K.-H.; Tabb, D. L.; Fenyö, D.; Bafna, V.; Wang, Y.; Rodriguez, H.; Boja, E. S.; Hiltke, T.; Rivers, R. C.; Sokoll, L.; Zhu, H.; Shih, I.-M.; Cope, L.; Pandey, A.; Zhang, B.; Snyder, M. P.; Levine, D. A.; Smith, R. D.; Chan, D. W.; Rodland, K. D.; CPTAC Investigators. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell* **2016**, *166* (3), 755–765. https://doi.org/10.1016/j.cell.2016.05.069.

(32) Mukherjee, S.; Sundfeldt, K.; Borrebaeck, C. A. K.; Jakobsson, M. E. Comprehending the Proteomic Landscape of Ovarian Cancer: A Road to the Discovery of Disease Biomarkers. *Proteomes* **2021**, *9* (2), 25. https://doi.org/10.3390/proteomes9020025.

(33) Khella, C. A.; Mehta, G. A.; Mehta, R. N.; Gatza, M. L. Recent Advances in Integrative Multi-Omics Research in Breast and Ovarian Cancer. *J Pers Med* **2021**, *11* (2), 149. https://doi.org/10.3390/jpm11020149.

(34) Ryu, J.; Thomas, S. N. Quantitative Mass Spectrometry-Based Proteomics for Biomarker Development in Ovarian Cancer. *Molecules* **2021**, *26* (9), 2674. https://doi.org/10.3390/molecules26092674.

(35) Nice, E. C. The Separation Sciences, the Front End to Proteomics: An Historical Perspective. *Biomed Chromatogr* **2021**, *35* (1), e4995. https://doi.org/10.1002/bmc.4995.

(36) Ponomarenko, E. A.; Poverennaya, E. V.; Ilgisonis, E. V.; Pyatnitskiy, M. A.; Kopylov, A. T.; Zgoda, V. G.; Lisitsa, A. V.; Archakov, A. I. The Size of the Human Proteome: The Width and Depth. *Int J Anal Chem* **2016**, *2016*, 7436849. https://doi.org/10.1155/2016/7436849.

(37) Aebersold, R.; Mann, M. Mass-Spectrometric Exploration of Proteome Structure and Function. *Nature* **2016**, *537* (7620), 347–355. https://doi.org/10.1038/nature19949.

(38) Su, M.; Zhang, Z.; Zhou, L.; Han, C.; Huang, C.; Nice, E. C. Proteomics, Personalized Medicine and Cancer. *Cancers (Basel)* **2021**, *13* (11), 2512. https://doi.org/10.3390/cancers13112512.

(39) Miles, H. N.; Delafield, D. G.; Li, L. Recent Developments and Applications of Quantitative Proteomics Strategies for High-Throughput Biomolecular Analyses in Cancer Research. *RSC Chem. Biol.* **2021**, *2* (4), 1050–1072. https://doi.org/10.1039/D1CB00039J.

(40) Key, M. A Tutorial in Displaying Mass Spectrometry-Based Proteomic Data Using Heat Maps. *BMC Bioinformatics* **2012**, *13 Suppl 16*, S10. https://doi.org/10.1186/1471-2105-13-S16-S10.
(41) Lanfredi, G. P.; Thomé, C. H.; Ferreira, G. A.; Silvestrini, V. C.; Masson, A. P.; Vargas, A. P.; Grassi, M. L.; Poersch, A.; Candido Dos Reis, F. J.; Faça, V. M. Analysis of Ovarian Cancer Cell Secretome during Epithelial to Mesenchymal Transition Reveals a Protein Signature Associated with Advanced Stages of Ovarian Tumors. *Biochim Biophys Acta Proteins Proteom* **2021**, *1869* (6), 140623. https://doi.org/10.1016/j.bbapap.2021.140623.
(42) McDermott, J. E.; Arshad, O. A.; Petyuk, V. A.; Fu, Y.; Gritsenko, M. A.; Clauss, T. R.; Moore, R. J.; Schepmoes, A. A.; Zhao, R.; Monroe, M. E.; Schnaubelt, M.; Tsai, C.-F.; Payne, S. H.; Huang, C.; Wang, L.-B.; Foltz, S.; Wyczalkowski, M.; Wu, Y.; Song, E.; Brewer, M. A.; Thiagarajan, M.; Kinsinger, C. R.; Robles, A. I.; Boja, E. S.; Rodriguez, H.; Chan, D. W.; Zhang, B.; Zhang, Z.; Ding, L.; Smith, R. D.; Liu, T.; Rodland, K. D. Proteogenomic Characterization of Ovarian HGSC Implicates Mitotic Kinases, Replication Stress in Observed Chromosomal Instability. *Cell Reports Medicine* **2020**, *1* (1), 100004. https://doi.org/10.1016/j.xcrm.2020.100004.
(43) Papp, E.; Hallberg, D.; Konecny, G. E.; Bruhm, D. C.; Adleff, V.; Noë, M.; Kagiampakis, I.; Palsgrove, D.; Conklin, D.; Kinose, Y.; White, J. R.; Press, M. F.; Drapkin, R.; Easwaran, H.; Baylin, S. B.; Slamon, D.; Velculescu, V. E.; Scharpf, R. B. Integrated Genomic, Epigenomic, and Expression Analyses of Ovarian Cancer Cell Lines. *Cell Rep* **2018**, *25* (9), 2617–2633. https://doi.org/10.1016/j.celrep.2018.10.096.
(44) Marquèze-Pouey, B.; Mailfert, S.; Rouger, V.; Goaillard, J.-M.; Marguet, D. Physiological Epidermal Growth Factor Concentrations Activate High Affinity Receptors to Elicit Calcium Oscillations. *PLoS One* **2014**, *9* (9), e106803. https://doi.org/10.1371/journal.pone.0106803.
(45) Shao, G.; Wang, J.; Li, Y.; Liu, X.; Xie, X.; Wan, X.; Yan, M.; Jin, J.; Lin, Q.; Zhu, H.; Zhang, L.; Gong, A.; Shao, Q.; Wu, C. Lysine-Specific Demethylase 1 Mediates Epidermal Growth Factor Signaling to Promote Cell Migration in Ovarian Cancer Cells. *Sci Rep* **2015**, *5*, 15344. https://doi.org/10.1038/srep15344.
(46) Stastna, M.; Van Eyk, J. E. Secreted Proteins as a Fundamental Source for Biomarker Discovery. *PROTEOMICS* **2012**, *12* (4–5), 722–735. https://doi.org/10.1002/pmic.201100346.
(47) Brown, K. J.; Seol, H.; Pillai, D. K.; Sankoorikal, B.-J.; Formolo, C. A.; Mac, J.; Edwards, N. J.; Rose, M. C.; Hathout, Y. The Human Secretome Atlas Initiative: Implications in Health and Disease Conditions. *Biochim Biophys Acta* **2013**, *1834* (11), 2454–2461. https://doi.org/10.1016/j.bbapap.2013.04.007.
(48) Schaaij-Visser, T. B. M.; de Wit, M.; Lam, S. W.; Jiménez, C. R. The Cancer Secretome, Current Status and Opportunities in the Lung, Breast and Colorectal Cancer Context. *Biochim Biophys Acta* **2013**, *1834* (11), 2242–2258. https://doi.org/10.1016/j.bbapap.2013.01.029.
(49) Pearce, O. M. T.; Delaine-Smith, R. M.; Maniati, E.; Nichols, S.; Wang, J.; Böhm, S.; Rajeeve, V.; Ullah, D.; Chakravarty, P.; Jones, R. R.; Montfort, A.; Dowe, T.; Gribben, J.; Jones, J. L.; Kocher, H. M.; Serody, J. S.; Vincent, B. G.; Connelly, J.; Brenton, J. D.; Chelala, C.; Cutillas, P. R.; Lockley, M.; Bessant, C.; Knight, M. M.; Balkwill, F. R. Deconstruction of a Metastatic Tumor Microenvironment Reveals a Common Matrix Response in Human Cancers. *Cancer Discovery* **2018**, *8* (3), 304–319. https://doi.org/10.1158/2159-8290.CD-17-0284.

(50) Huang, F.; Kirkpatrick, D.; Jiang, X.; Gygi, S.; Sorkin, A. Differential Regulation of EGF Receptor Internalization and Degradation by Multiubiquitination within the Kinase Domain. *Mol Cell* **2006**, *21* (6), 737–748. https://doi.org/10.1016/j.molcel.2006.02.018.
(51) Vivanco, I.; Rohle, D.; Versele, M.; Iwanami, A.; Kuga, D.; Oldrini, B.; Tanaka, K.; Dang, J.; Kubek, S.; Palaskas, N.; Hsueh, T.; Evans, M.; Mulholland, D.; Wolle, D.; Rajasekaran, S.; Rajasekaran, A.; Liau, L. M.; Cloughesy, T. F.; Dikic, I.; Brennan, C.; Wu, H.; Mischel, P. S.; Perera, T.; Mellinghoff, I. K. The Phosphatase and Tensin Homolog Regulates Epidermal Growth Factor Receptor (EGFR) Inhibitor Response by Targeting EGFR for Degradation. *Proc Natl Acad Sci U S A* **2010**, *107* (14), 6459–6464. https://doi.org/10.1073/pnas.0911188107.
(52) Konecny, G. E.; Wang, C.; Hamidi, H.; Winterhoff, B.; Kalli, K. R.; Dering, J.; Ginther, C.; Chen, H.-W.; Dowdy, S.; Cliby, W.; Gostout, B.; Podratz, K. C.; Keeney, G.; Wang, H.-J.; Hartmann, L. C.; Slamon, D. J.; Goode, E. L. Prognostic and Therapeutic Relevance of Molecular Subtypes in High-Grade Serous Ovarian Cancer. *J Natl Cancer Inst* **2014**, *106* (10), dju249. https://doi.org/10.1093/jnci/dju249.
(53) Ping, Z.; Siegal, G. P.; Harada, S.; Eltoum, I.-E.; Youssef, M.; Shen, T.; He, J.; Huang, Y.; Chen, D.; Li, Y.; Bland, K. I.; Chang, H. R.; Shen, D. ERBB2 Mutation Is Associated with a Worse Prognosis in Patients with CDH1 Altered Invasive Lobular Cancer of the Breast. *Oncotarget* **2016**, *7* (49), 80655–80663. https://doi.org/10.18632/oncotarget.13019.
(54) Wu, Z.-H.; Zhong, Y.; Zhou, T.; Xiao, H.-J. MiRNA Biomarkers for Predicting Overall Survival Outcomes for Head and Neck Squamous Cell Carcinoma. *Genomics* **2021**, *113* (1, Part 1), 135–141. https://doi.org/10.1016/j.ygeno.2020.12.002.
(55) Bhanot, H.; Reddy, M. M.; Nonami, A.; Weisberg, E. L.; Bonal, D.; Kirschmeier, P. T.; Salgia, S.; Podar, K.; Galinsky, I.; Chowdary, T. K.; Neuberg, D.; Tonon, G.; Stone, R. M.; Asara, J.; Griffin, J. D.; Sattler, M. Pathological Glycogenesis through Glycogen Synthase 1 and Suppression of Excessive AMP Kinase Activity in Myeloid Leukemia Cells. *Leukemia* **2015**, *29* (7), 1555–1563. https://doi.org/10.1038/leu.2015.46.
(56) Li, L.; Yang, L.; Cheng, S.; Fan, Z.; Shen, Z.; Xue, W.; Zheng, Y.; Li, F.; Wang, D.; Zhang, K.; Lian, J.; Wang, D.; Zhu, Z.; Zhao, J.; Zhang, Y. Lung Adenocarcinoma-Intrinsic GBE1 Signaling Inhibits Anti-Tumor Immunity. *Mol Cancer* **2019**, *18* (1), 108. https://doi.org/10.1186/s12943-019-1027-x.
(57) Wang, J.; Ye, C.; Chen, C.; Xiong, H.; Xie, B.; Zhou, J.; Chen, Y.; Zheng, S.; Wang, L. Glucose Transporter GLUT1 Expression and Clinical Outcome in Solid Tumors: A Systematic Review and Meta-Analysis. *Oncotarget* **2017**, *8* (10), 16875–16886. https://doi.org/10.18632/oncotarget.15171.
(58) Zhang, L.; Wang, B.; Wang, Z.-S.; Guo, Y.-L.; Shen, H. Construction of Glycolytic Regulator Gene Signature to Predict the Prognosis and Tumor Immune Cell Infiltration Levels for Prostate Cancer. *Comput Math Methods Med* **2022**, *2022*, 9273559. https://doi.org/10.1155/2022/9273559.
(59) Tang, B.; Xu, W.; Wang, Y.; Zhu, J.; Wang, H.; Tu, J.; Weng, Q.; Kong, C.; Yang, Y.; Qiu, R.; Zhao, Z.; Xu, M.; Ji, J. Identification of Critical Ferroptosis Regulators in Lung Adenocarcinoma That RRM2 Facilitates Tumor Immune Infiltration by Inhibiting Ferroptotic Death. *Clinical Immunology* **2021**, *232*, 108872. https://doi.org/10.1016/j.clim.2021.108872.
(60) Min, K.-W.; Kim, D.-H.; Son, B. K.; Moon, K. M.; Kim, S. M.; Intazur Rahaman, Md.; Kim, S. W.; Kim, E.-K.; Kwon, M. J.; Koh, Y. W.; Oh, I. H. High SLC2A1 Expression

Associated with Suppressing CD8 T Cells and B Cells Promoted Cancer Survival in Gastric Cancer. *PLoS One* **2021**, *16* (3), e0245075. https://doi.org/10.1371/journal.pone.0245075.
(61) Ding, X.; Liu, J.; Liu, T.; Ma, Z.; Wen, D.; Zhu, J. MiR‐148b Inhibits Glycolysis in Gastric Cancer through Targeting SLC2A1. *Cancer Med* **2017**, *6* (6), 1301–1310. https://doi.org/10.1002/cam4.1008.
(62) Shang, R.; Wang, M.; Dai, B.; Du, J.; Wang, J.; Liu, Z.; Qu, S.; Yang, X.; Liu, J.; Xia, C.; Wang, L.; Wang, D.; Li, Y. Long Noncoding RNA SLC2A1‐AS1 Regulates Aerobic Glycolysis and Progression in Hepatocellular Carcinoma via Inhibiting the STAT3/FOXM1/GLUT1 Pathway. *Mol Oncol* **2020**, *14* (6), 1381–1396. https://doi.org/10.1002/1878-0261.12666.
(63) Ujihira, T.; Ikeda, K.; Suzuki, T.; Yamaga, R.; Sato, W.; Horie-Inoue, K.; Shigekawa, T.; Osaki, A.; Saeki, T.; Okamoto, K.; Takeda, S.; Inoue, S. MicroRNA-574-3p, Identified by MicroRNA Library-Based Functional Screening, Modulates Tamoxifen Response in Breast Cancer. *Sci Rep* **2015**, *5*, 7641. https://doi.org/10.1038/srep07641.
(64) Imbastari, F.; Dahlmann, M.; Sporbert, A.; Mattioli, C. C.; Mari, T.; Scholz, F.; Timm, L.; Twamley, S.; Migotti, R.; Walther, W.; Dittmar, G.; Rehm, A.; Stein, U. MACC1 Regulates Clathrin-Mediated Endocytosis and Receptor Recycling of Transferrin Receptor and EGFR in Colorectal Cancer. *Cell Mol Life Sci* **2021**, *78* (7), 3525–3542. https://doi.org/10.1007/s00018-020-03734-1.
(65) Caballero-Díaz, D.; Bertran, E.; Peñuelas-Haro, I.; Moreno-Càceres, J.; Malfettone, A.; López-Luque, J.; Addante, A.; Herrera, B.; Sánchez, A.; Alay, A.; Solé, X.; Serrano, T.; Ramos, E.; Fabregat, I. Clathrin Switches Transforming Growth Factor-β Role to pro-Tumorigenic in Liver Cancer. *Journal of Hepatology* **2020**, *72* (1), 125–134. https://doi.org/10.1016/j.jhep.2019.09.012.
(66) Liu, C.-M.; Yu, C.-C.; Lin, T.; Liao, Y.-W.; Hsieh, P.-L.; Yu, C.-H.; Lee, S.-S. E3 Ligase STUB1 Attenuates Stemness and Tumorigenicity of Oral Carcinoma Cells via Transglutaminase 2 Regulation. *Journal of the Formosan Medical Association* **2020**, *119* (10), 1532–1538. https://doi.org/10.1016/j.jfma.2020.06.004.
(67) Xiong, Q.; Feng, D.; Wang, Z.; Ying, Y.; Xu, C.; Wei, Q.; Zeng, S.; Yang, L. Fatty Acid Synthase Is the Key Regulator of Fatty Acid Metabolism and Is Related to Immunotherapy in Bladder Cancer. *Front Immunol* **2022**, *13*, 836939. https://doi.org/10.3389/fimmu.2022.836939.
(68) de Angelis, C. M.; de Lima-Souza, R. A.; Scarini, J. F.; Egal, E. S. A.; do Amaral-Silva, G. K.; de Oliveira Gondak, R.; de Almeida, O. P.; Chone, C. T.; Kowalski, L. P.; Altemani, A.; Mariano, F. V. Immunohistochemical Expression of Fatty Acid Synthase (FASN) Is Correlated to Tumor Aggressiveness and Cellular Differentiation in Salivary Gland Carcinomas. *Head Neck Pathol* **2021**, *15* (4), 1119–1126. https://doi.org/10.1007/s12105-021-01319-3.
(69) Ricklefs, F. L.; Maire, C. L.; Matschke, J.; Dührsen, L.; Sauvigny, T.; Holz, M.; Kolbe, K.; Peine, S.; Herold-Mende, C.; Carter, B.; Chiocca, E. A.; Lawler, S. E.; Westphal, M.; Lamszus, K. FASN Is a Biomarker Enriched in Malignant Glioma-Derived Extracellular Vesicles. *Int J Mol Sci* **2020**, *21* (6), 1931. https://doi.org/10.3390/ijms21061931.
(70) Zhao, L.; Liu, H.; Luo, S.; Moorman, P. G.; Walsh, K. M.; Li, W.; Wei, Q. Associations between Genetic Variants of KIF5B, FMN1, and MGAT3 in the Cadherin Pathway and Pancreatic Cancer Risk. *Cancer Med* **2020**, *9* (24), 9620–9631. https://doi.org/10.1002/cam4.3603.

(71) Das, T. K.; Cagan, R. L. KIF5B-RET Oncoprotein Signals Through A Multi-Kinase Signaling Hub. *Cell Rep* **2017**, *20* (10), 2368–2383. https://doi.org/10.1016/j.celrep.2017.08.037.
(72) You, X.; Wu, J.; Wang, Y.; Liu, Q.; Cheng, Z.; Zhao, X.; Liu, G.; Huang, C.; Dai, J.; Zhou, Y.; Chen, D.; Chong, Y. Galectin-1 Promotes Vasculogenic Mimicry in Gastric Adenocarcinoma via the Hedgehog/GLI Signaling Pathway. *Aging (Albany NY)* **2020**, *12* (21), 21837–21853. https://doi.org/10.18632/aging.104000.
(73) Zhang, H.; Liu, B.; Cheng, J.; Ma, H.; Li, Z.; Xi, Y. Identification of Co-Expressed Genes Associated with MLL Rearrangement in Pediatric Acute Lymphoblastic Leukemia. *Biosci Rep* **2020**, *40* (5), BSR20200514. https://doi.org/10.1042/BSR20200514.
(74) Wu, T.-F.; Li, C.-F.; Chien, L.-H.; Shen, K.-H.; Huang, H.-Y.; Su, C.-C.; Liao, A. C. Galectin-1 Dysregulation Independently Predicts Disease Specific Survival in Bladder Urothelial Carcinoma. *Journal of Urology* **2015**, *193* (3), 1002–1008. https://doi.org/10.1016/j.juro.2014.09.107.
(75) Zhao, S.; Cheng, L.; Shi, Y.; Li, J.; Yun, Q.; Yang, H. MIEF2 Reprograms Lipid Metabolism to Drive Progression of Ovarian Cancer through ROS/AKT/MTOR Signaling Pathway. *Cell Death Dis* **2021**, *12* (1), 18. https://doi.org/10.1038/s41419-020-03336-6.
(76) Qi, L.; Kolodziej, T.; Rajfur, Z.; Huang, C. Roles of Talin2 in Traction Force Generation, Tumor Metastasis and Cardiovascular Integrity. *Current Protein & Peptide Science 19* (11), 1071–1078.
(77) Markowska, A.; Grybos, A.; Marszałek, A.; Bednarek, W.; Filas, V.; Gryboś, M.; Markowska, J.; Mądry, R.; Więckowska, B.; Nowalińska, D.; Szarszewska, M. Expression of Selected Molecular Factors in Two Types of Endometrial Cancer. *Adv Clin Exp Med* **2021**, *30* (10), 1057–1064. https://doi.org/10.17219/acem/137383.
(78) Zhong, Y.; Long, T.; Gu, C.-S.; Tang, J.-Y.; Gao, L.-F.; Zhu, J.-X.; Hu, Z.-Y.; Wang, X.; Ma, Y.-D.; Ding, Y.-Q.; Li, Z.-G.; Wang, X.-Y. MYH9-Dependent Polarization of ATG9B Promotes Colorectal Cancer Metastasis by Accelerating Focal Adhesion Assembly. *Cell Death Differ* **2021**, *28* (12), 3251–3269. https://doi.org/10.1038/s41418-021-00813-z.
(79) Pi, R.; Chen, Y.; Du, Y.; Dong, S. Comprehensive Analysis of Myoferlin in Human Pancreatic Cancer via Bioinformatics. *Biomed Res Int* **2021**, *2021*, 2602322. https://doi.org/10.1155/2021/2602322.
(80) An, H. J.; Song, D. H.; Koh, H. M.; Kim, Y.-M.; Ko, G. H.; Lee, J.-H.; Lee, J. S.; Yang, J. W.; Kim, M. H.; Seo, D. H.; Jang, S. M.; Kim, D. C. Myoferlin Silencing Inhibits VEGFR2-Mediated Proliferation of Metastatic Clear Cell Renal Cell Carcinoma. *Sci Rep* **2019**, *9*, 12656. https://doi.org/10.1038/s41598-019-48968-7.
(81) Lin, Y.-L.; Wang, Y.-L.; Fu, X.-L.; Li, W.-P.; Wang, Y.-H.; Ma, J.-G. Low Expression of Protocadherin7 (PCDH7) Is a Potential Prognostic Biomarker for Primary Non-Muscle Invasive Bladder Cancer. *Oncotarget* **2016**, *7* (19), 28384–28392. https://doi.org/10.18632/oncotarget.8635.
(82) Qi, Y.; Zhang, Y.; Peng, Z.; Wang, L.; Wang, K.; Feng, D.; He, J.; Zheng, J. SERPINH1 Overexpression in Clear Cell Renal Cell Carcinoma: Association with Poor Clinical Outcome and Its Potential as a Novel Prognostic Marker. *J Cell Mol Med* **2018**, *22* (2), 1224–1235. https://doi.org/10.1111/jcmm.13495.
(83) Kamikawaji, K.; Seki, N.; Watanabe, M.; Mataki, H.; Kumamoto, T.; Takagi, K.; Mizuno, K.; Inoue, H. Regulation of LOXL2 and SERPINH1 by Antitumor MicroRNA-29a in Lung Cancer with Idiopathic Pulmonary Fibrosis. *J Hum Genet* **2016**, *61* (12), 985–993. https://doi.org/10.1038/jhg.2016.99.

(84) Wang, Y.; Zheng, N.; Sun, T.; Zhao, H.; Chen, Y.; Liu, C. Role of TGM2 in T-Cell Lymphoblastic Lymphoma via Regulation of IL-6/JAK/STAT3 Signalling. *Mol Med Rep* **2022**, *25* (3), 76. https://doi.org/10.3892/mmr.2022.12592.
(85) Gu, C.; Cai, J.; Xu, Z.; Zhou, S.; Ye, L.; Yan, Q.; Zhang, Y.; Fang, Y.; Liu, Y.; Tu, C.; Wang, X.; He, J.; Li, Q.; Han, L.; Lin, X.; Li, A.; Liu, S. MiR-532-3p Suppresses Colorectal Cancer Progression by Disrupting the ETS1/TGM2 Axis-Mediated Wnt/β-Catenin Signaling. *Cell Death Dis* **2019**, *10* (10), 739. https://doi.org/10.1038/s41419-019-1962-x.
(86) Malkomes, P.; Lunger, I.; Oppermann, E.; Abou-El-Ardat, K.; Oellerich, T.; Günther, S.; Canbulat, C.; Bothur, S.; Schnütgen, F.; Yu, W.; Wingert, S.; Haetscher, N.; Catapano, C.; Dietz, M. S.; Heilemann, M.; Kvasnicka, H.-M.; Holzer, K.; Serve, H.; Bechstein, W. O.; Rieger, M. A. Transglutaminase 2 Promotes Tumorigenicity of Colon Cancer Cells by Inactivation of the Tumor Suppressor P53. *Oncogene* **2021**, *40* (25), 4352–4367. https://doi.org/10.1038/s41388-021-01847-w.
(87) Cho, S.-Y.; Oh, Y.; Jeong, E. M.; Park, S.; Lee, D.; Wang, X.; Zeng, Q.; Qin, H.; Hu, F.; Gong, H.; Liu, X.; Zhang, G.; Na, D.; Lee, J.; Chae, J.; Suh, Y.-S.; Kong, S.-H.; Lee, H.-J.; Kim, J.-I.; Park, H.; Zhang, C.; Yang, H.-K.; Lee, C. Amplification of Transglutaminase 2 Enhances Tumor-Promoting Inflammation in Gastric Cancers. *Exp Mol Med* **2020**, *52* (5), 854–864. https://doi.org/10.1038/s12276-020-0444-7.
(88) Li, C.; Cai, J.; Ge, F.; Wang, G. TGM2 Knockdown Reverses Cisplatin Chemoresistance in Osteosarcoma. *Int J Mol Med* **2018**, *42* (4), 1799–1808. https://doi.org/10.3892/ijmm.2018.3753.
(89) Fredolini, C.; Pathak, K. V.; Paris, L.; Chapple, K. M.; Tsantilas, K. A.; Rosenow, M.; Tegeler, T. J.; Garcia-Mansfield, K.; Tamburro, D.; Zhou, W.; Russo, P.; Massarut, S.; Facchiano, F.; Belluco, C.; De Maria, R.; Garaci, E.; Liotta, L.; Petricoin, E. F.; Pirrotte, P. Shotgun Proteomics Coupled to Nanoparticle-Based Biomarker Enrichment Reveals a Novel Panel of Extracellular Matrix Proteins as Candidate Serum Protein Biomarkers for Early-Stage Breast Cancer Detection. *Breast Cancer Res* **2020**, *22*, 135. https://doi.org/10.1186/s13058-020-01373-9.
(90) Luo, D.; Zhan, S.; Xia, W.; Huang, L.; Ge, W.; Wang, T. Proteomics Study of Serum Exosomes from Papillary Thyroid Cancer Patients. *Endocrine-Related Cancer* **2018**, *25* (10), 879–891. https://doi.org/10.1530/ERC-17-0547.
(91) Li, H.; Wang, C.; Lan, L.; Behrens, A.; Tomaschko, M.; Ruiz, J.; Su, Q.; Zhao, G.; Yuan, C.; Xiao, X.; Li, B.; Yan, L.; Wu, W.; Li, W.; Chen, J.; He, Y.; Zhang, C. High Expression of Vinculin Predicts Poor Prognosis and Distant Metastasis and Associates with Influencing Tumor-Associated NK Cell Infiltration and Epithelial-Mesenchymal Transition in Gastric Cancer. *Aging (Albany NY)* **2021**, *13* (4), 5197–5225. https://doi.org/10.18632/aging.202440.
(92) Wu, Z.; Wang, S.; Jiang, F.; Li, Q.; Wang, C.; Wang, H.; Zhang, W.; Xue, P.; Wang, S.-L. Mass Spectrometric Detection Combined with Bioinformatic Analysis Identified Possible Protein Markers and Key Pathways Associated with Bladder Cancer. *Gene* **2017**, *626*, 407–413. https://doi.org/10.1016/j.gene.2017.05.054.
(93) Huang, S.; Li, D.; Zhuang, L.; Sun, L.; Wu, J. Identification of Arp2/3 Complex Subunits as Prognostic Biomarkers for Hepatocellular Carcinoma. *Front Mol Biosci* **2021**, *8*, 690151. https://doi.org/10.3389/fmolb.2021.690151.
(94) Pei, L.; Liu, H.; Ouyang, S.; Zhao, C.; Liu, M.; Wang, T.; Wang, P.; Ye, H.; Wang, K.; Song, C.; Zhang, J.; Dai, L. Discovering Novel Lung Cancer Associated Antigens and the Utilization of Their Autoantibodies in Detection of Lung Cancer. *Immunobiology* **2020**, *225* (2), 151891. https://doi.org/10.1016/j.imbio.2019.11.026.

(95) Moradpoor, R.; Gharebaghian, A.; Shahi, F.; Mousavi, A.; Salari, S.; Akbari, M. E.; Ajdari, S.; Salimi, M. Identification and Validation of Stage-Associated PBMC Biomarkers in Breast Cancer Using MS-Based Proteomics. *Front Oncol* **2020**, *10*, 1101. https://doi.org/10.3389/fonc.2020.01101.
(96) Kobayashi, M.; Nagashio, R.; Jiang, S.-X.; Saito, K.; Tsuchiya, B.; Ryuge, S.; Katono, K.; Nakashima, H.; Fukuda, E.; Goshima, N.; Satoh, Y.; Masuda, N.; Saegusa, M.; Sato, Y. Calnexin Is a Novel Sero-Diagnostic Marker for Lung Cancer. *Lung Cancer* **2015**, *90* (2), 342–345. https://doi.org/10.1016/j.lungcan.2015.08.015.
(97) Alijaj, N.; Pavlovic, B.; Martel, P.; Rakauskas, A.; Cesson, V.; Saba, K.; Hermanns, T.; Oechslin, P.; Veit, M.; Provenzano, M.; Rüschoff, J. H.; Brada, M. D.; Rupp, N. J.; Poyet, C.; Derré, L.; Valerio, M.; Banzola, I.; Eberli, D. Identification of Urine Biomarkers to Improve Eligibility for Prostate Biopsy and Detect High-Grade Prostate Cancer. *Cancers (Basel)* **2022**, *14* (5), 1135. https://doi.org/10.3390/cancers14051135.
(98) Okada, R.; Koshizuka, K.; Yamada, Y.; Moriya, S.; Kikkawa, N.; Kinoshita, T.; Hanazawa, T.; Seki, N. Regulation of Oncogenic Targets by MiR-99a-3p (Passenger Strand of MiR-99a-Duplex) in Head and Neck Squamous Cell Carcinoma. *Cells* **2019**, *8* (12), 1535. https://doi.org/10.3390/cells8121535.
(99) Li, W.; Cui, X.; Qi, A.; Yan, L.; Wang, T.; Li, B. MiR-183-5p Acts as a Potential Prognostic Biomarker in Gastric Cancer and Regulates Cell Functions by Modulating EEF2. *Pathology - Research and Practice* **2019**, *215* (11), 152636. https://doi.org/10.1016/j.prp.2019.152636.
(100) Zhang, X.; Hu, L.; Du, M.; Wei, X.; Zhang, J.; Hui, Y.; Chen, C.; Li, G.; Hou, J. Eukaryotic Elongation Factor 2 (EEF2) Is a Potential Biomarker of Prostate Cancer. *Pathol. Oncol. Res.* **2018**, *24* (4), 885–890. https://doi.org/10.1007/s12253-017-0302-7.
(101) Ghafouri-Fard, S.; Khoshbakht, T.; Taheri, M.; Mokhtari, M. A Review on the Role of GAS6 and GAS6-AS1 in the Carcinogenesis. *Pathology - Research and Practice* **2021**, *226*, 153596. https://doi.org/10.1016/j.prp.2021.153596.
(102) Hattori, S.; Kikuchi, E.; Kosaka, T.; Miyazaki, Y.; Tanaka, N.; Miyajima, A.; Mikami, S.; Oya, M. Relationship Between Increased Expression of the Axl/Gas6 Signal Cascade and Prognosis of Patients with Upper Tract Urothelial Carcinoma. *Ann Surg Oncol* **2016**, *23* (2), 663–670. https://doi.org/10.1245/s10434-015-4848-x.
(103) LncRNA CDKN2BAS Aggravates the Progression of Ovarian Cancer by Positively Interacting with GAS6. *European Review*, 2020.
(104) Ceylan, Y.; Akpınar, G.; Doger, E.; Kasap, M.; Guzel, N.; Karaosmanoglu, K.; Kopuk, S. Y.; Yucesoy, I. Proteomic Analysis in Endometrial Cancer and Endometrial Hyperplasia Tissues by 2D-DIGE Technique. *Journal of Gynecology Obstetrics and Human Reproduction* **2020**, *49* (2), 101652. https://doi.org/10.1016/j.jogoh.2019.101652.
(105) Luo, H.-L.; Liu, H.-Y.; Chang, Y.-L.; Sung, M.-T.; Chen, P.-Y.; Su, Y.-L.; Huang, C.-C.; Peng, J.-M. Hypomethylated RRBP1 Potentiates Tumor Malignancy and Chemoresistance in Upper Tract Urothelial Carcinoma. *Int J Mol Sci* **2021**, *22* (16), 8761. https://doi.org/10.3390/ijms22168761.
(106) Ma, J.; Ren, S.; Ding, J.; Liu, S.; Zhu, J.; Ma, R.; Meng, F. Expression of RRBP1 in Epithelial Ovarian Cancer and Its Clinical Significance. *Biosci Rep* **2019**, *39* (7), BSR20190656. https://doi.org/10.1042/BSR20190656.
(107) Liu, S.; Lin, M.; Ji, H.; Ding, J.; Zhu, J.; Ma, R.; Meng, F. RRBP1 Overexpression Is Associated with Progression and Prognosis in Endometrial Endometrioid

Adenocarcinoma. *Diagn Pathol* **2019**, *14*, 7. https://doi.org/10.1186/s13000-019-0784-6.
(108) Liang, X.; Sun, S.; Zhang, X.; Wu, H.; Tao, W.; Liu, T.; Wei, W.; Geng, J.; Pang, D. Expression of Ribosome-Binding Protein 1 Correlates with Shorter Survival in Her-2 Positive Breast Cancer. *Cancer Sci* **2015**, *106* (6), 740–746. https://doi.org/10.1111/cas.12666.
(109) Pan, Y.; Cao, F.; Guo, A.; Chang, W.; Chen, X.; Ma, W.; Gao, X.; Guo, S.; Fu, C.; Zhu, J. Endoplasmic Reticulum Ribosome-Binding Protein 1, RRBP1, Promotes Progression of Colorectal Cancer and Predicts an Unfavourable Prognosis. *Br J Cancer* **2015**, *113* (5), 763–772. https://doi.org/10.1038/bjc.2015.260.
(110) Yang, L.; Hong, Q.; Xu, S.-G.; Kuang, X.-Y.; Di, G.-H.; Liu, G.-Y.; Wu, J.; Shao, Z.-M.; Yu, S.-J. Downregulation of Transgelin 2 Promotes Breast Cancer Metastasis by Activating the Reactive Oxygen Species/Nuclear Factor-KB Signaling Pathway. *Mol Med Rep* **2019**, *20* (5), 4045–4058. https://doi.org/10.3892/mmr.2019.10643.
(111) Chen, C.-L.; Chung, T.; Wu, C.-C.; Ng, K.-F.; Yu, J.-S.; Tsai, C.-H.; Chang, Y.-S.; Liang, Y.; Tsui, K.-H.; Chen, Y.-T. Comparative Tissue Proteomics of Microdissected Specimens Reveals Novel Candidate Biomarkers of Bladder Cancer. *Mol Cell Proteomics* **2015**, *14* (9), 2466–2478. https://doi.org/10.1074/mcp.M115.051524.
(112) Chen, C.-L.; Chung, T.; Wu, C.-C.; Ng, K.-F.; Yu, J.-S.; Tsai, C.-H.; Chang, Y.-S.; Liang, Y.; Tsui, K.-H.; Chen, Y.-T. Comparative Tissue Proteomics of Microdissected Specimens Reveals Novel Candidate Biomarkers of Bladder Cancer. *Mol Cell Proteomics* **2015**, *14* (9), 2466–2478. https://doi.org/10.1074/mcp.M115.051524
(113) Giampaolino, P.; Foreste, V.; Della Corte, L.; Di Filippo, C.; Iorio, G.; Bifulco, G. Role of Biomarkers for Early Detection of Ovarian Cancer Recurrence. *Gland Surg* **2020**, *9* (4), 1102–1111. https://doi.org/10.21037/gs-20-544.
(114) Muinao, T.; Deka Boruah, H. P.; Pal, M. Multi-Biomarker Panel Signature as the Key to Diagnosis of Ovarian Cancer. *Heliyon* **2019**, *5* (12), e02826. https://doi.org/10.1016/j.heliyon.2019.e02826.
(115) Manes, N. P.; Nita-Lazar, A. Application of Targeted Mass Spectrometry in Bottom-up Proteomics for Systems Biology Research. *Journal of Proteomics* **2018**, *189*, 75–90. https://doi.org/10.1016/j.jprot.2018.02.008.
(116) Terkelsen, T.; Krogh, A.; Papaleo, E. CAncer BioMarker Prediction Pipeline (CAMPP)—A Standardized Framework for the Analysis of Quantitative Biological Data. *PLOS Computational Biology* **2020**, *16* (3), e1007665. https://doi.org/10.1371/journal.pcbi.1007665.
(117) Dunn, W.; Burgun, A.; Krebs, M.-O.; Rance, B. Exploring and Visualizing Multidimensional Data in Translational Research Platforms. *Brief Bioinform* **2017**, *18* (6), 1044–1056. https://doi.org/10.1093/bib/bbw080.
(118) Lesur, A.; Bernardin, F.; Koncina, E.; Letellier, E.; Kruppa, G.; Dittmar, G. Highly Multiplexed Targeted Plasma Proteomics Quantifies Several Hundred Blood Proteins in Serum from Colorectal Carcinoma Patients. 29.

(1') *X! Tandem*. https://www.thegpm.org/tandem/ (acessado em 19-05-2022).
(2') *PeptideProphet Main*. http://peptideprophet.sourceforge.net/ (acessado em 19-05-2022).
(3') *ProteinProphet Main*. http://proteinprophet.sourceforge.net/ (acessado em 19-05-2022).
(4') *Skyline*. https://skyline.ms/project/home/software/Skyline/begin.view (acessado em 19-05-2022).
(5') *Perseus*. https://maxquant.net/perseus/ (acessado em 19-05-2022).