

UNIVERSIDADE FEDERAL DO PARANÁ

**ALEXANDRE LUIZ KORTE DE AZEVEDO**

PREDIÇÃO BIOINFORMÁTICA DO INTERATOMO DA PROTEÍNA *STRESS-INDUCED PHOSPHOPROTEIN 1* (STIP1) E VALIDAÇÃO DE SUA EXPRESSÃO EM DIFERENTES SUBTIPOS DE CÂNCER DE MAMA

CURITIBA

2019

ALEXANDRE LUIZ KORTE DE AZEVEDO

PREDIÇÃO BIOINFORMÁTICA DO INTERATOMA DA PROTEÍNA STRESS-  
INDUCED PHOSPHOPROTEIN 1 (STIP1) E VALIDAÇÃO DE SUA EXPRESSÃO  
EM DIFERENTES SUBTIPOS DE CÂNCER DE MAMA

Monografia apresentada como requisito parcial  
à obtenção do título de Bacharel,  
Curso de Ciências Biológicas,  
Universidade Federal do Paraná.

Orientadora: Jaqueline Carvalho de  
Oliveira

CURITIBA

2019

## RESUMO

O câncer de mama é o segundo tipo mais incidente de câncer, atrás apenas do câncer de pele não melanoma do Brasil e com maiores taxas de mortalidade em mulheres. No Brasil, o número de casos e as taxas de mortalidade apresentam um aumento nos últimos anos, principalmente devido à dificuldade de acesso a métodos de diagnóstico precoce e de tratamentos eficazes. Neste aspecto, a proteômica e a interatômica têm contribuído com o estudo da expressão proteica e seus conjuntos de interação, visando aumentar o entendimento e buscando marcadores para a doença. Neste estudo, a *Stress Induced Phosphoprotein 1* (STIP1), uma proteína considerada como possível biomarcadora para o câncer de mama, foi analisada juntamente com proteínas diferencialmente expressas no tumor primário de mama em relação ao tecido não tumoral, no contexto de seu interatoma e dos processos biológicos no qual ela e suas parceiras atuam. O interatoma predito possui alta robustez e, em conjunto com as análises de enriquecimento funcional, foi capaz de apontar que a STIP1 está relacionada com a regulação de diversos processos de relevância no câncer de mama. Além disso, foi possível relacionar sua super expressão com o aumento de agressividade da doença, validado por western blotting semi quantitativo em linhagens de subtipos contrastantes.

Palavras-chave: Câncer de mama. Interatômica. Proteômica. Stress Induced Phosphoprotein 1.

## ABSTRACT

The breast cancer is the second most common type of cancer behind only to non melanoma skin cancer in Brazil, and with the greatest rate of mortality in women. The number of cases and mortality rate show an increase in the last few years, mainly due access difficulties to early diagnosis methods and effective treatments. In this aspect, the proteomic and interactomic has contributed with the study of protein expression and their interaction network, aiming to increase the knowledge and looking for new biomarkers to breast cancer. In the present study, the *Stress Induced Phosphoprotein 1* (STIP1), a protein indicated to be a biomarker to breast cancer, was analyzed together with a set of differentially expressed proteins in mammary tumor when compared to non tumor tissue, in the context of our interaction network and the biological process in which STIP1 and their partners operate. The predicted interactome have great robustness and, together with the enrichment analysis, was capable of appoint the that STIP1 is related to the regulation of several relevant processes to breast cancer. Besides that, was possible correlate the super expression of STIP1 with aggressiveness increase of disease, validate by semi quantitative western blotting in cell lines of breast cancer contrasting subtypes.

Key words: Breast cancer. Proteomic. Interactomic. *Stress Induced Phosphoprotein*

1.

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>6</b>
<b>2</b>	<b>REVISÃO DE LITERATURA</b>	<b>7</b>
2.1	ASPECTOS EPIDEMIOLÓGICOS DO CÂNCER NO MUNDO E NO BRASIL	7
2.2	ASPECTOS ANATÔMICOS E FISIOLÓGICOS DA MAMA	8
2.3	ASPECTOS ETIOLÓGICOS DO CÂNCER DE MAMA	9
2.4	A PROTEÔMICA, INTERATÔMICA E BIOINFORMÁTICA NA PESQUISA DO CÂNCER.	11
2.5	STRESS-INDUCED PHOSPHOPROTEIN 1 (STIP1) E O CÂNCER	12
<b>3</b>	<b>METODOLOGIA</b>	<b>14</b>
3.1	OBTENÇÃO DO EXTRATO PROTEICO DE TECIDO	15
3.2	QUANTIFICAÇÃO E IDENTIFICAÇÃO LIVRE DE MARCAÇÃO POR ESPECTOMETRIA DE MASSAS (LFQ-MS)	15
3.3	ANÁLISE BIOINFORMÁTICA DE ANOTAÇÃO FUNCIONAL E ENRIQUECIMENTO DE VIAS	16
3.4	VALIDAÇÃO DE EXPRESSÃO DA STIP1 ATRAVÉS DE WESTERN BLOTTING.	17
3.4.1	OBTENÇÃO DAS CÉLULAS DE LINHAGEM	17
3.4.2	OBTENÇÃO DO EXTRATO PROTEICO DE LINHAGEM	17
3.4.3	SEPARAÇÃO DAS PROTEÍNAS POR SDS-PAGE	18
3.4.4	ETAPA DE WESTER BLOT	18
3.4.5	ANÁLISE DOS DADOS DO WESTERN BLOTTING	19
<b>4</b>	<b>RESULTADOS</b>	<b>20</b>
4.1	PROTEÍNAS SELECIONADAS PARA O ESTUDO.	20
4.2	ANÁLISE DE ENRIQUECIMENTO FUNCIONAL PARA PROCESSOS BIOLÓGICOS E FUNÇÕES MOLECULARES.	23
4.3	PREDIÇÃO DO INTERATOMA DA STIP1	25
4.4	VALIDAÇÃO DA EXPRESSÃO DA STIP1 NOS SUBTIPOS TUMORAIS LUMINAL-A E TRIPLO-NEGATIVO.	27
<b>5</b>	<b>DISCUSSÃO</b>	<b>29</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>30</b>
<b>7</b>	<b>REFERÊNCIAS</b>	<b>31</b>

## 1 INTRODUÇÃO

O avanço da ciência vem proporcionando um aumento geral na expectativa de vida da população, seja favorecendo a cura de doenças, aumentando o acesso à alimentação de qualidade, ou provendo tecnologias que facilitem a vida cotidiana: Mas fato é que o câncer é uma questão que foge a esta realidade de melhoras, com especial ênfase ao câncer de mama, uma vez que a taxa de mortalidade é crescente, bem como suas taxas de sobrevida são cada vez menores, de maneira geral. No Brasil, segundo o Instituto Nacional de Câncer (INCA), a taxa de sobrevida para mulheres diagnosticadas com câncer de mama, ou seja, o tempo de vida médio que essas mulheres têm após o diagnóstico, vem diminuindo, e são esperados novos 60 mil casos de câncer de mama para o biênio 2018-2019: Unidos, estes dois fatores indicam a relevância e a necessidade de novos estudos que possam favorecer o entendimento da doença.

O câncer é um conjunto complexo e heterogêneo de doenças, caracterizado pelo desenvolvimento e proliferação anormal de células, que podem vir a invadir outros tecidos: Em especial, o câncer de mama possui diversas classificações, sejam histológicas ou moleculares, e também diversos fatores que afetam sua ocorrência, como a dinâmica hormonal, uso e exposição a agentes que favoreçam a ocorrência da doença, menarca precoce e outros, tornando não apenas complicada a caracterização do estado clínico, mas também a criação de métodos de diagnóstico precoce de tratamentos terapêuticos não invasivos e eficazes. Porém, neste sentido, as abordagens pós-genômicas de proteômica e interatômica têm alcançado resultados relevantes.

A interatômica e a proteômica têm proposto biomarcadores, novos alvos de terapia e maior entendimento sobre a biologia do câncer de mama, pautando-se no fato de que as proteínas, por participarem ativamente dos processos biológicos que, quando alterados levam à tumorigênese, se constituem como alvos de terapia e marcadores moleculares acessíveis para a rotina clínica. Uma ferramenta que tem sido aliada aos estudos de proteômica e interatômica com grande êxito é a bioinformática, permitindo que a grande quantidade de dados gerados por estas tecnologias pós-genômicas seja analisada, facilitando a interpretação do seu significado biológico. Um dos grandes alvos de estudo para estas abordagens têm sido as proteínas que possuem expressão diferencial no tecido tumoral da neoplasia

mamária em relação ao tecido não tumoral correspondente, visto que as mudanças fisiológicas e moleculares que levam ao fenótipo canceroso podem, em parte, ser explicadas por estas desregulações de expressão

Tendo como principal objetivo de estudo a proteína *Stress-Induced Phosphoprotein 1* (STIP1), uma proteína que tem se mostrado de relevância na pesquisa do câncer, este estudo visa ampliar os conhecimentos, através de abordagens proteômicas e de interatômica, aliadas à bioinformática, sobre a função da proteína STIP1 no contexto da tumorigênese mamária, pretendendo predizer qual o conjunto de interações que compõem o interatoma desta proteína no câncer de mama, quais as funções biológicas podem estar sendo alteradas pelas proteínas diferencialmente expressas que fazem parte deste interatoma, e também validar a expressão da STIP1 nos subtipos mais contrastantes da doença.

## **2 REVISÃO DE LITERATURA**

### **2.1 ASPECTOS EPIDEMIOLÓGICOS DO CÂNCER NO MUNDO E NO BRASIL**

Nas últimas décadas, com avanço da ciência, sobreveio o desenvolvimento de tecnologias e estratégias que possibilitam maior sobrevida ou a cura de doenças que, antes, eram incuráveis e/ ou de alta letalidade, favorecendo, de maneira geral, o aumento da expectativa de vida da população mundial. Contudo, o câncer segue uma tendência contrária: Há um aumento crescente no número de casos da doença, devido principalmente ao crescimento da população mundial e pela prevalência dos principais fatores de risco, muitos dos quais estão associados ao desenvolvimento socioeconômico, mas também à desinformação a respeito da doença e às dificuldades que as mulheres possuem para o acesso aos métodos de diagnóstico e tratamentos adequados em estágios iniciais, contribuindo para que muitas das pacientes cheguem na clínica em estágios avançados da doença, com prognóstico e sobrevida desfavoráveis.

Disto resulta uma grande disparidade nas taxas de sobrevida entre países: Nos Estados Unidos da América e na Austrália, por exemplo, a taxa de sobrevida em cinco anos após o diagnóstico é de 90% enquanto que na África do Sul, que é de 40% (ALLEMANI, 2018). Fato é que o câncer tem sido uma das maiores causas de

morte no mundo: em 2018, por exemplo, foram 18.1 milhões de casos (IARC, 2018), com aproximadamente 9.6 milhões de mortes atribuídas à doença.

As taxas de sobrevivência no Brasil no período de cinco anos apresentaram queda, passando de 76.9% entre 2005 – 2009, para 75.2% nos anos 2010 – 2014 o que representa uma queda de aproximadamente 2%. Segundo o Instituto Nacional de Câncer (INCA, 2018), estima-se que 600 mil novos casos de câncer ocorram no Brasil no biênio 2018-2019, sendo que aproximadamente 60 mil destes compreendem o câncer de mama. Entre os tipos de câncer, o de mama é o mais incidente na população feminina brasileira, com exceção do câncer de pele não melanoma, e apresenta as maiores taxas de mortalidade, atrás apenas do câncer de pele não melanoma. Ressalta-se que a tendência no aumento da incidência deste tipo de neoplasia maligna é uma realidade mundial (INCA, 2018).

## 2.2 ASPECTOS ANATÔMICOS E FISIOLÓGICOS DA MAMA

Ao nascimento, a mama feminina e masculina não apresentam diferenças perceptíveis a nível estrutural ou funcional, permanecendo na infância como um conjunto de ductos lactíferos arranjados de maneira convergente e circunferencial em relação aos mamilos (ELLIS e MAHADEVAN, 2013). Do período da infância até a puberdade, a mama feminina irá permanecer sem alterações, o qual irá durar por aproximadamente até dois anos antes da menarca.

A partir desta fase, ocorrem diversas alterações morfológicas e funcionais no epitélio e estroma mamários, resultantes de respostas a várias influências hormonais no microambiente da mama, sendo o estrogênio e a progesterona os hormônios mais influentes (STINGL, 2010).

A presença do estrogênio é um dos sinais do início da puberdade feminina, e é produzido pelos folículos do ovócito quando os mesmos começam a amadurecer; a partir deste momento, ele será o desencadeador do desenvolvimento de diversas características sexuais primárias e secundárias da mulher, iniciando o desenvolvimento da mama feminina na puberdade. A progesterona também possui função vital para o desenvolvimento do tecido mamário, estimulando a proliferação celular e a formação de alvéolos e ductos (STINGL, 2010). Outros fatores também são responsáveis pelo desenvolvimento deste tecido, como os fatores de



crescimento EGF (*Epidermal growth factor*) e VEGF (*Vascular endothelial growth factor*), além das vias de sinalização WNT e Notch (ERCAN *et al.*, 2011).

A mama, uma vez madura, será constituída basicamente de estroma (adiposo, conjuntivo, vasos e nervos) e tecido parenquimático, que é o conjunto de glândulas circundado pelo estroma. O parênquima (ou tecido glandular) é formado por ácinos, os quais representam a menor parte do tecido e são responsáveis pela secreção de leite: Os conjuntos formados por ácinos são chamados de lóbulos, que em conjunto, serão conhecidos como lobos mamários. A secreção que se forma nos lobos mamários é transportada pelos ductos até a papila, pelo sistema lobular-ductal, e da papila estará pronta para ser exteriorizada pelos mamilos, (CASTRO, 2011), cumprindo a principal função das mamas, que é a lactação.

Na menopausa, a diminuição do aporte de hormônios como a progesterona e o estrogênio irá diminuir a estimulação do tecido, gerando um atrofiamento e declínio na presença de tecido glandular, o qual será substituído gradualmente por tecido adiposo (ELLIS & MAHADEVAN, 2013).

Percebe-se desta maneira que o tecido mamário apresenta diversos ciclos de estimulação hormonal, proliferação e involução, de maneira que a constante transformação que ocorre neste tecido favorece o aparecimento de alterações que podem resultar em uma proliferação incorreta de células epiteliais dos lóbulos ou ductos, gerando, respectivamente carcinomas lobulares ou ductais (ERCAN *et al.*, 2011). De fato, diversas vias vitais para o controle fisiológico e morfológico do desenvolvimento da mama encontram-se desregulados e/ou mutadas no câncer, como as vias de sinalização mediadas por WNT/beta-catenina, os receptores NRG3 de EGFR, e a sinalização Notch. Além disso, a desregulação das vias hormonais de progesterona e estrogênios também pode contribuir para alterar o desenvolvimento normal da mama e, conseqüentemente, para um maior risco de câncer (ERCAN *et al.*, 2011).

### 2.3 ASPECTOS ETIOLÓGICOS DO CÂNCER DE MAMA

O câncer é um conjunto de doenças que consiste no crescimento desordenado de células anormais, as quais podem invadir tecidos e órgãos, podendo se alastrar e causar novos tumores em um processo denominado metástase (MARTIN *et al.*, 2013). A evolução de um tumor pode ser rápida e

agressiva e, neste caso, de difícil controle, podendo levar a morte. Dependendo de fatores como agressividade, classificação, localização e estágio de desenvolvimento do tumor no momento em que se inicia o tratamento a sua evolução pode ser fatal (INCA, 2018).

As causas para o surgimento de um tumor são múltiplas, incluindo fatores genéticos, como mutações pontuais e disfunções imunes e hormonais, e também fatores ambientais externos, por meio da exposição à carcinógenos, como poluentes ambientais. Uma pequena fração dos tumores (em torno de 5%) são hereditárias e, neste caso, as alterações genéticas são transmitidas de uma geração a outra (APOSTOLOU *et al.*, 2013).

O câncer de mama apresenta múltiplos fatores envolvidos em sua etiologia, como idade da primeira menstruação, nuliparidade, primeira gravidez após os 30 anos, uso de anticoncepcionais e de terapias de reposição hormonal (TRH), bem como menopausa após os 55 anos, idade, obesidade e sedentarismo, os quais contribuem para a desregulação dos mecanismos naturais do desenvolvimento mamário (MACMAHON *et al.*, 1973).

As classificações dos carcinomas mamários indicam aspectos distintos de relevância clínica, auxiliando no diagnóstico, procedimentos terapêuticos e no prognóstico da doença. Perou e colaboradores, em 2000, inicialmente descreveram quatro subtipos, em que o luminal era classificado como um subtipo único. Nesta classificação inicial identificava-se: luminal/ RE+, correspondente à linhagem luminal; tipo basal (*basal-like*), podendo incluir a linhagem mioepitelial; HER2+, com superexpressão de HER2; e semelhante ao normal (*normal-like*). Em seguida, SORLIE *et al.*, 2001 e 2003 distinguiram no subtipo luminal dois outros: luminal A (tumores de baixo grau, RP+ e HER2-) e B (tumores de grau elevado, RP+/-, HER2+/- e alto grau de Ki-67, um marcador de proliferação celular).

Contudo, a análise por métodos moleculares para identificar os subtipos referidos apresenta um alto custo e pouca acessibilidade para a maioria dos hospitais e clínicas, sendo relevante simplificá-la. Assim, a classificação histopatológica, baseada na presença de marcadores e em aspectos imunohistoquímicos condizentes com cada subtipo, é uma alternativa mais acessível e de menor custo para classificar os tumores mamários, distinguindo-os em cinco subtipos: Luminal A, luminal B (HER2 negativo), luminal B (HER2 positivo), HER2

positivo (não luminal) e triplo negativo. Estes subtipos são classificados quanto a expressão dos receptores de estrogênio, progesterona e HER2, e de KI67, tornando o diagnóstico mais acessível à rotina clínica (GOLDHIRSCH *et al.*, 2013).

Considerando sua complexidade, o câncer de mama se apresenta como uma doença de difícil prevenção e tratamento, já que a grande variação entre os fatores de risco e as características genéticas que compreendem sua etiologia torna a completa prevenção da doença quase impossível, sendo relevante a busca por alternativas diagnósticas e terapêuticas mais eficazes.

#### 2.4 A PROTEÔMICA, INTERATÔMICA E BIOINFORMÁTICA NA PESQUISA DO CÂNCER.

Mesmo com a crescente taxa de incidência na população mundial, a atenção veiculada pela mídia e do contínuo esforço do meio científico, o câncer de mama continua apenas parcialmente compreendido, sendo que a ampliação deste conhecimento poderia favorecer novos métodos de diagnóstico precoce e de tratamentos terapêuticos não invasivos e eficazes, com conseqüente alta mortalidade relacionada ao acometimento deste tipo de doença (INCA, 2018). Assim, ressalta-se a importância de estudos que objetivem o desenvolvimento de novos métodos de diagnóstico, prognóstico e tratamento, baseados em uma ampliação do conhecimento sobre a doença. Neste sentido, estudos proteômicos aliados à interatômica têm alcançado resultados de grande relevância (ERLICH *et al.*, 2007; WANG *et al.*, 2010; CHAO *et al.*, 2013).

No cenário pós-genômico, dentre as diversas abordagens “ômicas”, têm-se destacado a proteômica e a interatômica. Os métodos proteômicos permitem a caracterização do conteúdo proteico de amostras biológicas, identificando proteínas expressas sob condições fisiológicas específicas (proteoma) e possibilitando elucidar com elevada precisão mecanismos intrínsecos da doença, muitos dos quais são responsáveis pelo acometimento e ganho de malignidade do tumor (SRINIVAS *et al.*, 2002). A análise das interações proteína-proteína, possibilitada através da interatômica, também constitui uma importante estratégia de pesquisa da doença, uma vez que os processos celulares mediados por proteínas geralmente envolvem diversas interações, com complexos multiproteicos e interações binárias essenciais, ou seja, as proteínas e suas interações participam ativamente de processos

biológicos de grande importância no microambiente celular e, eventualmente, resultam em alvos terapêuticos acessíveis, sendo estas uma das vantagens de estudá-las no contexto do câncer (SHRUTHI *et al.*, 2013). Alguns métodos possuem proeminência na pesquisa envolvendo a proteômica-interatômica do câncer. Destaca-se a tecnologia de espectrometria de massas com ionização por eletrospray (ESI-MS/MS), que tem sido amplamente utilizada no intuito de fornecer dados confiáveis para a identificação do conteúdo proteico e a mensuração da expressão destas proteínas (CHEN *et al.*, 2007); e a técnica de Western blotting (WB), que possibilita a validação da expressão de proteínas específicas, e é utilizada repetidamente em validações de proteínas diferencialmente expressas e de interações entre proteínas (FU-JUN *et al.*, 2012).

As abordagens “ômicas” possuem, sabidamente, técnicas que geram uma quantidade muito elevada de dados, sendo necessário desenvolver métodos que permitam organizar, classificar e interpretar, biologicamente, o significado dos dados adquiridos: A bioinformática, desta forma, é utilizada nos estudos “ômicos” para interpretar os complexos sistemas formados pelas proteínas e seu padrão de interações, auxiliando a integrar, expandir e elaborar respostas que expliquem a relevância da expressão diferencial e as interações no contexto do câncer de mama e outros (CRISTONI & MAZZUCA, 2011).

## 2.5 STRESS-INDUCED PHOSPHOPROTEIN 1 (STIP1) E O CÂNCER

Uma das proteínas que tem se destacado na pesquisa científica do câncer devido a sua participação na regulação de vários processos e vias biológicas, é a co-chaperona *Stress Induced Phosphoprotein 1* (STIP1). Essa proteína possui por três domínios TRP (*Tetratricopeptide Repeat*) e por dois domínios DP (*aspartic acid-proline repeat domain*) (FIGURA 1). Estes domínios são responsáveis pela sua capacidade de se ligar e interagir com outras proteínas, o que permite sua atuação como moduladora de atividade de diversas outras proteínas, através da regulação conformacional e estabilização de seus alvos, de maneira a mediar as interações moleculares de proteínas que possuam uma região de ligação para estes domínios TRP (DANIEL *et al.*, 2011). Especificamente, a STIP1 atua durante a fase intermediária, modulando a entrada de seus alvos em suas respectivas vias (DA COSTA, 2015).

FIGURA 1 – Esquema da proteína STIP1 e a distribuição de seus domínios



FONTE: Adaptado de BAINBUR-HUDSON et al., 2015

A STIP1, em mamíferos, é amplamente distribuída na célula, sendo encontrada no citoplasma em condições normais, devido à fosforilação pela proteína cdc2 cinase, mas podendo ser transferida para o núcleo quando fosforilada em outro sítio, pela proteína CKII (SHERIL *et al.*, 2008). Há evidências de que as interações da STIP1 com outras proteínas variam conforme a sua localização, e que a sua transferência não apenas serve para que ela possa agir individualmente no núcleo, mas que também sua translocação é responsável pelo transporte de várias de suas parceiras (principalmente *heat-shock proteins*) para o núcleo, aonde formam complexos de interação responsáveis por respostas moleculares ao choque térmico, como por exemplo, a sinalização para a entrada da célula na fase G1/S do ciclo celular (DANIEL *et al.*, 2011). Cita-se que a STIP1 também pode ser observada no complexo de Golgi, associada à membrana celular e no meio extracelular (HAJJ *et al.*, 2013 *apud* DA COSTA 2015).

O mais conhecido mecanismo de atuação da STIP1 é a formação de complexos com proteínas da família *heat-shock proteins*, os quais são responsáveis por regulação de *splicing* de RNA, transcrição, dobramento de proteínas, tradução de sinais e regulação do ciclo celular (TSAI *et al.*, 2016).

Em diversos tipos de câncer, a STIP1 tem sido relacionada a processos de importância no contexto do câncer, como a proliferação celular em linhagens de glioma e ovário através de ativação das vias de sinalização MAPK, ERK e ALK (ERLICH *et al.*, 2007; WANG *et al.*, 2010); invasão celular em câncer pancreático via ativação de MMP (WALSH *et al.*, 2011); promoção de metástase em câncer gástrico através da via de sinalização WNT/ $\beta$ -catenina (HUANG, 2018); entre outros.

No câncer de ovário, a STIP1 tem sido descrita como um possível biomarcador (CHAO *et al.*, 2013) e, em 2018, a proteína foi pela primeira vez indicada como possível biomarcador para o câncer de mama, principalmente por apresentar-se super expressa em tecidos mamários tumorais em relação aos não tumorais; e em linfonodos metastáticos em relação ao tumor primário. Além disso, há correlação entre a expressão da STIP1 com características clínicas e de sobrevida; Quanto maior a expressão, pior o prognóstico e menor a sobrevida (WU *et al.*, 2018; DA COSTA, 2015 ).

Contudo, apesar dos recentes estudos a respeito da STIP1, são poucos os trabalhos que descrevem a função desta proteína no câncer de mama ou as proteínas com as quais a mesma interage neste tecido. Assim, a falta de informações a respeito da STIP1 no câncer de mama indica a necessidade de se investigar a importância desta proteína neste contexto, seu interatoma e os processos celulares potencialmente alterados devido a sua super expressão. Deste modo, este trabalho visa contribuir para esta demanda de informações e verificar se a proteína STIP1 e suas parceiras de interação são relevantes também no câncer de mama. Caso se confirme a importância da STIP1 neste contexto, uma verificação de sua expressão em diferentes subtipos do câncer de mama será realizada, visando determinar se as diferenças de agressividade entre os subtipos podem estar relacionadas a uma maior ou menor expressão desta proteína.

### **3 METODOLOGIA**

Os dados de expressão proteica utilizados neste estudo são provenientes de um estudo prévio realizado no Laboratório de Citogenética Humana e Oncogenética do Departamento de Genética da UFPR, onde este estudo também foi realizado. Foram utilizados dados de quatorze amostras de tecido mamário (sete tumorais pareados com sete não tumorais) de sete pacientes, com idade média de  $56 \pm 14.63$  anos e diagnóstico de Carcinoma Ductal Invasivo. As pacientes foram atendidas no Hospital Nossa Senhora das Graças (Curitiba/PR). Desta forma, os itens 3.1 e 3.2 se referem à metodologia utilizada na tese de doutorado da Dra. Talita Helen Bombardelli Gomig. O projeto foi aprovado por Comitê de Ética do Setor de Ciências da Saúde da UFPR: CAAE n. 19870319.3.0000.0102.

### 3.1 OBTENÇÃO DO EXTRATO PROTEICO DE TECIDO

Para extração e solubilização das proteínas, 25 µg foram separados e a estes acrescidos 200 µL de tampão de lise (4% SDS, 0,1 M Tris-HCl pH 7,6 e 0,1 M DTT), seguido por um período de incubação de 5 minutos no gelo, para que o tampão possa agir.

Em seguida, o tecido foi homogeneizado com auxílio do TissueLyser, em três ciclos intercalados de aquecimento a 95°C, em banho seco, por 5 minutos. Após, a viscosidade dos extratos proteicos foi diminuída utilizando-se o banho ultrassônico por aproximadamente 1 hora. Em seguida, cada amostra foi centrifugada em 10.000 x g a 4°C por 10 minutos, para separação das proteínas do restante do conteúdo celular. Após a centrifugação, o sobrenadante contém o conteúdo proteico das células, o qual foi separado e armazenado em -80°C até sua utilização.

Ao final do processo, foram obtidos os extratos proteicos dos tecidos tumorais (T) e não tumorais (Na) que foram utilizados no presente estudo.

### 3.2 QUANTIFICAÇÃO E IDENTIFICAÇÃO LIVRE DE MARCAÇÃO POR ESPECTOMETRIA DE MASSAS (LFQ-MS)

As amostras obtidas após a extração foram preparadas para a espectrometria de massas. A técnica FASP (Filter-Aided Sample Preparation) (WIŚNIEWSKI *et al.*, 2009) foi utilizada previamente à mensuração da concentração dos extratos proteicos, e esta foi obtida por *Qubit 2.0 Fluorometer* (Life Technologies). Após, uma breve corrida eletroforética foi realizada em géis unidimensionais de poliacrilamida de empilhamento e de resolução (separação) a 10%. A digestão das proteínas foi em gel utilizando a enzima tripsina. A concentração e purificação dos peptídeos foi efetuada em ponteiras com coluna de fase reversa Zip Tip Pipette Tips (Millipore).

A LFQ-MS se realizou na Plataforma de Espectrometria de Massas do Instituto Carlos Chagas (ICC/Fio Cruz Paraná), através de nanocromatografia líquida de fase reversa no sistema *Easy-nLC 1000* (Thermo Fisher Scientific) seguida da ionização *electrospray* (ESI) e análise por espectrometria de massas sequencial (MS/MS) no espectrômetro LTQ-Orbitrap XL ETD (Thermo Fisher Scientific).

Triplicatas técnicas foram analisadas para cada triplicata biológica, e os espectros de massa foram analisados através do programa MaxQuant versão 1.5.8.3, sendo que a identificação das proteínas ocorreu utilizando-se o banco de dados Uniprot (UniProtKB - Homo sapiens, versão 24 de maio de 2017, 70.939 entradas).

A análise dos dados de LFQ-MS foi realizada no programa Perseus1.5.6.0 para a identificação de diferenças significativas de expressão ( $p < 0,05$ ) entre as proteínas identificadas nos tecidos tumorais e não tumorais.

### 3.3 ANÁLISE BIOINFORMÁTICA DE ANOTAÇÃO FUNCIONAL E ENRIQUECIMENTO DE VIAS

Os dados provenientes da espectrometria de massas e da análise pelo *Perseus* resultaram em uma lista de proteínas diferencialmente expressas entre os tecidos tumorais (T) e não tumorais (Na). Utilizando-se da métrica de Fold Change, foi possível identificar quais proteínas foram observadas com sub ou super expressão no tumor quando comparados aos tecidos não tumorais, sendo que a medida de corte o FoldChange de  $-1.5 > X > 1.5$ . Contudo, ressalta-se que as proteínas que não foram identificadas neste limiar também foram mantidas nas análises subsequentes, pois poderiam ser parte do interatoma. Assim, visando enriquecer o interatoma da STIP1 dentro do total de proteínas obtidas, foi realizada uma seleção com base na escolha de conjuntos de proteínas que estivessem comumente observadas em *hallmarks* do câncer (MSigDB v. 6.2), sendo o critério de seleção a significância com base no valor de  $p < 0.05$  para os conjuntos e, tendo em vista que o foco do trabalho é o interatoma da STIP1, foram selecionados apenas os conjuntos que contivessem esta proteína.

Plataformas como o *GSEA* e *FunRich* foram utilizadas para a anotação funcional dos genes que codificam as proteínas selecionadas, no intuito de obter informações sobre processos biológicos e funções moleculares (de acordo com termos da ontologia *Gene Ontology*) que estivessem enriquecidos em nosso conjunto, conforme os bancos de dados *KEGG* e *Reactome*.

A predição de interatoma se deu através de plataformas *online* como *STRING* versão 10.5, *Genemania* e outros, resultando em uma representação do conjunto de interação proteína-proteína das proteínas selecionadas, incluindo a STIP1 obrigatoriamente.



### 3.4 VALIDAÇÃO DE EXPRESSÃO DA STIP1 ATRAVÉS DE WESTERN BLOTTING.

#### 3.4.1 Obtenção das células de linhagem

Para a validação da expressão da proteína STIP1, duas linhagens representativas de cada subtipo (Luminal A e Triplo Negativo) serão cultivadas e utilizadas para a extração das proteínas.

Foram utilizadas as linhagens T47D e MCF-7 (Luminal A), e as linhagens MDA MB 468 e MDA MB 231 (Triplo Negativo), que foram cultivadas em Meio Dulbecco MEM (DMEM) com HEPES, 5% FBS e 1% L-Glutamina (Aminoácidos não essenciais) + insulina a 5% CO<sup>2</sup> e com meio RPMI-1640 com 5% FBS, 1% p/s e 1% Glutamina + insulina, a 5% CO<sup>2</sup>, respectivamente.

Após atingir a confluência esperada, as células foram desaderidas com tripsina 0.05% e EDTA em PBS 0.04%, centrifugadas por 4 minutos a 1500 rpm e, após tratados, os *pellets* foram armazenados a -80°C.

#### 3.4.2 Obtenção do extrato proteico de linhagem

Para extração e solubilização das proteínas, os *pellets* celulares obtidos (1x10<sup>6</sup> células) foram descongelados, centrifugados a 72 g em 4°C por 5 minutos e os resquícios de meio de cultura/armazenamento foram eliminados. Após, foram acrescentados 100 µL de tampão de lise (20 mM Tris 137 mM NaCl, 1% triton, 2 mM EDTA, HCl pH 8.0 e Inibidor de protease Roche) e uma rápida homogeneização foi realizada com a ponteira até o pellet de células se desfazer, seguindo-se por um período de incubação de 5 minutos no gelo.

Em seguida, cada amostra foi ser centrifugada em 10.000 x g a 4°C por 20 minutos, para separação das proteínas do restante do conteúdo celular. Após a centrifugação, o sobrenadante com o conteúdo proteico das células, o qual foi separado e quantificado através do método de Bradford (BRADFORD, 1976).

Ao final do processo, foram obtidos os extratos proteicos das quatro linhagens utilizadas no estudo para a validação via western blotting.

### 3.4.3 Separação das proteínas por SDS-PAGE

Para a separação das proteínas foi realizada uma corrida de eletroforese SDS- PAGE, de acordo com Harlow e Lane (1988) modificado. A técnica se baseia em romper as ligações covalentes entre as proteínas e possíveis moléculas que a elas podem estar ligadas e na neutralização das cargas inerentes às proteínas, de maneira que o único parâmetro que influenciará sua separação no gel será a massa molecular individual de cada proteína.

Os géis utilizados no SDS-PAGE são compostos de acrilamida/ bis-acrilamida, SDS e Tris, sendo os agentes catalizadores TEMED e APS, e separados em duas fases: Para o gel de empilhamento será utilizado um upper gel 8% e para o gel de separação, um lower gel 10% em um suporte de 1.5mm.

Em cada poço foi aplicada uma concentração final de 45 µg de amostra, em uma proporção padrão entre extrato, tampão de lise e tampão de amostra (0.3 M Tris-HCl, 50% glicerol, 10% SDS, azul de bromofenol e água ultrapura), aquecida a 95°C por 5 minutos, previamente à aplicação no gel. A corrida será feita em 120V por 1:30 horas, até o marcador de massa molecular atingir o final do gel. Foi utilizado o marcador PageRuler Prestained Protein Ladder (Thermo Scientific).

### 3.4.4 Etapa de western Blot

Após a separação eletroforética, as proteínas foram transferidas do gel para uma membrana de nitrocelulose na cuba de transferência Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad), com tampão de transferência (Tris, glicina, água ultrapura) por 2 horas a 100V, visando que as proteínas estejam em uma superfície acessível aos anticorpos. A eficiência da transferência foi verificada mediante coloração das proteínas pelo corante Ponceau (0.2% de corante, 3% ácido acético glacial). As membranas foram lavadas por duas vezes com TBST (1.5 M NaCl, 200 mM Tris-HCl e 0,05% Tween 20) até completa remoção do corante e, então, incubadas com 5 mL de TBST/leite (TBS-T contendo 5% de leite liofilizado, desnatado e desengordurado - Molico, Nestlé) por 40 minutos para bloqueio dos sítios livres. Após este período, a membrana foi secada e acrescida de 4 ml de solução de TBST/leite com o anticorpo primário anti-STIP1 (*anti-rabbit*) numa

concentração de 1:3000. A incubação ocorreu *overnight*, por no mínimo 16h, em câmara fria a 4°C, sob agitação orbital.

Após, a membrana foi retirada da câmara fria e lavada três vezes com TBST, e em seguida foi feita a incubação com anticorpo secundário *anti-rabbit*. A diluição do anticorpo será realizada em 1:4000 em TBST/leite, e a incubação ocorreu após a adição de 4 mL desta solução, em temperatura ambiente, por 1 hora sob agitação orbital. Como normalizador, foi utilizado um anticorpo anti-  $\beta$ -actina (*anti-rabbit*) em concentração 1:1000. Posteriormente, seguiu-se para o processo de revelação.

Como preparo para a revelação, as membranas foram lavadas por três vezes de 5 minutos em TBST, e então cobertas com 1 ml de substrato quimioluminescente (*GE Healthcare*) por cerca de 10 minutos e alocadas entre uma folha plástica em cassete fotográfico. O excesso do substrato foi removido e, em quarto escuro, a membrana foi exposta a um filme autoradiográfico (Kodak) por períodos de tempo determinados até a obtenção do sinal, este verificado após revelação e fixação do filme.

#### 3.4.5 Análise dos dados do *western blotting* – Análise semi - quantitativa

O resultado da revelação do filme após o *western blotting* foi analisado no programa *ImageJ* (versão. 1.52r), medindo-se a área do pico de densitometria de cada banda do filme: A intensidade deste pico corresponde à quantidade da proteína STIP1 na amostra. Após normalizar estes valores com base nos valores de densitometria encontrados para o controle (Proteína  $\beta$ -Actina, gene ACTB), foi possível comparar as expressões de abundância relativa entre os subtipos. As comparações foram previamente definidas, sendo que a linhagem MCF7 foi pareada com a linhagem MDA – MB – 231, e a linhagem T – 47D com a MDA – MB – 468.

Para verificar se a abundância relativa entre os subtipos era significativa, o Teste T foi utilizado, gerando um valor de p, e para quantificar a diferença, foi utilizada a medida de *Fold Change*, com significância para  $-1.5 > X > 1.5$ .

## 4 RESULTADOS

### 4.1 PROTEÍNAS SELECIONADAS PARA O ESTUDO

Após os processos de extração, quantificação, identificação das proteínas através da plataforma de espectrometria de massas e tratamento estatístico no programa Perseus, chegou-se a uma lista de proteínas, muitas das quais diferencialmente expressas.

A análise de enriquecimento funcional resultou na seleção de 57 proteínas de interesse dentre o total identificado, seguindo os critérios estabelecidos na metodologia. Destas, 23 (40%) apresentaram diferença de expressão nas comparações TxNa além do Fold Change  $-1.5 > X > 1.5$ , estabelecido como linha de corte.

Entre os genes que codificam essas proteínas, 13 foram classificados como envolvidos no desenvolvimento do câncer (*ACO1*, *FGA*, *GBE1*, *HNRNPK*, *HSP90AA1*, *HSPA4*, *IDH2*, *NUMA1*, *PDHB*, *PPP2R1A*, *PRKACA*, *PRKAR1A*, *SDHB*), e sendo dois destes classificados como oncogenes (*PRKACA*, *IDH2*). A Tabela 1 apresenta as 57 proteínas selecionadas, indicando as superexpressas (em rosa) e as subexpressas (em verde).

TABELA 1 – Relação das 57 proteínas selecionadas.

Símbolo do gene	Fold Change (Comparação TxNa)	Expressão
<i>ACLY</i>	1.057243895	Aumento
<i>ACO1</i>	-1.501537079	Sub expressa nos Tumerais
<i>AGL</i>	1.508716348	Super expressa nos tumerais
<i>AGT</i>	-1.761603685	Sub expressa nos Tumerais
<i>ATP5A1</i>	0.884190571	Aumento
<i>CIQBP</i>	1.563046995	Super expressa nos tumerais

<i>CAT</i>	-2.358020271	Sub expressa nos Tumerais
<i>DCTN2</i>	-0.909896529	Diminuição
<i>DLST</i>	-0.7276746	Diminuição
<i>FGA</i>	-1.093731357	Diminuição
<i>GBE1</i>	-2.71731096	Sub expressa nos Tumerais
<i>GFPT1</i>	1.513692619	Super expressa nos tumerais
<i>GNB2</i>	-1.034751864	Aumento
<i>HNRNPK</i>	1.906709783	Super expressa nos tumerais
<i>HSP90AA1</i>	2.545568071	Super expressa nos tumerais
<i>HSP90B1</i>	2.033294333	Super expressa nos tumerais
<i>HSPA12A</i>	-4.713232525	Sub expressa nos Tumerais
<i>HSPA4</i>	1.385305155	Diferencialmente expressa
<i>HSPA5</i>	2.273105024	Super expressa nos tumerais
<i>HSPA8</i>	2.14187981	Super expressa nos tumerais
<i>HSPA9</i>	1.401641071	Aumento
<i>HSPB6</i>	-6.477436276	Sub expressa nos Tumerais
<i>HSPD1</i>	2.121990738	Super expressa nos tumerais
<i>IDH2</i>	2.135811295	Super expressa nos tumerais
<i>ILK</i>	-2.203257011	Sub expressa nos Tumerais
<i>IQGAP1</i>	0.952145262	Aumento
<i>ITGA6</i>	-1.573030624	Sub expressa nos Tumerais
<i>KNG1</i>	-2.06926939	Sub expressa nos Tumerais
<i>NCL</i>	2.170861524	Super expressa nos tumerais
<i>NUMA1</i>	1.409589533	Aumento
<i>OGDH</i>	-1.212481912	Diminuição
<i>PDHB</i>	-0.853272764	Diminuição
<i>PGM1</i>	-1.471823729	Diminuição
<i>PPP2R1A</i>	0.519397738	Diferencialmente expressa

<i>PRKACA</i>	-1.399775424	Diminuição
<i>PRKARIA</i>	0.867069548	Aumento
<i>PRKAR2B</i>	-3.415413538	Aumento
<i>PSMA5</i>	0.515415022	Aumento
<i>PSMA7</i>	0.536879826	Aumento
<i>PSMC3</i>	0.403251914	Aumento
<i>PSMC4</i>	0.257195167	Aumento
<i>PSMC5</i>	0.530868167	Aumento
<i>PSMD11</i>	0.571118864	Aumento
<i>PSMD12</i>	0.570173376	Aumento
<i>PSMD2</i>	0.916113327	Aumento
<i>PSMD3</i>	0.661881923	Aumento
<i>PSMD5</i>	0.90333834	Aumento
<i>PSMD7</i>	0.919626524	Aumento
<i>RPL24</i>	2.114531418	Super expressa nos tumorais
<i>SDHB</i>	-0.800911857	Diminuição
<i>SLC25A12</i>	-0.543194071	Diminuição
<i>SLC25A13</i>	1.394718614	Aumento
<i>STIP1</i>	1.4543960	Aumento
<i>SLC25A5</i>	1.2616244	Aumento
<i>SUCLA2</i>	-1.487049352	Diminuição
<i>TUBB4B</i>	1.716053786	Super expressa nos tumorais
<i>UGP2</i>	-1.280370286	Diminuição

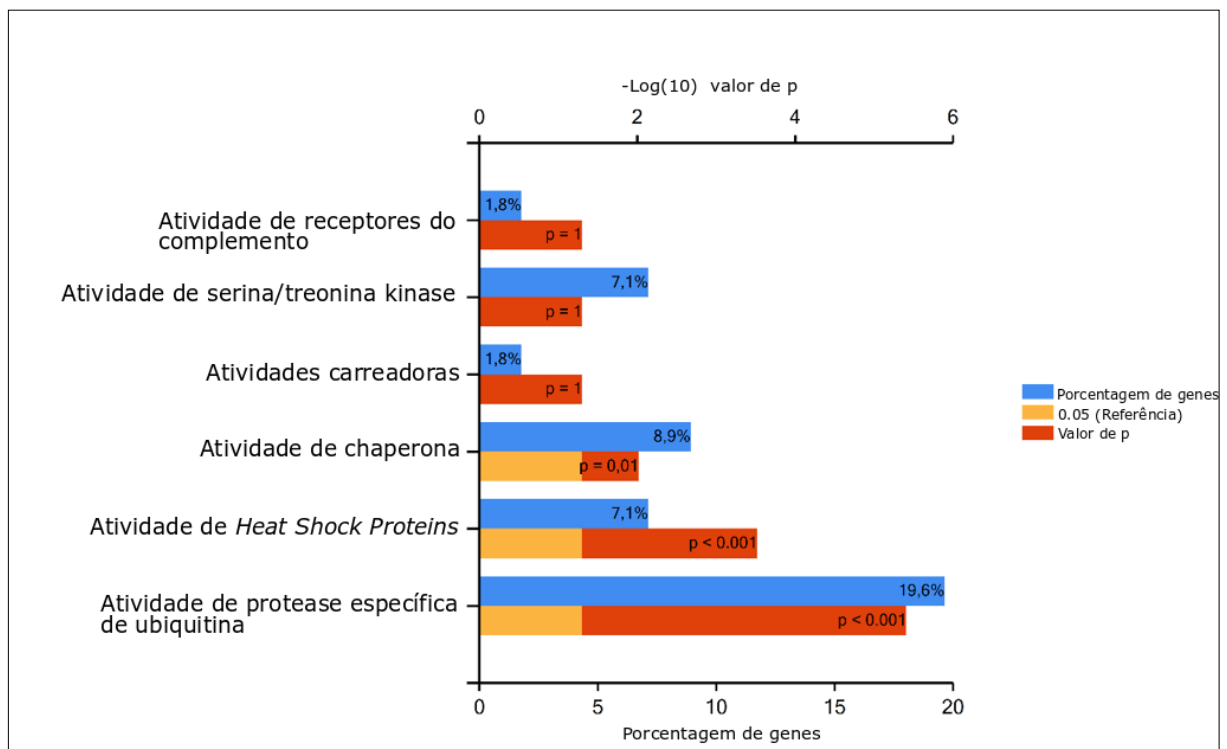
FONTE: O autor, 2019

## 4.2 ANÁLISE DE ENRIQUECIMENTO FUNCIONAL PARA PROCESSOS BIOLÓGICOS E FUNÇÕES MOLECULARES.

A análise de enriquecimento funcional é uma estratégia de estudo baseada na análise e interpretação de um conjunto de moléculas (no caso deste estudo, proteínas) e que permite determinar quais processos biológicos e funções moleculares estão enriquecidas em determinado tecido sob uma condição fisiológica específica. Os programas FunRich (v. 3.1.3) e GSEA permitem este tipo de análise classificando biologicamente o conjunto de proteínas de acordo termos do Gene Ontology (GO).

As proteínas utilizadas neste trabalho foram as 57 proteínas selecionadas, que compõem a predição do interatoma da STIP1. O resultado da análise de enriquecimento para a categoria GO de funções moleculares está representado na FIGURA 2.

FIGURA 2 – Funções moleculares enriquecidas para as proteínas preditas no interatoma de STIP1.

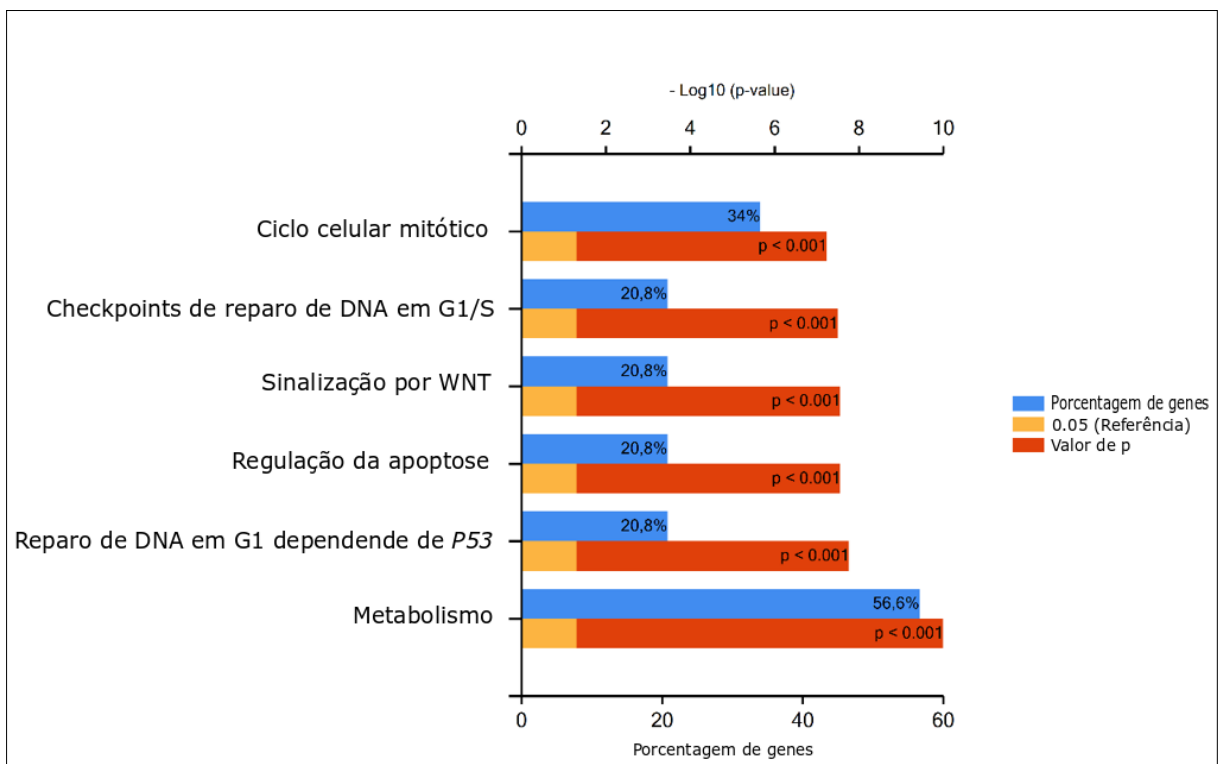


Fonte: O autor, 2019.

O resultado indica que, das seis principais funções moleculares encontradas como enriquecidas no conjunto de proteínas preditas no interatoma da STIP1, três obtiveram um valor de  $p$  significativo ( $p < 0.05$ ): Atividade de chaperona, atividade de *Heat Shock Protein* e atividade de protease específica de ubiquitina. Este resultado demonstra as principais atividades exercidas por estas proteínas, e assim, a maneira como elas atuam para exercer suas funções biológicas. Como veremos a seguir, estas três atividades são vitais para a regulação dos processos biológicos encontrados nas análises.

Para o mesmo conjunto, foram evidenciados os principais processos biológicos enriquecidos, segundo a terminologia GO (FIGURA 3).

FIGURA 3 – Processos Biológicos enriquecidos para as proteínas preditas no interatoma de STIP1.



Fonte: O autor, 2019.

Foram identificados seis processos biológicos com maior enriquecimento, todas com valor de  $p$  significativo ( $p < 0.001$ ): Ciclo celular mitótico, checkpoints de reparo de DNA em G1/S, sinalização por WNT, regulação de apoptose, reparo de DNA em G1 dependente de p53 e metabolismo. Considerando que as proteínas do



conjunto são diferencialmente expressas em tecido tumoral em relação ao tecido mamário não tumoral, é possível considerar que estes processos biológicos podem estar desregulados (*up* ou *down* regulados), interferindo negativamente no desenvolvimento das células, e favorecendo a ocorrência do câncer.

Assim, a análise de enriquecimento funcional indica que o interatoma da STIP1 representa um conjunto de proteínas que participa ativamente de vias biológicas que, uma vez desreguladas pelas expressões alteradas de alguns de seus componentes, podem contribuir para o processo do câncer.

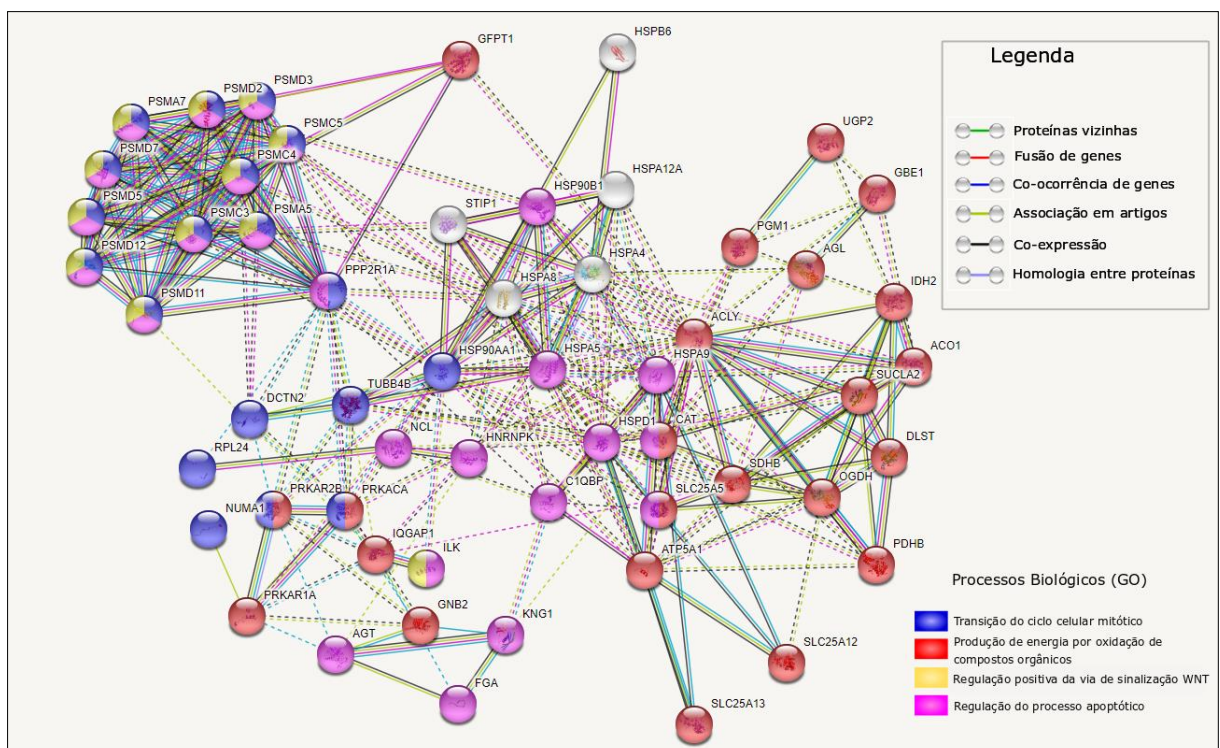
#### 4.3 PREDIÇÃO DO INTERATOMA DA STIP1

Para que a descrição e consequente compreensão da função biológica de uma determinada proteína seja devidamente alcançada, o conhecimento a respeito dos parceiros específicos de interação desta proteína é necessário, levando em conta que o conceito de “função proteica” depende de uma questão hierárquica dentro do contexto de interações que ocorrem na célula e nas principais vias e processos biológicos em geral. Além disso, o correto funcionamento de uma proteína depende de fatores como recrutamento de proteínas parceiras, controle alostérico, interações com receptores e moléculas sinalizadoras, modificações pós transcricionais, além das condições de microambiente celular, de modo que compreender não apenas a estrutura de uma proteína, mas sim as interações e relações inter-proteínas, é o objetivo principal a ser atingido para se entender a função de uma proteína.

Para identificar quais as interações proteína-proteína que fazem parte do interatoma da STIP1 dentro da neoplasia mamária, no intuito de compreender como a STIP1 desempenha suas funções, uma predição do perfil de interações desta proteína foi realizada no programa STRING v. 10.5, a partir do conjunto de 57 proteínas selecionadas neste estudo. Utilizando-se de informações oriundas de diversos bancos de dados, o programa une informações de associações físicas estáveis e/ou transientes, compartilhamento de substrato, retransmissão de informações e dados de co-expressão, baseados em dados curados de variedade bioquímica, biofísica e de técnicas genéticas experimentais em geral. Com estes dados, é possível usar algorítmicos próprios do programa para analisar o conjunto de proteínas e atribuir um valor de significância para as interações.

O resultado desta predição (FIGURA 4) representa o interatoma da STIP1 juntamente com as 56 proteínas selecionadas, colocando mais próximas umas das outras aquelas com as maiores valores de significância na interação. Cada círculo na representação corresponde a uma proteína, e os traços que ligam esses círculos representam as diferentes categorias de interação entre as proteínas. O valor de p para o interatoma, calculado a partir do enriquecimento de interações proteína-proteína (enrichment PPI), foi de  $1.0e-16$ . Este valor indica que a distribuição dos nós (as proteínas) não é aleatória, e que o número de interações é maior do que o esperado, indicando também a robustez desta predição e a alta possibilidade da predição ser uma representação próxima da realidade que ocorre nas células.

FIGURA 4 – Resultado da predição de interatoma subdividido em clusters.



Fonte: O Autor, 2019.

Foi possível agrupar as proteínas deste interatoma em quatro grandes clusters segundo os processos biológicos comuns às proteínas, seguindo os termos do GO: *Cluster 1* - Transição do ciclo celular mitótico; *Cluster 2* - produção de energia por oxidação de compostos orgânicos; *Cluster 3* - regulação positiva da via de sinalização WNT e *Cluster 4* - regulação do processo apoptótico. Estes quatro clusters representam uma categorização das funções biológicas específicas

encontradas na análise de enriquecimento funcional (*Cluster 1* - Ciclo celular mitótico, checkpoints de reparo de DNA em G1/S e reparo de DNA em G1 dependente de *p53*; *Cluster 2* – Metabolismo; *Cluster 3* – Sinalização por WNT; *Cluster 4* – Regulação de apoptose). A correspondência entre os processos biológicos encontrados de maneira independente em diferentes análises (de enriquecimento funcional e de perfil de interações) era esperada, uma vez que o conjunto de proteínas é o mesmo, mas mesmo assim demonstra a robustez do interatoma.

É interessante ressaltar que a proteína STIP1 apresenta como principais parceiras de interação as *Heat Shock Proteins* (HSPs), como esperado, e que a STIP1 e as HSPs ocupam posições centrais dentro do interatoma. O significativo valor de *p* para o enriquecimento de PPI deste interatoma permite afirmar que essa configuração da disposição espacial da STIP1 e das HSPs não é ao acaso, e o fato de que todos os clusters possuem sempre alguma interação com estas proteínas, direta ou indiretamente, nos permite concluir que a STIP1, através de suas interações com as proteínas HSPs, pode contribuir para a regulação dos processos biológicos aqui representados pelos quatro clusters periféricos, indicando que não apenas o interatoma aqui proposto possui grande relevância para o contexto da tumorigênese mamária, mas também que a proteína STIP1 possui uma função especial como reguladora de todos estes processos.

#### 4.4 VALIDAÇÃO DA EXPRESSÃO DA STIP1 NOS SUBTIPOS TUMORAIS LUMINAL-A E TRIPLO-NEGATIVO.

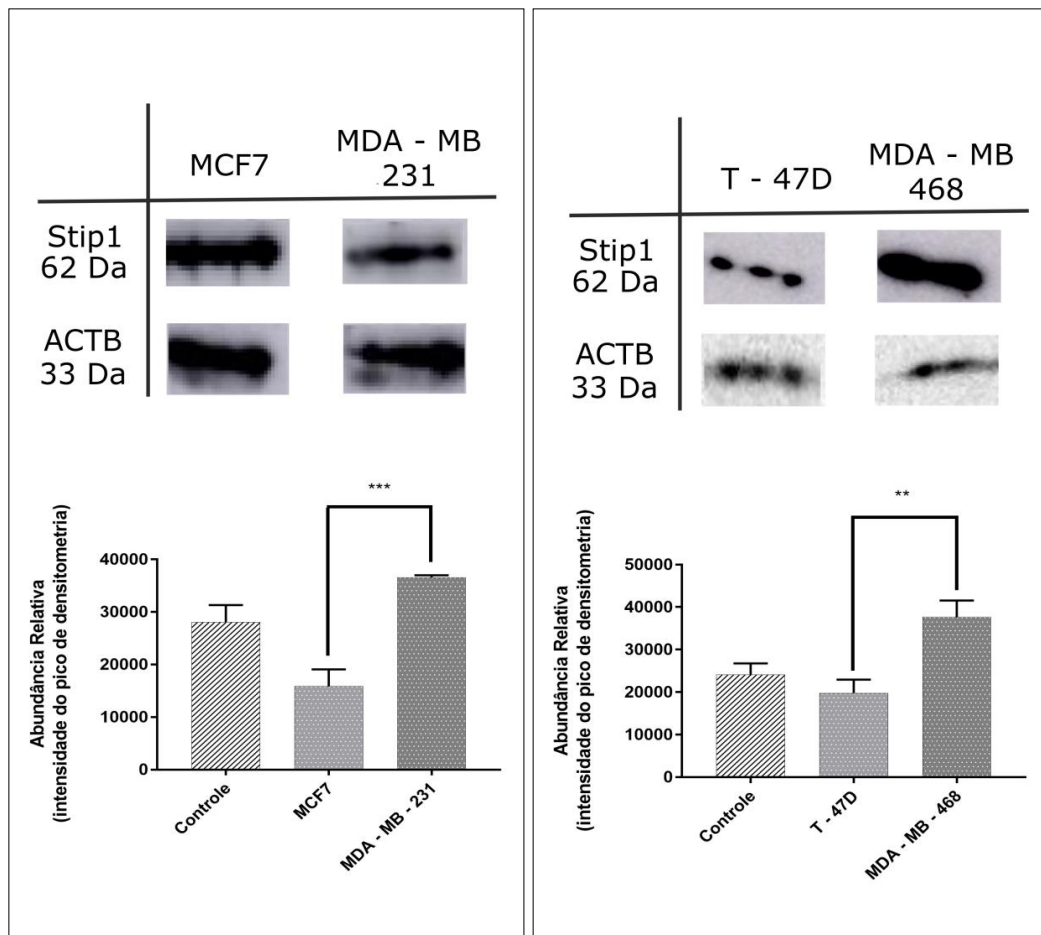
Os resultados anteriores confirmam a hipótese de que a proteína STIP1 e seu interatoma, assim como ocorre em outros tipos de câncer, possuem uma função de alta relevância para processos que influenciam no ganho de malignidade nas células mamárias e para o câncer de mama como um todo.

Porém, considerando que mesmo dentro da mesma neoplasia (a mamária), existem diversos subtipos com diferentes graus de agressividade, levanta-se a questão de que o aumento de expressão da STIP1 pode estar mais acentuado em subtipos mais agressivos, considerando que a STIP1 contribui para este estado de aumento de agressividade.

Assim, para responder a esta pergunta, a técnica de *Western Blotting* semi-quantitativo foi utilizada para mensurar se havia diferença de expressão entre os subtipos do câncer de mama. Os subtipos selecionados para a validação foram os considerados menos e mais agressivos, respectivamente luminal A e triplo negativo. As linhagens MCF7 e T – 47D representam o subtipo luminal A e as linhagens MDA – MB – 231 e MDA – MB – 468 triplo negativo.

O resultado da comparação, considerando a expressão relativa da proteína STIP1 em cada subtipo está representado na FIGURA 5.

FIGURA 5 – Comparação da expressão da STIP1 em linhagens do subtipo Luminal A e Triplo Negativo.



Fonte: O autor, 2019.

A diferença de expressão da STIP1 entre as linhagens foi significativa, como pode ser observado na tabela 2.

TABELA 2 – Significância da diferença de expressão da STIP1 na comparação entre subtipos.

	<i>Fold Change</i>	Valor de p
MDA – MB – 468 X T – 47D	1,900208309	0.0036 ***
MDA – MB – 231 X MCF7	2,305366251	0.0004 **

Fonte: O autor, 2019

## 5 DISCUSSÃO

A proteína STIP1 tem sido repetidamente relacionada com processos biológicos de relevância em diversos tipos de câncer, mas para o câncer de mama poucos são os trabalhos que analisam sua importância no contexto da tumorigênese e sua contribuição para o desenvolvimento e progressão da doença: Os poucos trabalhos existentes indicam uma maior expressão desta proteína em tecido tumoral e linfonodo metastático quando comparado com o tecido normal, e a relação desta proteína com baixa sobrevida.

Assim, este trabalho objetivou contribuir para o entendimento da função biológica desta proteína no câncer de mama, a partir da predição do interatoma da STIP1, determinação dos processos biológicos alterados pela STIP1 e suas parceiras de interação, e pela verificação e validação da expressão desta proteína em dois dos diferentes subtipos de câncer de mama.

Após seleção via métodos bioinformáticos, o interatoma da STIP1 foi predito, tendo 57 proteínas, das quais 23 são diferencialmente expressas além do limite de corte de *fold change*, 13 possuindo seus genes classificados como envolvidos do câncer e 2 oncogenes. Este interatoma foi separado em quatro *clusters* principais, que compartimentalizam as proteínas conforme seus valores de p para enriquecimento de interação proteína-proteína e os processos biológicos em comum entre as proteínas. A proteína STIP1 e suas principais parceiras, as HSPs, ocupam posições centrais no interatoma, e possuem interações com todos os quatro clusters. A posição central destas proteínas, aliado com as funções moleculares encontradas na análise de enriquecimento funcional para a categoria função molecular (Atividade de chaperonas, atividade de *Heat shock*, e atividade de

protease específica de Ubiquitina), que são atividades tipicamente de regulação, sugerindo que a STIP1 e as HSPs estão regulando a atividade das outras proteínas do interatoma, e conseqüentemente dos processos biológicos em que elas participam.

Os quatro *clusters* observados representam, cada um, um processo biológico, reunindo e categorizando os resultados da análise de enriquecimento funcional. As quatro categorias são: Transição do ciclo celular mitótico, produção de energia por oxidação de compostos orgânicos, regulação positiva da via de sinalização WNT e regulação do processo apoptótico. Todos estes processos estão altamente relacionados com o câncer e seus *hallmarks*, sugerindo que o interatoma regulado pela STIP1 possui alta relevância para a tumorigênese mamária, uma vez que estes processos tendem a estar alterados devido à expressão diferenciada de boa parte das proteínas deste interatoma.

Considerando a importância biológica da STIP1 no câncer de mama, indicada neste estudo, levantou-se a hipótese de que a maior expressão da STIP1 acompanharia os subtipos de maior agressividade. Para responder a esta questão, a expressão da STIP1 foi verificada através de *western blotting* semi quantitativo, no subtipo mais agressivo, o Triplo Negativo, em relação ao menos agressivo (Luminal A). O resultado demonstra uma expressão significativamente maior desta proteína no subtipo mais agressivo, o que sugere uma correlação entre o aumento de expressão desta proteína com a maior agressividade do subtipo mais agressivo da doença.

## 6 CONCLUSÃO

A STIP1 teve uma predição de interatoma bastante robusta, a qual permitiu propor suas parceiras de interação e, juntamente com as análises de enriquecimento funcional, os processos nos quais elas estão envolvidas. Assim, foi possível determinar que a STIP1 não apenas possui uma expressão aumentada no tumor mamário em relação ao tecido não tumoral (segundo os dados da espectrometria de massas), mas pode contribuir, juntamente com suas parceiras de interação, para o acometimento e progressão da doença, e para a diferença de malignidade entre os subtipos.

## 7 REFERÊNCIAS

APOSTOLOU, P.; FOSTIRA, F. Hereditary Breast Cancer: The Era of New Susceptibility Genes. Hindawi Publishing Corporation. BioMed Research International Volume 2013, Article ID 747318, 11 pages.

BAINBUR-HUDSON, S.; EDKINS, A, L.; BLATCH, G, L. Hsp70/Hsp90 Organising Protein (Hop): Beyond Interactions with Chaperones and Prion Proteins. Subcell Biochem. 2015;78:69-90.

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

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

CHAO A.; LAI, C, H.; TSAI, C,L.; HSUEH S.; HSUEH C.; Tumor Stress-Induced Phosphoprotein1 (STIP1) as a Prognostic Biomarker in Ovarian. Cancer. PLoS ONE 8(2): e57084. doi:10.1371/journal.pone.0057084. 2013

CHEN, Y.; CHOONG, L.; LIN, Q.; PHILP, R.; WONG, C.; ANG, B.; TAN, Y.; LOH, M, S.; HEW, C.; SHAH, N.; DRUKER, B, J.; CHONG, P.; LIM Y. Differential Expression of Novel Tyrosine Kinase Substrates during Breast Cancer Development. Molecular & Cellular Proteomics December, 2007.

CRISTONI, S.; MAZZUCA, S. Bioinformatics Applied to Proteomics. Available from: <<https://www.intechopen.com/books/systems-and-computational-biology-bioinformatics-and-computational-modeling/bioinformatics-applied-to-proteomics>> Acesso em 16/11/2019.

DA COSTA, G, G. Análise comparativa de expressão proteica entre amostras tumorais e não tumorais de carcinomas mamários. Tese de Doutorado. UFPR, 2015.

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

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

ERLICH, R. B.; KAHN, S. A.; LIMA, F. R. S.; MURAS, A. G.; MARTINS, R. A. P.; LINDEN, R.; CHIARINI, L. B.; MARTINS, V. R.; NETO, V. M. STI1 Promotes Glioma Proliferation Through MAPK and PI3K Pathways. Glia, v. 55, p. 1690–1698, 2007.

FU-JUN, L.; SHAO-HUA, J.; XIAO-FANG, S. Differential proteomic analysis of pathway biomarkers in human breast cancer by integrated bioinformatics. Oncology Letters 4: 1097-1103, 2012.

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

HARLOW, E.; LANE, D. Antibodies: A laboratory manual. Cold Spring Harbor Laboratory,1988.

HUANG L.; ZHAI E.; CAI S. LIN Y.; LIAO, J.; SUI,H, J.; PENG; XU L.; CHEN L.; ZENG, Z. Stress-inducible Protein-1 promotes metastasis of gastric cancer via Wnt/ $\beta$ - catenin signaling pathway. Journal of Experimental & Clinical and Cancer Research, 2018.

IARC. Latest global cancer data: Cancer burden rises to 18.1 million new cases and 9.6 million cancer deaths in 2018. Press release n° 263. 2018.

INCA. Estimativa 2018: incidência de câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva. Rio de Janeiro, 2018.



MACMAHON, B.; COLE, P.; BROWN, J. Etiology of Human Breast Cancer: A Review, *JNCI: Journal of the National Cancer Institute*, Volume 50, Issue 1, 1 January 1973, Pages 21–42.

PEROU, C., SØRLIE, T., EISEN, M. Molecular portraits of human breast tumours. *Nature* 406, 747–752 (2000).

SHERIL, D.; GRAEME, B.; LONGSHAW, V, M.; SÖTI, C.; CSERMELY, P.; BLATCH, G, L. Nuclear translocation of the phosphoprotein Hop (Hsp70/Hsp90 organizing protein) occurs under heat shock, and its proposed nuclear localization signal is involved in Hsp90 binding. *Biochimica et Biophysica Acta* 1783 1003–1014, 2008.

SHRUTHI, B, S.; VINODHKUMAR, P.; SELVAMANI. Proteomics: A new perspective for cancer. *Adv Biomed Res* 2016;5:67. 2013

SØRLIE, T.; TIBSHIRANI, R.; PARKER, J.; HASTIE, H.; MARRON, J, S.; NOBEL, A.; DENG, S.; JOHNSEN,H.; PESICH, R.; GEISLER, S.; DEMETER, J.; PEROU, C, M.; LØNNING, P,E.; BROWN,P,O.; BØRRESEN-DALE, A.; BOTSTEIN,D. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proceedings of the National Academy of Sciences* Jul 2003.

SRINIVAS,P, R.; VERMA, M.; ZHAO,Y.; SRIVASTAVA, S. Proteomics for Cancer Biomarker Discovery. *Clinical Chemistry* 1160–1169. 2002

STINGL, J. Estrogen and Progesterone in normal mammary gland development and in cancer. *HORM CANC* (2011) 2:85–90

TSAI C.; CHAO A.; JUNG S.; TSAI C.; LIN C.; CHEN S.; SUE S.; WANG, T.; WANG H.; LAI C. Stress-induced phosphoprotein-1 maintains the stability of JAK2 in cancer cells. *Oncotarget*, Vol. 7, No. 31, 2016.

WANG, T. H.; CHAO, A.; TSAI, C. L.; CHANG, C. L.; CHEN, S. H.; LEE, Y. S.; CHEN, J. K.; LIN, Y. J.; CHANG, P. Y.; WANG, C. J.; CHAO, A. S.; CHANG, S. D.; CHANG, T. C.; LAI, C. H.; WANG, H. S. Stress-induced phosphoprotein 1 as a

secreted biomarker for human ovarian cancer promotes cancer cell proliferation. *Mol. Cell. Proteomics*, v. 9, n. 9, 1873-1884, 2010.

WALSH, N.; O'DONOVAN, N.; KENNEDY, S.; HENRY, M.; MELEADY, P.; CLYNES, M.; DOWLING, P. Identification of pancreatic cancer invasion-related proteins by proteomic analysis. *Proteome Sci.*, v. 7, n. 3, 2009.

WU, R.; LIU, R.; PENG, P.; QIU, H.; XIONG, H.; YU, S.; HUANG, X.; ZHANG, H.; ZHUANG, L. Tumor stress-induced phosphoprotein 1 as a prognostic biomarker for breast cancer. *Ann Transl Med* ;6(15):302, 2018.