



Universidade Federal do Rio de Janeiro  
Instituto de Química  
Programa de Pós-Graduação em Bioquímica



PRISCILA FERREIRA DE AQUINO

AVALIAÇÃO PROTEÔMICA DA MARGEM DE RESSECÇÃO DE  
PACIENTES COM CÂNCER GÁSTRICO POR  
ESPECTROMETRIA DE MASSAS

Rio de Janeiro  
2015

PRISCILA FERREIRA DE AQUINO

AVALIAÇÃO PROTEÔMICA DA MARGEM DE RESSECÇÃO DE  
PACIENTES COM CÂNCER GÁSTRICO POR  
ESPECTROMETRIA DE MASSAS

Tese de doutorado apresentada ao Programa de Pós-graduação em Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de doutor em Ciências (Bioquímica).

Orientadores: Dr. Paulo Costa Carvalho

Dr. Gilberto Barbosa Domont

Rio de Janeiro  
2015

A657

Aquino, Priscila Ferreira de

Avaliação proteômica da margem de ressecção de pacientes com câncer gástrico por espectrometria de massas. / Priscila Ferreira de Aquino. – Rio de Janeiro: UFRJ, 2015.

127 f.

Tese (doutorado) – Universidade Federal do Rio de Janeiro, Instituto de Química, 2015.

Orientadores: Paulo Costa Carvalho e Gilberto Barbosa Domont.

1. Bioquímica. 2. Proteômica. 3. Câncer gástrico. 4. Espectrometria de massas. I. Carvalho, Paulo Costa. II. Domont, Gilberto Barbosa. III. Universidade Federal do Rio de Janeiro. Instituto de Química. Programa de Pós-Graduação em Bioquímica. IV. Título.

CDD 547

AVALIAÇÃO PROTEÔMICA DA MARGEM DE RESSECÇÃO DE  
PACIENTES COM CÂNCER GÁSTRICO POR  
ESPECTROMETRIA DE MASSAS

Tese de doutorado apresentada ao Programa de Pós-graduação em Bioquímica do Instituto de Química da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de doutor em Ciências (Bioquímica).

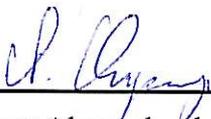
Aprovado em 26 de outubro de 2015.

COMISSÃO EXAMINADORA



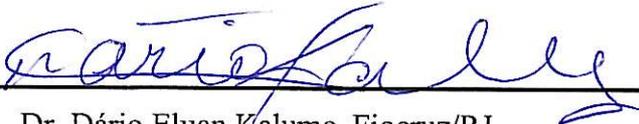
---

Dr. Paulo Costa Carvalho, Fiocruz/PR – Presidente



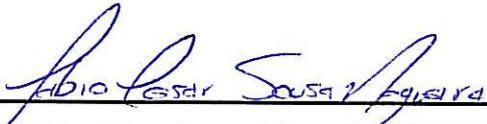
---

Dr. Donat Alexander de Chapeaurouge, Fiocruz/RJ



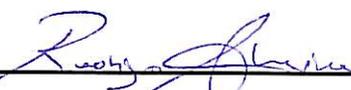
---

Dr. Dário Eluan Kalume, Fiocruz/RJ



---

Dr. Fabio César Souza Nogueira, DBQ/IQ-UFRJ



---

Dr. Rodrigo Volcan Almeida, DBQ/IQ-UFRJ

*Dedico aos meus pais, Cirene e Manuel, que  
sempre me apoiaram em meus estudos.*

## AGRADECIMENTOS

Em primeiro lugar a Deus, pela minha vida, saúde e por guiar meus passos proporcionando grandes oportunidades de aprendizado.

À minha família, pela compreensão nos momentos difíceis, apoio nas minhas decisões, dedicação, carinho, amor e educação.

Aos meus orientadores Dr. Paulo Costa Carvalho e Dr. Gilberto Barbosa Domont pela orientação, apoio, conhecimentos transmitidos, atenção, paciência, confiança, amizade, dedicação e ajuda durante esses anos de doutorado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela bolsa de estudos concedida e aos órgãos financeiros que apoiaram este projeto.

Aos médicos Sidney Raimundo Silva Chalub, Marcus Vinicius Motta Valadão da Silva, Antonio Felipe Santa Maria Coquillard Ayres, Rodrigo Panno Basílio de Oliveira, a aluna de doutorado Elzalina Soares, e as equipes de residência dos centros cirúrgicos e da patologia dos Hospitais Fundação Centro de Controle em Oncologia do Estado do Amazonas (FCECON) e Hospital Universitário Gaffrée e Guinle (HUGG), pela compreensão, auxílio na coleta e análise histopatológica das amostras utilizadas nesse trabalho.

Aos pacientes da FCECON e do HUGG que aceitaram voluntariamente em participar desse trabalho.

Aos professores membros da banca desta tese.

Aos antigos e novos integrantes da Unidade Proteômica, em especial ao Dr. Fabio Nogueira, Dr. Magno Junqueira, Dra. Livia Goto, Dr. Aniel Sanchez, Gabriel Duarte, Rafael Melani, Erika Velasquez, Jimmy Esneider, Nina Daddario, Livia Zamagna e Isabel de Queiroz que me auxiliaram nos experimentos proteômicos, na discussão do experimento e dos dados obtidos, pelo aprendizado mútuo assim como também agradeço pela amizade, carinho e apoio nos momentos mais difíceis.

As minhas amigas Kênia Costa, Bárbara Campos e Roberta Guimarães (*in memoriam*) pela oportunidade de dividir diversas experiências ao morar com vocês, pelos momentos de descontração, confiança, amizade, amor, compreensão, ajuda, e apoio nas situações mais complicadas e inesperadas.

À Dra. Juliana de Saldanha da Gama Fischer pelo apoio, amizade e a troca de conhecimento científico em proteômica.

À Dra. Maria da Glória Costa Carvalho pelo apoio e ensinamentos em biologia molecular.

Ao MSc. Diogo Borges Lima pela amizade, apoio, compreensão, paciência e auxílio durante os anos de doutorado e na análise de dados dessa tese.

Ao Arthur Pezzin Ton pela amizade, compreensão, carinho, paciência e apoio nos momentos que mais precisei.

As minhas amigas de Manaus, Fabiana Almeida, Aimêe Oliveira, Luciana Freire, Adriana Silva e Mayane Souza pelo apoio, compreensão e amizade.

À Fundação Oswaldo Cruz, em especial ao Instituto Leônidas e Maria Deane (ILMD) pelo apoio.

Por fim, a todos que de alguma forma contribuíram para a realização deste trabalho.

"Talvez não tenha conseguido fazer o melhor, mas lutei para que o melhor fosse feito"

Martin Luther King

## RESUMO

AQUINO, Priscila Ferreira de. Avaliação Proteômica da Margem de Ressecção de Pacientes com Câncer Gástrico por Espectrometria de Massas. Tese (Doutorado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

O câncer gástrico é um tipo particular de câncer, que apresenta uma alta taxa de malignidade e mortalidade no mundo inteiro, afetando pessoas de todas as idades, especialmente acima dos 50 anos. O diagnóstico é geralmente obtido nos estágios mais avançados, devido aos sintomas não específicos em seus estágios iniciais, diminuindo drasticamente as opções de tratamento e sobrevida. Um problema comum com o câncer é a recorrência da doença, que para o tipo gástrico é acima de 50% dos casos. Neste contexto, a margem de ressecção é uma região de grande interesse para a compreensão da biologia da doença. Com tais motivações, o presente trabalho objetiva investigar e caracterizar possíveis modificações nos perfis proteicos de amostras de tecidos tumorais e de diferentes margens de ressecção em pacientes diagnosticados com câncer gástrico. Uma análise proteômica foi realizada e dividida em duas etapas: a primeira, um fragmento de tecido de um paciente compreendendo toda a região anatômica do estômago foi seccionada em dez partes e individualmente submetidas à análise por *MudPIT*. Nosso grupo criou um *software*, chamado de *Shotgun Imaging Analyzer (SIA)*, que nos permitiu investigar diferentes características proteicas desse tecido. Nesta abordagem, foi possível obter uma visualização do cenário proteômico das regiões estudadas correlacionada com identificação e quantificação de 1.936 proteínas, onde foi possível demonstrar a expressão de proteínas sabidamente correlacionadas com tumores; adicionalmente, nossos dados permitiram discriminar a região tumoral da margem de ressecção. Já na segunda etapa, quatro fragmentos de tecidos (tumor e margens de ressecção a diferentes distâncias) de dois pacientes foram submetidas à análise histopatológica e proteômica empregando múltiplas separações cromatográficas e quantificação por marcação com isóbaros (iTRAQ). Na análise histopatológica, verificou-se que as margens de ressecção mais distantes do tumor (3 e 5 cm) estavam livres de neoplasia. Já em nossa análise proteômica foram identificadas 2.468 proteínas entre os tecidos estudados. Adicionalmente, nossos dados quantitativos apontaram que apesar das margens de ressecção apresentarem ausência de câncer pela avaliação histológica, a nível molecular, essas já se encontravam modificadas. Exemplos de proteínas que corroboram com essa afirmação, e apresentam aumento na abundância, são a

lumican, a glutathione S-transferase e a talina-1. Em conclusão, apresentamos uma análise proteômica detalhada da região tumoral e das diferentes margens de ressecção, disponibilizando informações sobre a abundância de proteínas que podem auxiliar na maior compreensão dos processos ligados à presença da doença nos tecidos adjacentes ao tumor.

Palavras-chave: Proteômica, câncer gástrico, margem de ressecção, espectrometria de massas.

**ABSTRACT**

AQUINO, Priscila Ferreira de. Avaliação proteômica da Margem de Ressecção de Pacientes com câncer Gástrico por Espectrometria de Massas. Tese (Doutorado em Bioquímica) – Instituto de Química, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2015.

Gastric cancer is a particular type of cancer that presents a high rate of malignancy and mortality worldwide; it affects people of all ages, especially those above 50. It is usually diagnosed in an advanced stage due to symptoms that are not specific to early stages, dramatically decreasing the options of treatment and survival rate. A common problem for gastric cancer is having an incidence of recurrence greater than 50%. The resection margin poses as a region of interest for understanding this disease's biology. This study investigates and characterizes molecular changes in proteomic profiles of tissue samples from tumor and their corresponding resection margins from patients diagnosed with gastric cancer. A proteomic analysis was performed in two stages; the first comprised sectioning a patient's gastric cancer biopsy into ten parts and subjecting each one to MudPIT analysis. For this, our group developed a software, named Shotgun Imaging Analyzer (SIA), which allowed creating a quantitative proteomic landscape of this biopsy. This analysis disclosed 1,936 proteins and allowed us to list a panel of abundance alterations of which many proteins have been previously described to correlate with cancer; thus, our mass spectrometry data allowed us to delineate the tumor from its resection margin region. The second stage consisted in analyzing four biopsies (tumor area, resection margin of 1, 3 and 5 cm) from two different patients by histopathology and a multiplexing proteomic approach (iTRAQ-8 plex). The histopathological analysis showed that the further margins (i.e., 3 & 5 cm) presented no apparent morphological evidence of tumor. Yet, our quantitative proteomic analysis identified 2,468 proteins, many supporting the hypothesis that regions in the resection margin were already compromised at a molecular level. Examples of proteins corroborating with such are: lumican, glutathione S-transferase, and talin-1. These observations suggest that the molecular modifications related to oncogenesis are occurring even in the regions not detected by the histopathological analysis. In conclusion, we provide an in-depth proteomic analysis of the gastric cancer together with its corresponding resection margins, disclosing building blocks for future studies aiming on constructing diagnostic and prognostic panels, besides shedding light on the underlying biological processes that are occurring due to the presence this disease.

Keywords: Proteomics, gastric cancer, resection margin, mass spectrometry.

## SUMÁRIO

RESUMO .....	IX
ABSTRACT .....	XI
ABREVIACÕES, SÍMBOLOS E UNIDADES .....	XIV
LISTA DE FIGURAS.....	XVIII
LISTA DE TABELAS.....	XXII
1. INTRODUÇÃO .....	23
1.1 Câncer gástrico .....	23
1.1.1 Incidência e epidemiologia.....	24
1.1.2 Classificação e tipos histológicos do câncer gástrico .....	26
1.1.3 Fatores de risco .....	27
1.2 Margem de ressecção .....	28
1.3 Proteômica.....	30
1.3.1 Espectrometria de massas.....	33
1.3.2 Proteômica quantitativa.....	34
1.4 Projeto Proteoma Humano (Chromosome-Centric Human Proteome Project, C-HPP e Biology/Disease-driven Human Proteome Project, B/D-HPP).....	37
2. OBJETIVOS.....	39
2.1 GERAL.....	39
2.2 ESPECÍFICOS.....	39
3. MATERIAL E MÉTODOS.....	40
3.1 Seleção de pacientes.....	40
3.2 Caracterização dos tumores .....	40
3.3 Coleta e preservação das amostras.....	41
3.4 Etapa A – Análise do perfil proteico de tecidos das margens de ressecção e do tecido canceroso com a ferramenta computacional <i>Shotgun Imaging Analyzer</i> utilizando uma quantificação <i>label-free</i> .....	43

3.4.1 Pulverização do tecido, extração e dosagem de proteínas.....	43
3.4.2 Digestão trípica em solução.....	44
3.4.3 Dessalinização e concentração das amostras.....	44
3.4.4 Cromatografia de troca catiônica forte.....	44
3.4.5 Dessalinização, concentração e quantificação das amostras .....	45
3.4.6 Aquisição de dados de LC- MS/MS.....	45
3.4.7 Análise de dados .....	46
3.5 Etapa B - Análise histopatológica e proteômica, empregando a quantificação por marcação com isóbaros – iTRAQ, de tecidos coletados da mesma região (tumor e margem de ressecção). .....	49
3.5.1 Análise histopatológica .....	49
3.5.2 Análise proteômica.....	50
3.5.3 Análise dos dados proteômicos.....	54
4. RESULTADOS E DISCUSSÃO.....	57
4.1 Etapa A – Análise do perfil proteico de tecidos das margens de ressecção e do tecido canceroso com a ferramenta computacional <i>Shotgun Imaging Analyzer</i> utilizando uma quantificação <i>label-free</i> . .....	58
4.2 Etapa B - Análise histopatológica e proteômica, empregando a quantificação por marcação com isóbaros – iTRAQ, de tecidos coletados da mesma região (tumor e margem de ressecção). .....	67
4.2.1 Análise histopatológica .....	67
4.2.2 Análise proteômica.....	69
5. CONCLUSÃO.....	83
6. COLABORAÇÕES E/ OU PRODUÇÃO CIENTÍFICA.....	85
7. PERSPECTIVAS.....	87
REFERÊNCIAS BIBLIOGRÁFICAS.....	88
ANEXOS.....	100

**ABREVIACÕES, SÍMBOLOS E UNIDADES**

1-D	Unidimensional
2-D	Bidimensional
AC	Corrente alternada, do inglês, <i>Alternating Current</i>
ACN	Acetonitrila
AGC	Controle de ganho automático, do inglês, <i>Automatic Gain Control</i>
AJCC	<i>American Joint Committee on Cancer</i>
AQUA	Quantificação absoluta, do inglês, <i>Absolute quantification</i>
B/D-HPP	Projeto Proteoma Humano dirigido à Biologia/Doença, do inglês, <i>Biology/Disease-driven Human Proteome Project</i>
CA	Antígeno carboidrato
CAAE	Certificado de Apresentação para Apreciação Ética
CEA	Antígeno carcinoembrionário
CECON	Centro de Controle em Oncologia
CEP	Comitê de Ética em Pesquisa
C-HPP	Projeto proteoma humano centrado no cromossomo, do inglês, <i>Chromosome-Centric Human Proteome Project</i>
CID	Dissociação induzida por colisão, do inglês, <i>Collision Induced Dissociation</i>
CSCs	Células tronco-tumorais, do inglês, <i>Cancer Stem Cells</i>
DDA	Aquisição dependente de dados, do inglês, <i>Data Dependent Acquisition</i>
DESI	Dessorção de ionização por electrospray, do inglês, <i>Desorption Electrospray Ionization</i>
DESI-MS	Dessorção de ionização por electrospray - Espectrometria de massas, do inglês, <i>Desorption Electrospray Ionization- Mass Spectrometry</i>
DNA	Ácido desoxirribonucleico, do inglês, <i>Deoxyribonucleic Acid</i>
dNIAF	Distribuição normalizada do fator de abundância do íon, do inglês, <i>distributed Normalized Ion Abundance Factor</i>
dNSAF	Distribuição normalizada de fatores de abundância espectrais, do inglês, <i>distributed Normalized Spectral Abundance Factor</i>
DTT	Ditiotreitól
EBV	Vírus Epstein-Barr, do inglês, <i>Epstein Barr virus</i>
ESI	Ionização por electrospray, do inglês, <i>Electrospray Ionization</i>

ETD	Dissociação por Transferência de Elétron; do inglês, <i>Electron Transfer Dissociation</i>
FDR	Taxa de Falsos Positivos, do inglês, <i>False Discovery Rate</i>
FFPE	Fixados em formalina e emblocados em parafina, do inglês, <i>Formalin-Fixed, Paraffin-Embedded</i>
GKN1	<i>Gastrokine-1</i>
GKN2	<i>Gastrokine-2</i>
GUI	Interface gráfica do usuário
HCD	Dissociação por alta energia, do inglês, <i>High Energy Collision Dissociation</i>
HE	Hematoxilina e Eosina
HIC	Cromatografia de interação hidrofóbica, do inglês, <i>Hydrophobic Interaction Chromatography</i>
HILIC	Cromatografia líquida de interação hidrofílica, do inglês, <i>Hydrophilic Interaction Liquid Chromatography</i>
HIV	Vírus da Imunodeficiência Humana, do inglês, <i>Human Immunodeficiency Virus</i>
HPLC	Cromatografia líquida de alta eficiência, do inglês, <i>High Performance Liquid Chromatography</i> .
HPP	Projeto proteoma humano, do inglês, <i>Human Proteome Project</i>
HUCFF/UFRJ	Hospital Universitário Clementino Fraga Filho/Universidade Federal do Rio de Janeiro
HUGG	Hospital Universitário Gaffrée e Guinle
HUPO	Organização do Proteoma Humano, do inglês, <i>Human Proteome Organization</i>
IAA	Iodoacetamida
IEX	Cromatografia de troca iônica, do inglês, <i>Ion Exchange Chromatography</i>
iTRAQ	Marcadores isóbaros para a quantificação relativa e absoluta, do inglês, <i>Isobaric Tags for Relative and Absolute Quantification</i>
KCl	Cloreto de potássio
KH <sub>2</sub> PO <sub>4</sub>	Dihidrogenofosfato de potássio
LC-MS/MS	Cromatografia líquida-espectrometria de massas sequencial, do inglês, <i>Liquid Chromatography Tandem-Mass Spectrometry</i>
Ln	Logaritmo neperiano

LTQ	<i>Linear Trap Quadrupole</i>
<i>m/z</i>	Razão massa/carga
MALDI	Ionização por dessorção a laser auxiliada por matriz, do inglês, <i>Matrix-Assisted Laser Desorption/Ionization</i>
MR	Margem de ressecção
MRM	Monitoramento de reações múltiplas, do inglês, <i>Multiple Reaction Monitoring</i>
MS	Espectrometria de massas, do inglês, <i>Mass Spectrometry</i>
MS/MS	Espectrometria de massas sequencial, do inglês, <i>Tandem Mass Spectrometry</i>
MS1	Perfil do espectro de massas
MS2	Varredura de íons produtos
MudPIT	Tecnologia de identificação de proteínas multidimensional, do inglês, <i>Multidimensional Protein Identification Technology</i>
nLC	Nano-Cromatografia líquida, do inglês, <i>nano-Liquid Chromatography</i>
PL	<i>PatternLab for proteomics</i>
ppm	Parte por milhão
p/v	Razão peso/volume
PSM	<i>Peptide Spectrum Matching</i>
RF	Radiofrequência
RIC	Cromatograma reconstruído de íons, do inglês, <i>Reconstructed Ion Chromatogram</i>
ROI	Região de interesse, do inglês, <i>Region of Interest</i>
RP	Fase reversa, do inglês, <i>Reverse Phase</i>
PCR	Reação em cadeia pela polimerase, do inglês, <i>Polymerase Chain Reaction</i>
SCX	Troca catiônica forte, do inglês, <i>Strong Cation Exchange</i>
SEC	Cromatografia de exclusão molecular, do inglês, <i>Size Exclusion Chromatography</i>
SEPro	<i>Search engine processor</i>
SIA	<i>Shotgun-Imaging Analyzer</i>
SILAC	Marcação de isótopos estáveis em aminoácidos em cultura de células, do inglês, <i>Stable Isotope Labelling with Amino acids of Cells in culture</i>

SRM	Monitoramento de reação selecionada, do inglês, <i>Selected-Reaction Monitoring</i>
STRING	<i>Search Tool for the Retrieval of Interacting Genes/Proteins</i>
TCEP	Cloridrato de tris (carboxietil) fosfina, do inglês, <i>tris (2-carboxyethyl) phosphine hydrochloride</i>
TEAB	Bicarbonato de trietilamônio, do inglês, <i>triethylammonium bicarbonate</i>
TFA	Ácido trifluoroacético
TFF1	<i>Trefoil factor 1</i>
TFF2	<i>Trefoil factor 2</i>
TMT	Marcadores de massa sequencial, do inglês, <i>Tandem Mass Tags</i>
TNM	Classificação de tumores malignos Tumor, Nódulo e Metástase
TOF	Tempo de voo, do inglês, <i>Time-of-Flight</i>
UFAM	Universidade Federal do Amazonas
UFRJ	Universidade Federal do Rio de Janeiro
v/v	Razão volume/volume
WHO	Organização mundial de saúde, do inglês, <i>World Health Organization</i>
x g	Veze gravidade, do inglês, <i>Times gravity</i>
XIC	Cromatograma de íon extraído, do inglês, <i>Extracted-Ion Chromatogram</i>

## LISTA DE FIGURAS

Figura 1- Distribuição mundial da estimativa sobre a incidência de casos (x1000) de câncer gástrico prevalente em cinco anos na população adulta, de acordo com a Organização Mundial de Saúde. Fonte: modificado de "International Agency for Research on Cancer".....	24
Figura 2- Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para o biênio 2014-2015 por sexo, exceto câncer de pele não melanoma. Fonte: "Instituto Nacional de Câncer José Alencar Gomes da Silva".....	25
Figura 3 - Corte histopatológico corado com hematoxilina e eosina de dois tipos sub-histológicos de câncer gástrico A) Tipo difuso de adenocarcinoma enfatizando a presença de células em anéis de sinete (aumento da objetiva 200X); B) Tipo intestinal de adenocarcinoma exibindo a formação de glândulas (aumento da objetiva 100X). Fonte: modificado de YAKIREVICH; RESNICK, 2013. ....	27
Figura 4 - Visualização da margem de ressecção a partir da delimitação da região tumoral. Fonte: OPPENHEIMER et al., 2010.....	29
Figura 5 - Esquema da tecnologia de identificação de proteínas multidimensional (MudPIT), a qual utiliza uma cromatografia de nanofluxo em uma coluna bifásica com resina de troca catiônica forte (SCX) seguido de fase reversa (RP) acoplada a espectrometria de massas em tandem. Fonte: Modificado de FISCHER, 2010. ....	32
Figura 6 - A) Imagem ilustrativa de um espectrômetro de massas <i>Linear trap</i> quadrupolo (LTQ) Orbitrap Velos® e B) Esquema desse instrumento. Fonte: modificado de “ <a href="http://www.thermoscientific.com">http://www.thermoscientific.com</a> ”; OLSEN et al., 2009. ....	34
Figura 7 – Representação de um cromatograma de íons extraídos (XIC). Fonte: modificado de ONG; MANN, 2005.....	35
Figura 8 - A) Estrutura química dos componentes do iTRAQ; B) Diagrama mostrando a localização dos íons marcadores do iTRAQ 8-plex no espectro de MS/MS. Fonte: modificado de “ <a href="http://tools.thermofisher.com/content/sfs/manuals/cms_049786.pdf">http://tools.thermofisher.com/content/sfs/manuals/cms_049786.pdf</a> ”; “National Institute on Aging - Proteomics”.....	37
Figura 9 – Fluxograma geral das análises realizadas nas amostras coletadas neste trabalho. .	42
Figura 10 - A) Imagem obtida através da ressecção cirúrgica do estômago completo de um paciente; B) Tecido desse estômago dividida verticalmente em 10 seções.....	43
Figura 11 - Interface gráfica do usuário (GUI) exibindo a imagem de um tecido com câncer gástrico analisado no trabalho. ....	48

Figura 12 - A) Imagem do estômago de um paciente compreendendo os quatro fragmentos estudados (tumor, margem de ressecção - MR 1, 3 e 5 cm); B) Seção do estômago dividida horizontalmente em duas partes, uma para análise histopatológica e outra para a análise proteômica. ....	49
Figura 13 - Esquema da marcação com isóbaros (iTRAQ – 8 <i>plex</i> ) nas amostras estudadas..	51
Figura 14 – Imagem da solução de hidróxido de amônio 5% posicionada na fonte de ionização entre a saída do capilar e a entrada no espectrômetro de massas. ....	54
Figura 15 - Espectro de massas sequencial do peptídeo único <i>K.DVDWFLLGSPIEK.L</i> encontrado na sequência da proteína <i>gastrokine-2</i> . ....	60
Figura 16 – Resultado do <i>Shotgun Imaging Analyzer</i> para a proteína NX_Q86XP6, isoforma 2 da <i>gastrokine</i> . A) Mapa de calor baseado na quantificação. B) O eixo y representa a distribuição normalizada de fatores de abundância espectrais (dNSAF) e o eixo x está alinhado com a imagem fornecida no <i>Shotgun Imaging</i> e os números correspondentes de cada seção do tecido. Um regressor foi usado para traçar o gráfico da área em azul. Seções 1-4 representa regiões tumorais, seções 5-7 foram classificadas como intermediárias e as seções 8-10 como margem de ressecção. ....	60
Figura 17 – Rede de interações da proteína <i>gastrokine-2</i> (GKN2) através do programa <i>STRING 10.0</i> . As linhas roxas representam confiança de evidência na interação, sendo que quanto mais intenso é a cor e a linha mais alta é a confiança ( <i>score</i> > 0,9). Fonte: “ <a href="http://string-db.org/">http://string-db.org/</a> ”. ....	62
Figura 18 - Resultado do <i>Shotgun Imaging Analyzer</i> para a proteína NX_P21266, isoforma 1 da glutatona S-transferase Mu 3. A) O mapa de calor baseado na quantificação pelo cromatograma extraído de íons e B) A representação gráfica da abundância dessa proteína através das seções do tecido. ....	63
Figura 19 - Resultado do <i>Shotgun Imaging Analyzer</i> da proteína NX_P04083, isoforma 1 da anexina A1. A) O mapa de calor e B) Representação gráfica da abundância dessa proteína através das seções do tecido. ....	64
Figura 20 - Resultado do <i>Shotgun Imaging Analyzer</i> da proteína NX_P12830, isoforma 1 da E-caderina. A) O mapa de calor e B) Representação gráfica da abundância dessa proteína através das seções do tecido. ....	65
Figura 21 - Resultado do <i>Shotgun Imaging Analyzer</i> para a proteína NX_P67936, isoforma 1 da tropomiosina cadeia alpha-4. A) O mapa de calor, B) Representação gráfica e C) Percentagem da abundância dessa proteína nas seções do tecido. ....	66

Figura 22 - Corte histopatológico do fragmento tumoral do paciente dois corado com hematoxilina e eosina, onde é possível observar a presença de células em anéis de sinete indicados na figura.....	68
Figura 23 - Corte histopatológico do fragmento da margem de ressecção a 1 cm do paciente dois corado com hematoxilina e eosina, onde verificou-se adenocarcinoma pouco diferenciado. ....	68
Figura 24 - Corte histopatológico do fragmento da margem de ressecção a 3 cm do paciente três corado com hematoxilina e eosina, livre de neoplasia. ....	68
Figura 25 - Corte histopatológico do fragmento da margem de ressecção a 5 cm do paciente três corado com hematoxilina e eosina, livre de neoplasia.....	69
Figura 26 – Perfil cromatográfico de uma das amostras analisadas (HILIC_01) no estudo....	72
Figura 27 – Um exemplo da região de baixa massa de um espectro de MS2 obtido neste estudo, enfatizando a presença dos íons marcadores característicos do iTRAQ-8 <i>plex</i> . ....	72
Figura 28 – Gráfico da quantidade de espectros identificados na triplicata técnica das treze frações cromatográficas. ....	73
Figura 29 – Representação gráfica da quantidade de espectro <i>multiplex</i> obtido em todas as amostras estudadas.....	73
Figura 30 – Representação gráfica da normalização realizada pelo módulo de quantificação para o iTRAQ do <i>PatternLab (Isobaric Analyzer)</i> . A) Visualização dos sinais totais dos MS2 identificados nos oito canais; B) Visualização dos sinais totais depois da normalização nos oito canais; C) Visualização dos sinais totais do arquivo <i>RAW</i> . ....	74
Figura 31 – Representação gráfica da abundância do peptídeo <i>R.IINEPTAAAIAYGLDKR.E</i> encontrado nas réplicas da amostra HILIC_01, com 5 espectros entre os diferentes canais de marcação isóbara estudada. O sinal dos marcadores, decorrente de cada espectro, estão categorizados por cores.....	75
Figura 32 - Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 1 cm. Nessa, cada ponto representa um peptídeo que foi mapeado como um ponto no gráfico de acordo com $-\log(p\text{-valor})$ (eixo das abscissas) e $\ln(\text{Fold Change})$ (eixo das coordenadas). Os pontos em vermelho representam peptídeos que estão mais abundantes na margem de ressecção e que passaram no critério do $p\text{-valor}$ e do $\text{fold change}$ ; os pontos em verde demonstram os peptídeos que estão mais abundantes na região tumoral e que também passaram nos critérios estatísticos. Já os pontos em cinza significam peptídeos que não passaram pelos filtros estatísticos.....	77

Figura 33 – Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 5 cm. ....	78
Figura 34 - Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 3 cm. ....	78
Figura 35 – Representação gráfica da análise da tendência de proteínas pelo <i>TrendQuest</i> entre a região tumoral e as três margens de ressecção estudadas. Nessa, cada linha representa uma proteína.....	79
Figura 36 - Representação gráfica da tendência de proteínas relativamente uniforme nas quatro regiões estudadas pelo <i>TrendQuest</i> .....	80
Figura 37 - Visualização da cobertura de peptídeos de proteínas <i>gastrokine -1</i> (A) e desmina (B). Nestas, as barras em vermelho indicam peptídeos que tiveram decréscimos em sua abundância; as barras verdes sinalizam peptídeos que tiveram um aumento na sua abundância e as barras cinzas representam peptídeos que não passaram nos critérios estatísticos. ....	81

## LISTA DE TABELAS

Tabela 1 – Características clínicas dos pacientes estudados. ....	57
Tabela 2 – Características histopatológicas dos pacientes estudados. ....	57
Tabela 3 – Leitura da concentração de proteínas pelo ensaio fluorimétrico – <i>Qubit 2.0</i> ® ( <i>Invitrogen</i> ). ....	58
Tabela 4 – Leitura da concentração dos peptídeos presente nas frações resultante dos passos salinos realizados. ....	59
Tabela 5 – Leitura da concentração de proteínas pelo ensaio fluorimétrico – <i>Qubit 2.0</i> ® ( <i>Invitrogen</i> ). ....	70
Tabela 6 – Leitura da concentração dos peptídeos antes da marcação pelo ensaio fluorimétrico – <i>Qubit 2.0</i> ® ( <i>Invitrogen</i> ). ....	70
Tabela 7 – Leitura da concentração dos peptídeos presente nas frações resultante da cromatografia líquida de interação hidrofílica (HILIC). ....	71
Tabela 8 – Leitura da concentração dos peptídeos presente nas frações resultante da cromatografia líquida de troca catiônica forte (SCX). ....	71
Tabela 9 – Resumo dos resultados obtidos da quantificação dos peptídeos para o iTRAQ do <i>PatternLab (Isobaric Analyzer)</i> nas três abordagens empregadas. ....	75
Tabela 10 – Dados referentes a quantificação do peptídeo <i>R.IINEPTAAAIAYGLDKR.E</i> encontrado mais de uma vez nas réplicas da amostra HILIC_01 durante a eluição cromatográfica, os quais foram usados para calcular o valor do <i>ln fold change</i> (-1.670). ....	76
Tabela 11 – Dados referentes aos valores de <i>ln fold change</i> obtidos na quantificação do peptídeo <i>R.IINEPTAAAIAYGLDKR.E</i> encontrado mais de uma vez nas réplicas da amostra HILIC_01 durante a eluição cromatográfica, que foram usados para calcular o <i>p-valor</i> (0,001). ....	76

## 1.INTRODUÇÃO

A associação da química com a biologia e as técnicas ômicas (*e.g.* genômica, metabolômica, proteômica etc.) revolucionaram a forma com que podemos vislumbrar, ao nível molecular, patologias como, por exemplo, o câncer (JIMENO; HIDALGO, 2006). Estas tecnologias nos permitem identificar, caracterizar e quantificar biomoléculas diferencialmente abundantes quando comparado diferentes estados patológicos. Neste contexto, a margem de ressecção, um tecido morfológicamente sadio que é retirado durante a remoção cirúrgica do câncer, torna-se uma região promissora para ser explorada por técnicas ômicas. A margem de ressecção localiza-se ao redor do tumor e estudos apontam possíveis alterações moleculares nessa região que podem estar diretamente ligadas à recorrência da doença (OPPENHEIMER et al., 2010). Portanto, nosso grupo postula que o estudo proteômico da margem de ressecção a diferentes distâncias do tumor pode contribuir para um melhor entendimento desta patologia.

O câncer gástrico é um problema de saúde pública por resultar em elevada taxa de mortalidade ao redor do mundo; o seu diagnóstico é geralmente realizado apenas nos estágios mais avançados (LIN et al., 2012). Em trabalho anterior, nosso grupo comparou biópsias de câncer gástrico com margens de ressecção e tecidos saudáveis, e mostramos que as margens de ressecção apresentavam um perfil proteômico mais semelhante a biópsias de câncer do que as de tecidos saudáveis. Estes resultados aumentaram o nosso interesse em estudar a margem de ressecção por possivelmente desempenhar algum papel importante para as células cancerosas; servindo como um microambiente de nutrição para o crescimento e propagação de tais células (AQUINO et al., 2012). O presente trabalho é uma continuação deste estudo; aqui, nos aprofundamos em caracterizar alterações moleculares vinculadas ao processo de tumorigênese.

### 1.1 Câncer gástrico

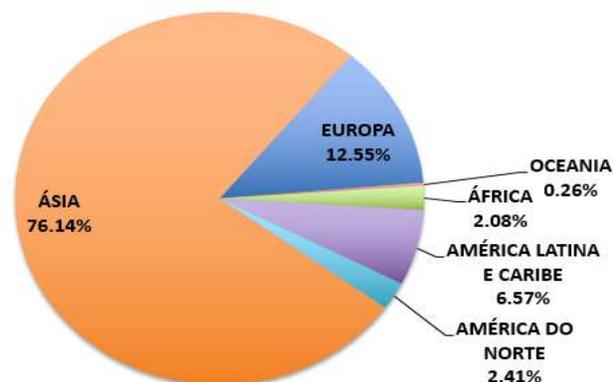
O câncer é uma doença multifatorial, envolvendo diversas alterações genéticas e epigenéticas que englobam fatores externos como o meio ambiente e os hábitos ou costumes de uma população. A doença é caracterizada pelo acúmulo anormal de células, que se propagam e dividem sem controle, invadindo e destruindo tecidos adjacentes ou/e tecidos e órgãos mais distantes, através de um processo denominado metástase (GARTNER, 2007; NUSSBAUM, 2008). Entre os diversos tipos, o câncer gástrico é o quinto mais comum e o terceiro em mortes relacionados à essa doença no mundo inteiro (FERLAY et al., 2015). Este geralmente produz sintomas não específicos em seus estágios iniciais, como a dispepsia, fazendo com que seja

diagnosticado tardiamente, onde entre 80% e 90% dos pacientes apresentam os tumores localmente em estágios avançados ou metastáticos, por conseguinte, uma baixa taxa de ressecabilidade do tumor e tempo de sobrevida ao paciente (YAKIREVICH; RESNICK, 2013).

A endoscopia é o principal método de diagnóstico do câncer gástrico, porém pode-se observar também na clínica médica o uso de biomarcadores para tumores gastrointestinais como: o CEA (antígeno carcinoembrionário), o CA 19-9 (antígeno carboidrato 19-9) e o CA 72-4 (antígeno carboidrato 72-4). Contudo, a especificidade destes é menor que 40% em pacientes em estado avançado e menor que 20% em pacientes em estágios iniciais. Além disso, cada um desses biomarcadores podem apresentar níveis de expressão distintos quando diferentes estágios do câncer gástrico são investigados (LIU et al., 2014; UCAR et al., 2008).

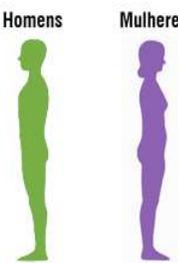
### 1.1.1 Incidência e epidemiologia

A incidência do câncer gástrico varia substancialmente nas diferentes regiões geográficas, refletindo uma distribuição bem heterogênea entre os países, com altas taxas na Ásia, na Europa, América Latina e Caribe, de acordo com a última estimativa da Organização Mundial de Saúde (do inglês *World Health Organization* – WHO) (Figura 1). Neste contexto, o Japão é o país que apresenta a maior taxa de incidência deste tipo de câncer, sendo o risco em homens entre 1.8 e 2.0 vezes maior do que em mulheres (LEUNG et al., 2008; YANG, 2006). Mais de 70% dos casos ocorrem em países em desenvolvimento e, apesar de um pequeno declínio mundialmente, este tipo de câncer representa 7% do total de mortes quando comparado aos demais tipos (FERLAY et al., 2015).



**Figura 1** - Distribuição mundial da estimativa sobre a incidência de casos (x1000) de câncer gástrico prevalente em cinco anos na população adulta, de acordo com a Organização Mundial de Saúde. Fonte: modificado de "International Agency for Research on Cancer".

No Brasil, o câncer de estômago em homens é o quarto mais frequente e o sexto para as mulheres (Figura 2), desconsiderando-se os casos de câncer de pele não melanoma. Ao fazer uma análise mais detalhada entre as regiões brasileiras, observa-se que, para os homens, é o segundo mais frequente nas regiões Norte e Nordeste. Nas regiões Centro-Oeste e Sul, o câncer de estômago é o quarto e na região Sudeste, ocupa a quinta posição. Com relação às mulheres, é o terceiro mais frequente na Região Norte. Nas regiões Sudeste e Nordeste, ocupa a quinta posição. Já nas regiões Sul e Centro-Oeste, é o sexto tipo de câncer mais frequente (“Estimativa 2014 - Incidência de Câncer no Brasil”). Estima-se que, para o biênio 2014-2015, o número de casos novos de câncer de estômago seja de 12.870 entre homens e de 7.520 nas mulheres. Estes valores correspondem a um risco estimado de 13,19 casos novos a cada 100 mil homens e 7,41 para cada 100 mil mulheres (“Instituto Nacional de Câncer José Alencar Gomes da Silva”).

Localização primária	casos novos	%			Localização primária	casos novos	%
Próstata	68.800	22,8%			Mama Feminina	57.120	20,8%
Traqueia, Brônquio e Pulmão	16.400	5,4%		Cólon e Reto	17.530	6,4%	
Cólon e Reto	15.070	5,0%		Colo do Útero	15.590	5,7%	
Estômago	12.870	4,3%		Traqueia, Brônquio e Pulmão	10.930	4,0%	
Cavidade Oral	11.280	3,7%		Glândula Tireoide	8.050	2,9%	
Esôfago	8.010	2,6%		Estômago	7.520	2,7%	
Laringe	6.870	2,3%		Corpo do Útero	5.900	2,2%	
Bexiga	6.750	2,2%		Ovário	5.680	2,1%	
Leucemias	5.050	1,7%		Linfoma não Hodgkin	4.850	1,8%	
Sistema Nervoso Central	4.960	1,6%		Leucemias	4.320	1,6%	

\*Números arredondados para 10 ou múltiplos de 10.

**Figura 2-** Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para o biênio 2014-2015 por sexo, exceto câncer de pele não melanoma. Fonte: “Instituto Nacional de Câncer José Alencar Gomes da Silva”.

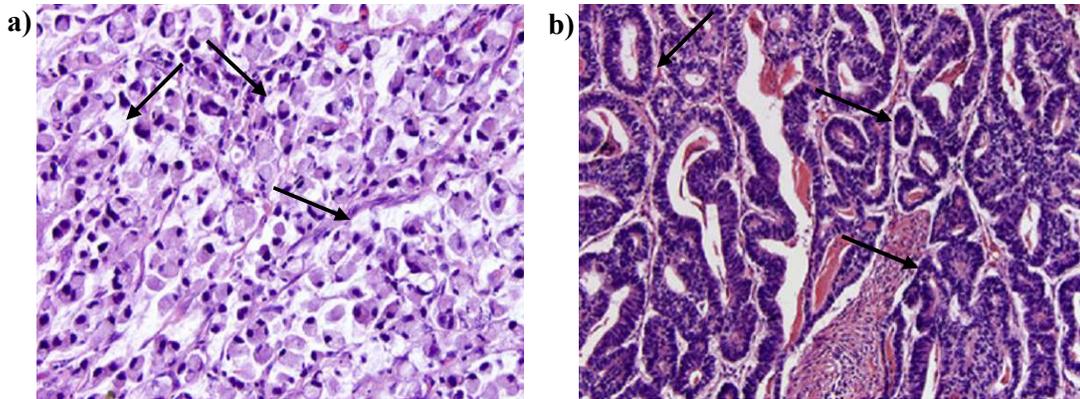
Dados epidemiológicos mostram que câncer gástrico é raro antes dos 40 anos, porém o risco é maior com o aumento de idade, sendo mais frequente acima dos 50 anos de idade (DICKEN et al., 2005; EBERT et al., 2006). A média de sobrevivência é em torno de cinco anos para 10-28% dos casos com tumores ressecados, oito meses para pacientes submetidos a procedimentos paliativos e somente cinco meses para estados muito avançados da doença e que não podem ser submetidos a uma cirurgia. De maneira que o diagnóstico precoce, assim como novas formas de tratamento não invasivos são cruciais para o rastreamento da população de risco deste tipo de câncer e o aumento do tempo de sobrevida (DICKEN et al., 2005; LIU et al., 2014).

### 1.1.2 Classificação e tipos histológicos do câncer gástrico

Os tumores gástricos se apresentam, predominantemente, sob a forma de três tipos histológicos: o adenocarcinoma, responsável por 95% dos tumores gástricos, o linfoma, diagnosticado em cerca de 3% dos casos, e leiomiossarcoma, que se inicia em tecidos que dão origem aos músculos e aos ossos (INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA; SCHULZ, 2005) . Vários sistemas de classificação têm sido propostos para ajudar na descrição do estadiamento do câncer gástrico, seja através das características macroscópicas ou microscópicas. Os dois sistemas mais utilizados são: a classificação Tumor, Nódulo e Metástase (TNM) da *American Joint Committee on Cancer* (AJCC) e a classificação de Lauren.

O sistema TNM descreve a extensão anatômica da doença com base na avaliação de três componentes: o tumor (T) avaliando a extensão do tumor primário; os linfonodos (N) verificando a ausência ou presença e a extensão de metástase em linfonodos regionais; e metástase (M) observando a ausência ou presença de metástase à distância (INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA).

A classificação de Lauren é um sistema utilizado para adenocarcinomas gástricos que separa em dois tipos sub-histológicos: intestinal e difuso. Tumores que possuem características de ambos os tipos são designados como um tipo misto de adenocarcinoma (YAKIREVICH; RESNICK, 2013). Este sistema descreve os tumores de acordo com a configuração microscópica e o padrão de crescimento. O tipo difuso apresenta células tumorais não coesivas difundindo e infiltrando o estroma do estômago; geralmente possuem uma infiltração da parede do estômago com pouca ou nenhuma formação de glândula e possuem células conhecidas como *anel de sinete* (Figura 3a). Este tipo tem demonstrado um aumento na incidência relativa, ocorrendo mais frequentemente em pacientes mais jovens e está associado a um pior prognóstico. Já o tipo intestinal é caracterizado pela formação de glândulas que exibem vários graus de diferenciação e cujo crescimento é expansivo e não infiltrativo como no tipo difuso (Figura 3b) (DICKEN et al., 2005; YAKIREVICH; RESNICK, 2013).



**Figura 3** - Corte histopatológico corado com hematoxilina e eosina de dois tipos sub-histológicos de câncer gástrico **A)** Tipo difuso de adenocarcinoma enfatizando a presença de células em anéis de sinete (aumento da objetiva 200X); **B)** Tipo intestinal de adenocarcinoma exibindo a formação de glândulas (aumento da objetiva 100X). Fonte: modificado de YAKIREVICH; RESNICK, 2013.

### 1.1.3 Fatores de risco

Diversos fatores podem contribuir individualmente ou em conjunto para o surgimento dessa neoplasia, desde fatores genéticos a epigenéticos e até hábitos alimentares. O maior fator de risco conhecido até o momento é a infecção pela bactéria *Helicobacter pylori*. Esta é um dos agentes infecciosos mais comuns em humanos (> 50% da população mundial é infectada); é considerada o principal agente etiológico em mais de 95% das gastrites crônicas, sendo responsável pelo aumento do risco de câncer em até nove vezes e classificada como agente cancerígeno do grupo I, ou seja, cancerígeno para humanos (CÉSAR et al., 2002; YAKIREVICH; RESNICK, 2013). A ação dessa bactéria é de forma indireta, onde a interação entre as células epiteliais da mucosa do estômago e a bactéria ocorre mediante a atuação de várias proteínas bacterianas (*e.g.* proteína VacA) que induzem à vacuolização e morte das células do hospedeiro. Isso leva à liberação de nutrientes facilitando a introdução da *H. pylori* através da barreira mucosa e, como essa proteína também é imunossupressora, diminui as respostas por macrófagos e células-T (SCHULZ, 2005).

Outros fatores que podem ser citados são: o baixo consumo de frutas frescas e vegetais, a ingestão de álcool, o fumo e as nitrosaminas, ( $R_2N-N=O$ ), potentes agentes cancerígenos relacionadas a diversos gêneros alimentícios; estes, ao reagirem com as bases nitrogenadas do DNA podem causar mutações, sendo que algumas dessas mutações já foram encontradas em cânceres gástricos nos genes: TP53, CDH1 (E-caderina), e CTNNB1 ( $\beta$ -catenina) (MACDONALD et al., 2004; SCHULZ, 2005).

Já entre alguns dos principais fatores genéticos tem-se: a pré-disposição genética pelo alto risco de mutações nos genes – CDH1; que é responsável por codificar a E-caderina, uma

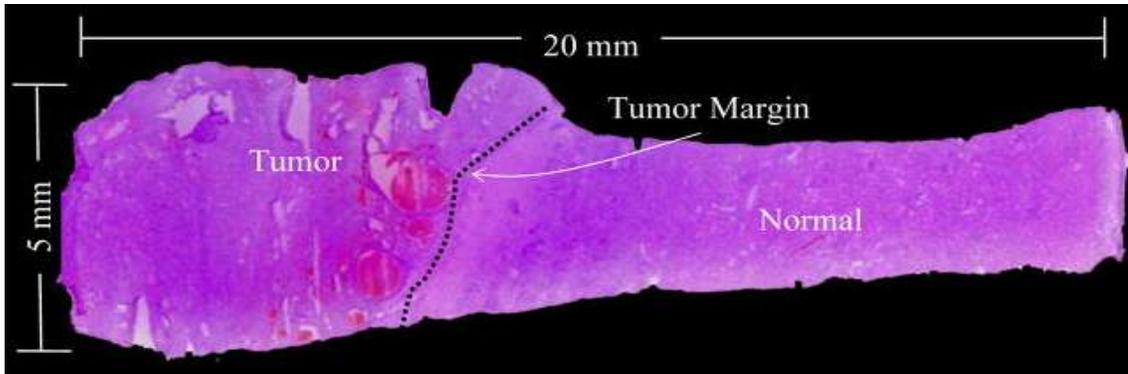
glicoproteína que auxilia na conexão do citoesqueleto através das cateninas, e cuja perda da funcionalidade dessa proteína pode contribuir para uma desregulação da via WNT e acarretar um padrão de crescimento difuso das células tumorais (ALBERTS et al., 2002; SCHULZ, 2005); e os possíveis polimorfismos no gene IL1B, responsável por codificar as citocinas, um grupo de proteínas envolvidas na sinalização de respostas imunes do organismo (ALBERTS et al., 2002; MACDONALD et al., 2004).

Além disso, alterações epigenéticas (*i.e.*, não altera o código de DNA mas pode modificar a expressão gênica) como por exemplo, a metilação de genes promotores que consequentemente causam o silenciamento desses genes, também têm sido mostrados na literatura correlacionadas ao câncer gástrico e a progressão da gastrite crônica (KANG et al., 2015; LIANG et al., 2015; YAKIREVICH; RESNICK, 2013).

## 1.2 Margem de ressecção

A principal forma de tratamento curativo do câncer gástrico é a ressecção cirúrgica, mais comumente a gastrectomia subtotal ou total. A escolha do procedimento depende do tamanho, localização e da capacidade de obter margens de ressecção livres da doença de forma macroscópica e microscópica (WU; CHUNG, 2013; YAKIREVICH; RESNICK, 2013). Em geral, para tumores localizados em mais de um terço do estômago, este é completamente removido; já para um comprometimento menor da mucosa gástrica, a retirada é parcial.

Neste contexto, a margem de ressecção pode ser entendida como a área em torno da região tumoral (Figura 4). A partir da avaliação de um patologista, esta pode ser classificada como negativa ou positiva, ou seja, pode apresentar possíveis células cancerígenas (positiva) ou não (negativa), assim como é possível avaliar, juntamente com tecido proveniente da lesão tumoral, o grau de disseminação da neoplasia (LIANG et al., 2015; SHEN et al., 2006). Tais avaliações auxiliam a equipe médica na necessidade de outros procedimentos adicionais ao tratamento do paciente.



**Figura 4** - Visualização da margem de ressecção a partir da delimitação da região tumoral. Fonte: OPPENHEIMER et al., 2010.

A obtenção de uma margem de ressecção negativa é um importante fator no tratamento do paciente. A distância adequada das margens até o tumor é algo bastante questionado, visto que esta pode variar amplamente, dependendo de vários fatores, o que dificulta a definição e o estabelecimento de padrões para o corte (SHEN et al., 2006). No entanto, há alguns estudos e recomendações para a obtenção das margens proximal e distal. Entre eles tem-se as diretrizes da associação japonesa de câncer gástrico de 2010. Nesta, recomendam-se ressecções proximais de 2 cm para tumores T1, 3 cm para tumores com um padrão de crescimento mais expansivo (T2-T4) ou do tipo histológico intestinal, e 5 cm para tumores T2-T4 ou do tipo histológico difuso (JAPANESE GASTRIC CANCER ASSOCIATION, 2011; KIM et al., 2014; SHIN; PARK, 2013; SQUIRES et al., 2015).

Apesar da distância da margem proximal ser mais variável, segundo Squires III et al. (2015), um valor adequado seria  $>3,0$  cm visto que se tem verificado que margens de ressecção da extremidade proximal de tamanhos mais extensos ( $> 5$ cm) não necessariamente estão associadas a uma melhora significativa nas taxas de sobrevida. Já para as margens distais, estas geralmente têm sido determinadas como, pelo menos, 2 a 4 cm distais do piloro (porção terminal do estômago responsável por controlar a passagem do alimento para o duodeno, durante a digestão) (SHIN; PARK, 2013).

A incidência de 1-20% de margem de ressecção positiva tem sido encontrada nos pacientes; estas estão associadas com a recorrência local da doença e correlacionadas a um pior tempo de sobrevida em cinco anos (18%) comparado às margens negativas (45%) (LIANG et al., 2015; NEOPTOLEMOS et al., 2001).

Em um trabalho com câncer gástrico, a taxa de recorrência foi de 50,7% em aproximadamente 17 a 19 meses após o procedimento cirúrgico (MARRELLI et al., 2005). O estadiamento tumoral é um fator que é levado em consideração ao ser observado a presença de

células malignas nesses tecidos, onde tumores do tipo não diferenciado, que invadem dois terços ou mais do estômago e tumores de estágios mais avançados representam processos biológicos mais agressivos. Além disso, nestes pacientes com margens positivas, uma nova operação cirúrgica para a retirada do tecido residual com estas células deve ser avaliada minuciosamente, sendo a quimioterapia e a radioterapia, as alternativas mais empregadas nesses casos (LIANG et al., 2015).

### 1.3 Proteômica

Os avanços na área da genômica, principalmente com a disponibilização de diversos genomas sequenciados, catalisaram os estudos proteômicos. O termo proteoma foi descrito pela primeira vez em 1994 referindo-se ao completo conteúdo proteico expresso pelo genoma (LANE, 2005; WILKINS et al., 1996; YATES; OSTERMAN, 2007). Progressos na biologia molecular, na química de proteínas e na bioinformática possibilitaram o surgimento da proteômica na década de 90. Esta pode ser entendida como o conjunto de técnicas que são utilizadas para se estudar proteomas e como eles reagem às mudanças com o tempo e às várias condições submetidas (EIDHAMMER et al., 2007). Portanto, a proteômica permite realizar estudos sistemáticos de diversas propriedades de proteínas; como por exemplo: estrutural, das modificações pós-traducionais, função, localização celular, interação proteína-proteína, entre outros; possibilitando assim, uma visão global e integrada de um determinado sistema biológico (HUBER, 2003).

O estudo do proteoma requer metodologias que possuam sensibilidade, reprodutibilidade, rapidez, facilidade e automação, fatores esses que resultam em uma análise de custo relativamente alta. O desenvolvimento de novas tecnologias para a separação de peptídeo/proteína, a análise por espectrometria de massas, a quantificação por marcação com isóbaros e a análise dos dados pela bioinformática tem possibilitado o progresso nessa área (AEBERSOLD; MANN, 2003; YATES et al., 2009; ZHANG et al., 2013). Atualmente, a espectrometria de massas (MS) é a principal ferramenta para a análise de proteínas em larga-escala e tem proporcionado análises em poucas horas, como por exemplo no estudo de Hebert et al. (2014) que identificou em uma hora aproximadamente 4.000 proteínas de levedura.

Uma das estratégias proteômicas mais utilizadas é a *bottom-up*. Nessa, a identificação de proteínas é feita pela análise dos peptídeos derivados da proteólise. Entre as técnicas pertencentes a essa abordagem tem-se a *Shotgun proteomics*, que é a mais adotada na caracterização de misturas complexas de proteínas; e foi criada pelo renomado cientista Dr.

*John R. Yates, III*, cujo nome é uma analogia ao sequenciamento genômico por *shotgun* (ZHANG et al., 2013).

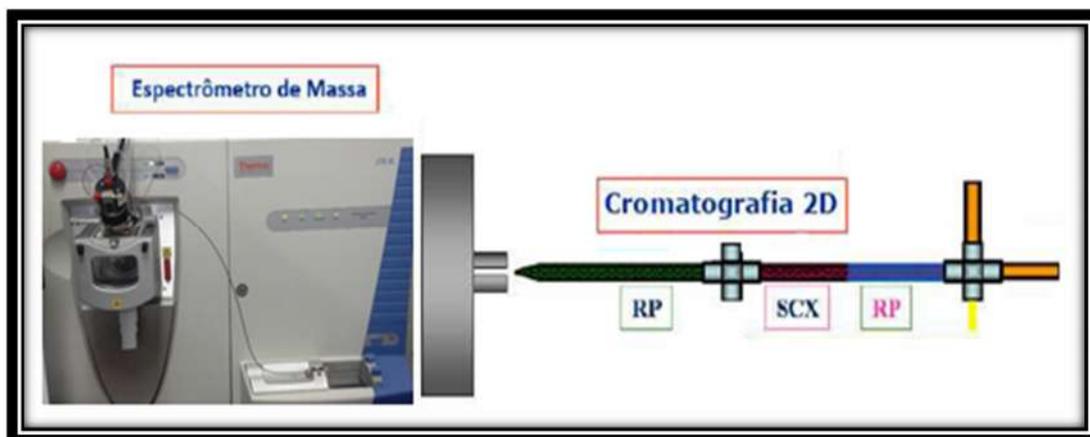
Um típico experimento de proteômica *bottom-up* engloba a identificação e quantificação de milhares de proteínas, proporcionando dados para auxiliar na compreensão de uma dada questão biológica investigada. Existem várias metodologias de preparo de amostras que abrangem desde proteínas purificadas até misturas proteicas complexas oriundas de cultura de células, bactérias, organelas, tecidos, entre outros.

A preparação de amostras é um passo crítico e deve levar em consideração as diferentes características físico-químicas que as proteínas possuem, assim como as possíveis interações com outras biomoléculas. O método de preparo de amostras deve ser reprodutível, com alto rendimento, idealmente minimizando os passos de manipulação, evitar contaminações e deve usar reagentes que sejam compatíveis com a etapa de espectrometria de massas (COTTRELL, 2011; NOGUEIRA, 2012; ZHANG et al., 2013). A digestão proteolítica é realizada, em geral, com enzimas que diferem na especificidade de clivagem entre os aminoácidos. A clivagem ocorre através da hidrólise da ligação amida antes ou depois de resíduos específicos ou a combinação deles. A tripsina é uma das enzimas mais usadas na digestão proteica para experimentos do tipo *shotgun proteomics*, cuja clivagem ocorre após a arginina e lisina (no sentido C-terminal), desde que não estejam acompanhados de prolina. Dependendo da complexidade da amostra, pode-se utilizar também uma combinação de enzimas, o que permite aumentar a cobertura proteica e a sensibilidade do resultado (OLSEN et al., 2004; ZHANG et al., 2013).

Já quanto ao fracionamento e a separação de proteínas e peptídeos, os principais métodos utilizados são: eletroforese uni e bidimensional (1-D e 2-D, respectivamente) e a cromatografia líquida (LC). As técnicas baseadas em gel foram bastante empregadas na década de 90, sendo hoje praticamente substituídas pelas metodologias de LC. A complementaridade entre a cromatografia e a espectrometria de massas tem proporcionado cada vez mais a utilização da abordagem sem a utilização de gel na análise proteômica (NOGUEIRA, 2012; ZHANG et al., 2013).

Os principais métodos cromatográficos usados são: interação hidrofóbica (HIC), hidrofílica (HILIC), troca iônica (IEX), exclusão molecular (SEC) e por afinidade; sendo que a nanocromatografia líquida (nLC) de fase reversa (RP) é o método cromatográfico mais empregado. Atualmente, também existem métodos cromatográficos combinados como, por exemplo, a tecnologia multidimensional de identificação de proteínas (MudPIT), que associa a cromatografia de troca catiônica forte (SCX) intercalada com cromatografia de fase reversa

(RP) diretamente acoplada a espectrometria de massas em tandem (Figura 5) (WASHBURN et al., 2001; ZHANG, X. et al., 2010). Estes métodos podem ser *on-line* (com uma coluna bifásica) ou *off-line* (em colunas separadas; e.g. a cromatografia de troca iônica é feita separadamente para então submeter as frações dessa primeira dimensão à cromatografia de fase reversa). O modo *on-line*, é um processo automatizado diminuindo portanto, as perdas e a contaminação de amostra (MOTOYAMA; YATES, 2008). O MudPIT e outros métodos cromatográficos bidimensionais, proporcionam uma melhoria na separação cromatográfica, o que conseqüentemente resulta em um aumento nas identificações proteicas. Nessa tese, utilizamos métodos cromatográficos bidimensionais, combinando SCX com RP e HILIC com RP.



**Figura 5-** Esquema da tecnologia de identificação de proteínas multidimensional (MudPIT), a qual utiliza uma cromatografia de nanofluxo em uma coluna bifásica com resina de troca catiônica forte (SCX) seguido de fase reversa (RP) acoplada a espectrometria de massas em tandem. Fonte: Modificado de FISCHER, 2010.

A identificação de espectros oriundos de peptídeos pode ser feita mediante ao auxílio de ferramentas de proteômica computacional, onde os espectros com a razão massa/carga dos íons precursores juntamente com o espectro de dissociação destes serão os dados analisados. As três formas canônicas de estratégias para identificação são: sequenciamento *de novo*, *sequence tag search*, e o *peptide spectrum match* (PSM), sendo essa última, utilizada neste trabalho, e considerada o padrão-ouro, por ser a metodologia que apresenta a maior sensibilidade na existência de um banco de sequências proteicas. Alguns exemplos de ferramentas de busca que utilizam o PSM são: o SEQUEST (ENG et al., 1994), o X-Tandem (CRAIG; BEAVIS, 2004), o Mascot (PERKINS et al., 1999), o ProLuCID (XU et al., 2015) e o Comet (ENG et al., 2013).

### 1.3.1 Espectrometria de massas

A espectrometria de massas (MS) é uma ferramenta analítica com aplicações em diversos campos da ciência como: na biologia, medicina, biotecnologia, forense, física, entre outros. Para a proteômica, é a principal ferramenta para identificação de peptídeos e proteínas podendo até mesmo caracterizar modificações pós-traducionais, perfis de expressão, interações e estruturas proteicas, entre outros (FERREIRA et al., 2009; LIEBLER, 2001).

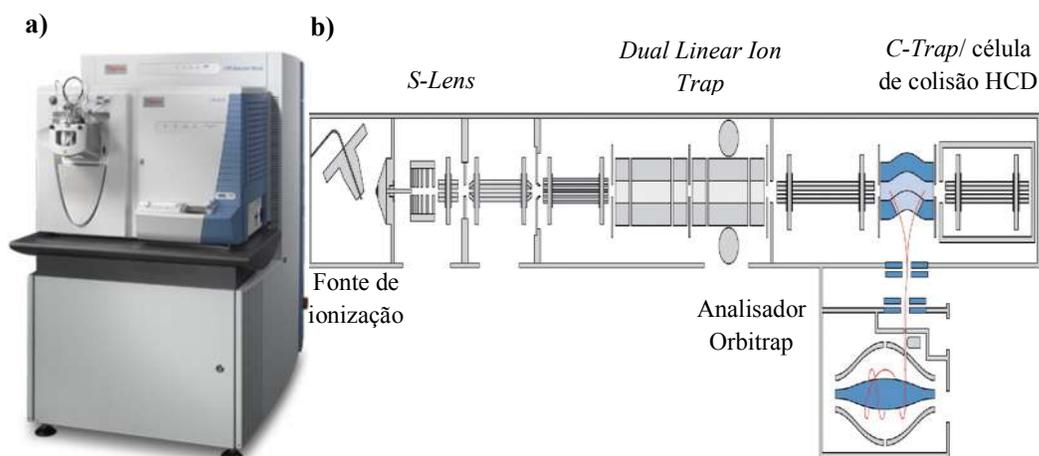
A aplicação da espectrometria de massas na biologia tem sido alavancada principalmente pela criação de técnicas de ionização consideradas brandas, onde se destacam: o ESI –*Electrospray ionization* (FENN et al., 1989) e o MALDI - *Matrix-assisted laser desorption/ionization* (TANAKA et al., 1988). A ionização por *eletrospray* (ESI) produz íons a partir de uma amostra em fase líquida. Já a ionização por MALDI é associada principalmente com os analisadores de tempo de voo (TOF) e ioniza o analito que se encontra co-cristalizado com uma matriz composta de um ácido orgânico de baixo peso molecular (HILLENKAMP et al., 1991; HILLENKAMP; KARAS, 1990).

Avanços tecnológicos possibilitaram a criação de espectrômetros híbridos (*i.e.*, com mais de um analisador de massas) capazes de aplicar metodologias complementares de dissociação de biomoléculas (*i.e.*, a dissociação por colisão induzida - CID, a dissociação por colisão em alta energia – HCD, e a dissociação por transferência de elétron - ETD), os quais têm proporcionado cada vez mais vislumbrar perfis proteômicos mais complexos. Avanços também têm sido realizados principalmente na melhoria da acurácia nas medidas de massa/carga dos íons (ZHANG et al., 2013).

Resumidamente, um espectrômetro de massas, em sua forma mais simples, é constituído por três componentes: uma fonte de ionização, responsável pela vaporização e ionização da amostra; um analisador de massas que separa os íons de acordo com  $m/z$ , e o detector, que registra os sinais dos íons.

Neste trabalho, foi utilizado o LTQ Orbitrap Velos (*Thermo Fisher Scientific*) (Figura 6a). Este instrumento possui uma fonte de ionização nano-ESI, dois tipos de analisadores: um *Ion Trap* linear e um Orbitrap, e detectores conforme demonstrado na Figura 6b. O analisador *Ion Trap* linear é composto por duas células lineares similares a do analisador quadrupolo que proporcionam um isolamento rápido e captura/seleção de íons para dissociação. Ambas as células têm a mesma fonte comum de RF e AC, mas tensões de corrente contínua independentes, o que permite a transferência de íons de uma célula para outra. A primeira célula do *ion trap* linear é operada em uma pressão maior por estar otimizada em aprisionar e dissociar

íons. Já a segunda, está otimizada para separação dos íons pelo  $m/z$  fornecidos pela primeira célula (“<http://www.thermoscientific.com>”; OLSEN et al., 2009). O analisador Orbitrap é caracterizado pela sua alta resolução e acurácia (MAKAROV, 2000). Este aprisiona os íons fazendo com que orbitem em torno de um eletrodo central enquanto oscilam harmonicamente ao longo do eixo com frequências relativas aos valores de  $m/z$ . Tal aprisionamento deve-se ao campo eletrostático aplicado, no qual a atração eletrostática do eletrodo interno se iguala à força centrípeta. A determinação da razão massa/carga dos íons aprisionados se faz aplicando a transformada de *Fourier* no sinal proveniente do detector (MAKAROV, 2000). Portanto, a presença de ambos os analisadores neste espectrômetro, proporciona tirar proveito das vantagens de cada um; geralmente obtendo a massa dos precursores com alta resolução pelo Orbitrap e associando a alta velocidade do LTQ na dissociação e análise dos íons parentais.



**Figura 6 - A)** Imagem ilustrativa de um espectrômetro de massas *Linear trap* quadrupolo (LTQ) Orbitrap Velos® e **B)** Esquema desse instrumento. Fonte: modificado de “<http://www.thermoscientific.com>”; OLSEN et al., 2009.

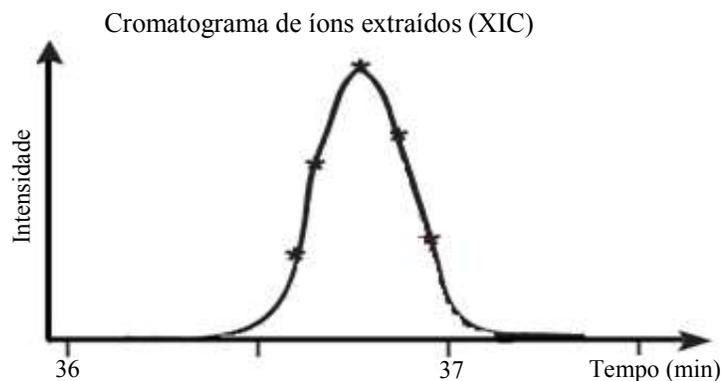
### 1.3.2 Proteômica quantitativa

Cada vez mais, métodos quantitativos têm sido partes integrantes dos atuais estudos de proteômica. Dois tipos principais se destacam: a quantificação absoluta e a quantificação relativa (ZHANG et al., 2013).

A quantificação absoluta tem como princípio a introdução de padrões na amostra, cuja as concentrações são previamente conhecidas para a determinação da quantidade de uma substância (proteína ou peptídeo) em questão. Esses padrões, podem ser proteínas marcadas isotopicamente com um número limitado de peptídeos e concentrações conhecidas. Entre os principais métodos desse tipo de quantificação tem-se o AQUA (*absolut quantification*) (GERBER et al., 2003). Nesse, os padrões internos são peptídeos sintetizados a partir de

isótopos estáveis incorporados entre suas moléculas, assemelhando-se aos peptídeos primários formados na proteólise. Esses peptídeos sintéticos também podem ser preparados levando em consideração a análise das modificações pós-traducionais. Dessa forma, o AQUA baseia-se na comparação relativa entre o peptídeo marcado com isótopos e o respectivo da amostra. A quantificação absoluta tem se mostrado promissora na proteômica direcionada (*targeted proteomics*). No entanto, atualmente, ela é limitada a um número reduzido de peptídeos, principalmente devido a necessidade do padrão interno (BANTSCHIEFF et al., 2007; ZHANG et al., 2013).

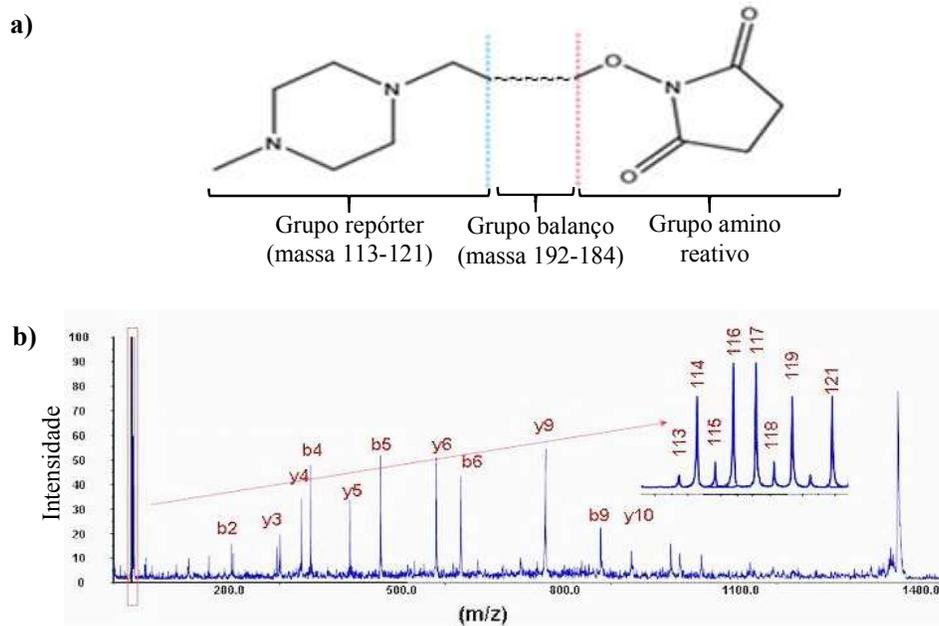
A quantificação relativa possibilita uma estimativa da abundância relativa de peptídeos em diferentes amostras. Há dois métodos principais desse tipo de quantificação: a livre ou isenta de marcação (*label-free*) e a dependente de marcação. A estratégia *label-free* avalia as amostras comparando a intensidade dos íons oriundos dos peptídeos, entre diferentes corridas no espectrômetro de massas. As principais técnicas de quantificação utilizadas em *label-free* são: cromatograma de íons extraídos, ou XIC – também conhecido por cromatograma reconstruído de íons (RIC), onde um ou mais valores  $m/z$  representando um determinado peptídeo é (são) “extraído (s)” do cromatograma. Neste, a intensidade do íon precursor ao longo do tempo é plotada em um gráfico, representando os pontos da análise (Figura 7). A tolerância na leitura do  $m/z$  deste precursor depende da acurácia e da resolução do espectrômetro de massas utilizado. Finalmente, o XIC é obtido calculando a área abaixo da curva representada pelos pontos plotados. Este valor, pode então ser comparado com outros XIC's do mesmo peptídeo em amostras diferentes. Alternativamente, pode-se usar o número de espectros de MS2 adquiridos e mapeados para uma proteína como um indicador da abundância relativa desta na amostra (*spectral counts*).



**Figura 7** – Representação de um cromatograma de íons extraídos (XIC). Fonte: modificado de ONG; MANN, 2005.

Nas estratégias dependentes de marcação, estes marcadores podem ser introduzidos metabolicamente (*e.g.* SILAC), quimicamente (*e.g.* iTRAQ, TMT) ou enzimaticamente (*e.g.* marcação com  $^{18}\text{O}$ ) (BANTSCHIEFF et al., 2007; ONG; MANN, 2005; ZHANG et al., 2013). Estas abordagens permitem analisar múltiplas amostras em diferentes estágios simultaneamente, diminuindo assim, a variabilidade experimental e o tempo de análise por espectrometria de massas. Um exemplo desse tipo de marcação é o iTRAQ (*Isobaric Tags for Relative and Absolute Quantification*) – marcadores isóbaros (BONDARENKO et al., 2002; ZHANG et al., 2013) que são constituídos por um grupo amino reativo (éster N-hidroxisuccinimida) que reage com as aminas primárias dos peptídeos, um grupo balanceador de massas e um grupo marcador único (*e.g.* o *iTRAQ 8-plex* produz os fragmentos iônicos:  $m/z$  113, 114, 115, 116, 117, 118, 119 e 121) (Figura 8a). Esses reagentes são incorporados na parte N-terminal e nas cadeias laterais de lisinas dos peptídeos digeridos de uma mistura complexa. Os peptídeos derivados são cromatograficamente indistinguíveis. Contudo, estes produzem íons de assinatura ou íons marcadores diferentes, quando dissociados no MS2 (Figura 8b). De maneira que a identificação e a quantificação dos peptídeos originais na amostra são possíveis a partir da intensidade do sinal desses íons marcadores (ROSS et al., 2004; ZHANG et al., 2013).

Nesta tese, dois tipos de quantificação foram usados: livre ou isenta de marcação (*label-free*) e a marcação por isóbaros com o *iTRAQ 8-plex*.



**Figura 8 - A)** Estrutura química dos componentes do iTRAQ; **B)** Diagrama mostrando a localização dos íons marcadores do iTRAQ 8-plex no espectro de MS/MS. Fonte: modificado de “[http://tools.thermofisher.com/content/sfs/manuals/cms\\_049786.pdf](http://tools.thermofisher.com/content/sfs/manuals/cms_049786.pdf)”; “National Institute on Aging - Proteomics”.

#### 1.4 Projeto Proteoma Humano (Chromosome-Centric Human Proteome Project, C-HPP e Biology/Disease-driven Human Proteome Project, B/D-HPP)

Os diversos avanços na proteômica culminaram no surgimento de projetos ambiciosos como o projeto do proteoma humano – HPP, que objetiva mapear e identificar todo o proteoma humano (C-HPP) e o B/D-HPP (projeto proteoma humano dirigido à Biologia/Doença) que busca proteínas correlacionadas a doenças como o câncer. Este projeto é uma iniciativa da *Human Proteome Organization* (HUPO), no qual o consórcio é composto por várias equipes internacionais, sendo que a Unidade Proteômica da Universidade Federal do Rio de Janeiro (UFRJ) foi eleita para representar o Brasil, tendo como objetivo caracterizar proteínas do cromossomo 15. Neste sentido, o presente trabalho se integra principalmente no C-HPP; e em um recente estudo de perfis epigenéticos e de transcrição observou-se a associação da região genômica do cromossomo 15q25 ao câncer gástrico, no qual identificou-se um novo grupo de genes com metilação no DNA específico para esse tipo de câncer (KANG et al., 2015).

Entre os pontos centrais do C-HPP, destaca-se a busca pelas “*missing proteins*”. Estas podem ser definidas como proteínas que possuem apenas evidência por transcriptômica ou identificadas por anticorpos, mas sem confirmação por espectrometria de massas (“<http://www.c-hpp.org/>”).

Portanto, a informação obtida no projeto do proteoma humano auxiliará na compreensão dos processos moleculares que regulam o organismo na presença ou ausência de determinada patologia (AEBERSOLD et al., 2013) e com o intuito de auxiliar no fornecimento dessas informações será realizado nessa tese, a categorização de proteínas identificadas por cromossomos e quanto ao seu status de “*missing*”.

## 2. OBJETIVOS

### 2.1 GERAL

Investigar diferenças no perfil proteico de tecidos tumorais e das margens de ressecção para uma maior compreensão das possíveis modificações decorrentes do processo de malignização.

### 2.2 ESPECÍFICOS

- Evidenciar, por espectrometria de massas, mudanças proteicas ocorridas a diferentes distâncias anatômicas do tumor primário (adenocarcinoma gástrico).
- Avaliar o perfil proteico de fragmentos de tecido tumoral e das margens de ressecção com a ferramenta computacional *Shotgun Imaging Analyzer* utilizando a quantificação *label-free*.
- Correlacionar os dados das áreas (tumor e margem de ressecção) por histopatologia e por proteômica, empregando a quantificação por marcação com isóbaros - iTRAQ.
- Avaliar os resultados dos perfis proteômicos por ambos os tipos de quantificação utilizada – *label-free* e marcação com isóbaros (iTRAQ).

### 3. MATERIAL E MÉTODOS

O projeto foi submetido e aprovado pelo Comitê de Ética em Pesquisa (CEP) do Hospital Universitário Clementino Fraga Filho da Universidade Federal do Rio de Janeiro (HUCFF/UFRJ) sob N° 10252913.5.0000.5257-CAAE (Anexo 1).

#### 3.1 Seleção de pacientes

Foram selecionados três pacientes de acordo com os seguintes critérios:

❖ Critérios de inclusão:

- a) Pacientes com cânceres gástricos diagnosticados através de biópsia e endoscopia, tendo como diagnóstico histológico adenocarcinoma; e que tivessem indicação cirúrgica pelo corpo clínico dos hospitais responsáveis por cada coleta (Fundação CECON e/ou Hospital Universitário Gaffrée e Guinle - HUGG);
- b) Pacientes maiores de idade que concordassem em participar da pesquisa e independente do sexo.

❖ Critérios de exclusão:

- a) Os pacientes não deveriam apresentar histórico ou qualquer evidência de HIV, vírus da hepatite B ou C e a presença de outros carcinomas;
- b) A ocorrência prévia ou atual das doenças citadas acima foi anotada e a amostra correspondente foi descartada.

#### 3.2 Caracterização dos tumores

A caracterização e a classificação do tipo histológico de câncer gástrico dos pacientes selecionados foram realizadas mediante endoscopia digestiva, onde o material obtido (biópsias) foi analisado por HE (hematoxilina e eosina) para o diagnóstico de adenocarcinoma. O estadiamento do tecido tumoral, dos diferentes pacientes selecionados, foi realizado de acordo com a classificação de Lauren, a qual categorizou em dois tipos sub-histológicos: intestinal e difuso.

### 3.3 Coleta e preservação das amostras

Em todos os casos, o consentimento escrito dos pacientes (Termo de Consentimento Livre e Esclarecido – Anexo 2) foi obtido mantendo-se a devida privacidade. Além disso, foi obtido um questionário detalhado (Anexo 3) que permitiu, ao fim da pesquisa, correlacionar os dados experimentais e laboratoriais com a situação particular de cada paciente.

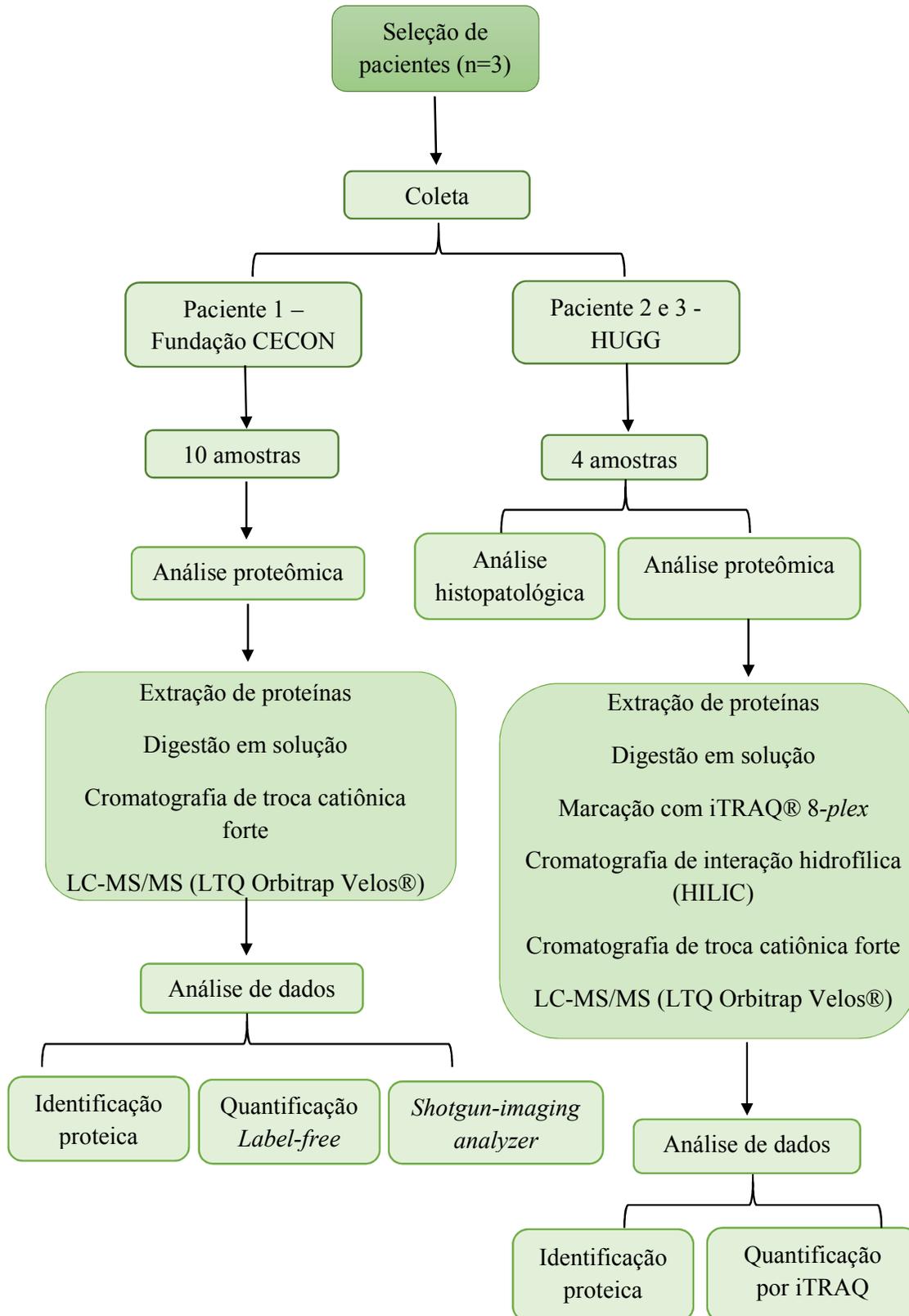
As amostras foram coletadas na Fundação CECON (Amazonas) e no Hospital Universitário Gaffrée e Guinle (HUGG), da Universidade Federal do Estado do Rio de Janeiro (UNIRIO – Rio de Janeiro), sob a orientação de um cirurgião do aparelho digestivo, especialista em Oncologia e de um patologista, ambos integrantes da equipe de pesquisa do projeto. Neste, foram obtidos tecidos coletados de peça cirúrgica do estômago de três pacientes, compreendendo a margem de ressecção e a região tumoral. O estadiamento do tumor foi determinado de acordo com a classificação de Lauren.

Os fragmentos coletados foram armazenados em tubos estéreis de 15 mL com tampa e transportados em gelo. O material coletado foi armazenado a -80°C na Universidade Federal do Amazonas – UFAM, quando a coleta foi realizada no Amazonas, e na Universidade Federal do Rio de Janeiro – UFRJ, quando coletado no Rio de Janeiro.

A seguir, encontra-se um fluxograma geral (Figura 9) de como foi feita a análise das amostras; este trabalho é composto por duas etapas majoritárias:

**Etapas A** – Análise do perfil proteico de tecidos das margens de ressecção e do tecido canceroso com a ferramenta computacional *Shotgun Imaging Analyzer* utilizando a quantificação *label-free*.

**Etapas B** – Análise histopatológica e proteômica, empregando a quantificação por marcação com isóbaros – iTRAQ, de tecidos coletados da mesma região (tumor e margem de ressecção).

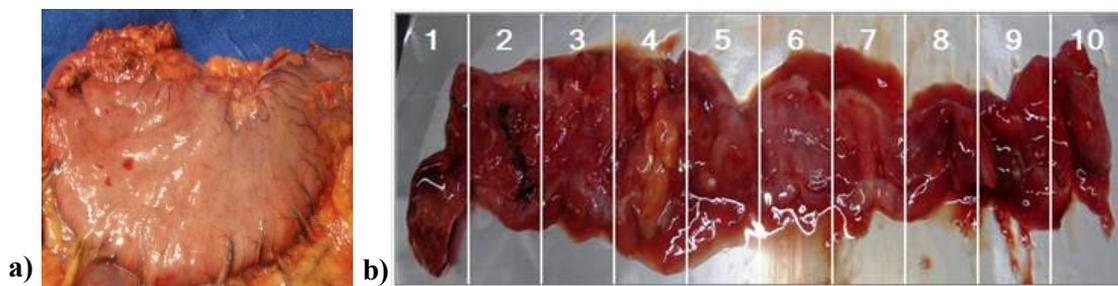


**Figura 9** – Fluxograma geral das análises realizadas nas amostras coletadas neste trabalho.

### 3.4 Etapa A – Análise do perfil proteico de tecidos das margens de ressecção e do tecido canceroso com a ferramenta computacional *Shotgun Imaging Analyzer* utilizando uma quantificação *label-free*.

#### 3.4.1 Pulverização do tecido, extração e dosagem de proteínas

Esta etapa consiste na análise de um tecido medindo 20 cm de um paciente (paciente 1) da Fundação CECON, compreendendo toda a região do estômago (Figura 10). Este foi dividido verticalmente em dez seções do mesmo tamanho (2 cm). Cada seção foi pulverizada e macerada com nitrogênio líquido, e se adicionou RapiGest® (*Waters*) a uma concentração de 0,1% (p/v) preparado com bicarbonato de amônio 50 mM. A quantidade de RapiGest adicionado em cada fragmento foi normalizada da seguinte forma: para cada 100 mg de amostra foram adicionados 200 µL de RapiGest 0,1%. Feito isso, as amostras foram submetidas a uma agitação por vinte segundos, depois centrifugadas a 4°C, por 30 min, a 20800 x g. Posteriormente, o sobrenadante foi retirado. A quantificação de proteínas presente em cada seção (amostra) foi realizada pelo ensaio fluorimétrico *Qubit*® 2.0 (*Invitrogen*), de acordo com as instruções do fabricante. Resumidamente, para cada microtubo, é adicionado 1 µL de cada amostra e mais 199 µL de solução de trabalho do *Qubit* contendo o reagente fluoróforo. Estes são agitados e incubados a temperatura ambiente por 15 minutos; e em seguida, a leitura da absorbância das amostras é realizada. A leitura de um branco, contendo apenas RapiGest 0,1%, foi realizada de maneira similar às amostras. Já a curva padrão foi obtida através da leitura de três padrões contidos no kit, preparados também de forma semelhante, modificando-se apenas os volumes do padrão (10 µL) e da solução de trabalho do *Qubit* (190 µL).



**Figura 10 - A)** Imagem obtida através da ressecção cirúrgica do estômago completo de um paciente; **B)** Tecido desse estômago dividida verticalmente em 10 seções.

### 3.4.2 Digestão trípica em solução

Cem microgramas de proteínas foram reduzidos com adição de 20 mM de ditioneitol (DTT) por trinta minutos a uma temperatura de 60°C. Em seguida, após atingirem a temperatura ambiente, as amostras foram alquiladas com a adição de iodoacetamida (IAA) a 66,7 mM por trinta minutos à temperatura ambiente e no escuro. Mediu-se o pH e em seguida, todas as amostras foram digeridas com a enzima tripsina (Promega) na proporção de 1/50 (p/p) de enzima/substrato. Incubou-se a 37°C por 20 horas. Parou-se a reação com ácido fórmico a 10% (v/v) para uma concentração final de 1% (v/v) deixando por quarenta minutos a temperatura ambiente. Após esse tempo, houve a degradação do RapiGest®, então centrifugou-se a 6°C por trinta minutos a 20800 x g para a remoção de todo material insolúvel.

### 3.4.3 Dessalinização e concentração das amostras

Sessenta microgramas de peptídeos trípticos de cada amostra foram dessalinizados e concentrados mediante o uso de mini-colunas feitas em *tips* de 200 µL, empacotadas com a resina Poros® R2 (*Applied Biosystems*). Cada coluna foi ativada com acetonitrila, equilibrada com TFA 1% e as amostras preparadas em TFA 1%. Carregou-se as amostras na coluna e esta foi lavada com TFA 0,1%. A eluição dos peptídeos da coluna foi realizada com a solução de acetonitrila 70% e TFA 0,1%. Estes foram concentrados no *Speed Vacuum* (*Thermo Fisher Scientific*).

### 3.4.4 Cromatografia de troca catiônica forte

Os peptídeos concentrados na etapa anterior foram ressuspensos na solução A (5 mM KH<sub>2</sub>PO<sub>4</sub> em 25% ACN, pH 3) e aplicados em uma “*macro spin column*” de troca catiônica forte (*Harvard Apparatus*, nº de referência 744202) de acordo com as instruções do fabricante. Resumidamente, as “*macro spin column*” foram ativadas duas vezes com 500 µL da solução A por 10 minutos a temperatura ambiente. Centrifugou-se cada coluna a 1500 x g por três minutos. As amostras foram carregadas na coluna e centrifugou-se novamente conforme descrito anteriormente. A eluição dos peptídeos ocorreu em etapas através da aplicação sequencial de 300 µL da solução A com concentrações crescentes de KCl, as quais foram 85 mM, 150 mM, 250 mM e 400 mM. Após a aplicação de cada passo salino, centrifugou-se as colunas contendo

as amostras a 1500 x g por três minutos. Foram obtidas 40 frações, as quais foram concentrados no *Speed Vacuum* (*Thermo Fisher Scientific*).

#### 3.4.5 Dessalinização, concentração e quantificação das amostras

Cada fração resultante da cromatografia de troca catiônica (item 3.4.4) foi dessalinizada e concentrada mediante o uso de mini-colunas feitas em *tips* de 200  $\mu\text{L}$ , empacotadas com a resina Poros® R2 (*Applied Biosystems*) conforme o procedimento descrito no item 3.4.3. Os peptídeos foram ressuspensos em 20  $\mu\text{L}$  de solução de ácido fórmico 0,1%, sendo a quantificação realizada através do ensaio fluorimétrico – *Qubit 2.0*® (*Invitrogen*) de acordo com as instruções do fabricante (item 3.4.1).

#### 3.4.6 Aquisição de dados de LC- MS/MS

Em triplicata técnica, 1  $\mu\text{g}$  dos peptídeos provenientes da digestão e da troca catiônica forte foram injetados em um volume de 4  $\mu\text{L}$  numa pré-coluna de fase reversa empacotadas *in house*, com fase estacionária *ReprosilPur C18 Acqua* (esferas de 5  $\mu\text{m}$  de diâmetro e tamanho de poro 120 Å, Dr. Maisch), com comprimento de 3 cm e diâmetro interno de 150  $\mu\text{m}$ . Em seguida, estes peptídeos foram submetidos a uma coluna de fase reversa empacotadas também *in house*, com fase estacionária *ReprosilPur C18 Acqua* (esferas de 3  $\mu\text{m}$  de diâmetro e tamanho de poro 120 Å, Dr. Maisch), com comprimento de 20 cm e diâmetro interno de 75  $\mu\text{m}$ , utilizando um sistema de nanocromatografia, *Easy-nLC II* (*Proxeon*). Os peptídeos foram eluídos da coluna com um gradiente cromatográfico de 5-50% de fase B (95% de acetonitrila em 0,1% de ácido fórmico) por 100 minutos. Posteriormente, a coluna foi lavada com o gradiente de 50-95% da mesma fase móvel por 20 minutos, com fluxo de 200 nL/min, diretamente no espectrômetro de massas LTQ Orbitrap Velos (*Thermo Fisher Scientific*).

Aquisição dos espectros no LTQ Orbitrap Velos foi realizada no modo *data dependent acquisition* (DDA), alternando automaticamente entre a aquisição do *full scan* MS e MS/MS, com uma lista de exclusão dinâmica de 90s. Após a varredura inicial do MS no sistema com resolução de 60,000 a 400  $m/z$ , os seis íons mais intensos, com carga igual ou superior a 2+ foram isolados, um de cada vez, e fragmentados pela dissociação induzida por colisão (CID) e dissociação por colisão em alta energia (HCD), usando energias de colisão normalizadas de 35 e 50, respectivamente. Os parâmetros empregados para a operação do espectrômetro de massas foram: voltagem no spray de 2,5 kV, fluxo zero de *sheath* e *auxiliary gas*, temperatura de 235°C

no capilar, a habilitação de *predictive automatic gain control* (AGC) e um nível de 70% de RF na *S-lens*, sendo as funções de *scan* do espectrômetro de massas e os gradientes dos solventes no nLC controlados pelo *software* Xcalibur 2.0 (*Thermo Fisher Scientific*).

### 3.4.7 Análise de dados

#### 3.4.7.1 *Peptide Spectrum Matching* – PSM

Todos os passos de análise usados nessa etapa A foram realizados usando o *PatternLab for proteomics* (PL) versão 3.2. Sequências proteicas do *Homo sapiens* foram obtidas do NeXtProt em 15 de julho de 2013. Em seguida, o módulo de preparação de banco de dados do PL foi usado para preparar um banco de sequências do tipo “*target-decoy*”, e também incluir as sequências de 127 contaminantes comuns a experimentos de espectrometria de massas (exemplos: queratina, tripsina). A ferramenta de busca *ProLuCID* (versão 1.3) foi usada para comparar os espectros de massas em *tandem* obtidos experimentalmente com os espectros teóricos gerados a partir do banco de dados, selecionando os peptídeos candidatos mais prováveis (PSMs) (XU et al., 2015).

Resumidamente, a busca foi limitada a candidatos trípticos e semi-trípticos; a carbamidometilação foi imposta como uma modificação fixa e a oxidação da metionina como variável. Nesta etapa, qualquer peptídeo candidato com massa de até 50 ppm da massa mensurada foi considerado, e o *XCorr* e o *Z-Score* foram utilizados como métricas de semelhanças primária e secundária, respectivamente.

#### 3.4.7.2 Validação do PSMs

A validação das identificações proteicas foi realizada pelo *Search Engine Processor – SEPro* (CARVALHO; FISCHER; XU; COCIORVA; et al., 2012). Resumidamente, as identificações foram agrupadas pelo estado de carga do precursor ( $2+$  e  $\geq 3+$ ) e, em seguida, pelo *status* tríptico (tríptico ou semi-tríptico), resultando em quatro subgrupos distintos. Para cada grupo, os valores do *XCorr*,  $\Delta CN$  e *Z-Score* resultante do ProLuCID foram usados para gerar um discriminador Bayesiano a fim de determinar uma função discriminante entre as identificações de não-*decoys* e *decoys*. As identificações foram classificadas em uma ordem crescente de acordo com os valores do discriminador.

Os valores mínimos (*cut-off*) foram estabelecidos para que a lista final de identificações contivesse 1% de *decoys* (*i.e.*, 1% FDR). Este procedimento foi realizado de forma independente em cada subconjunto de dados, resultando em uma taxa de falso-positivo que era independente do estado triptico ou do estado de carga. Adicionalmente, uma restrição a sequências com seis ou mais resíduos de aminoácidos foi imposta. Os resultados foram pós-processados para aceitar PSMs com menos de 5 ppm de erro, e proteínas que tivessem duas ou mais evidências independentes na identificação (por exemplo, a identificação de um peptídeo com diferentes estados de carga, uma versão modificada e não modificada do mesmo peptídeo, ou dois peptídeos diferentes). Este último filtro acabou resultando em 0% de FDR nas identificações. Para a análise quantitativa de *shotgun-imaging*, foram consideradas apenas proteínas que tivessem pelo menos 1 peptídeo único identificado.

#### 3.4.7.3 Quantificação relativa de proteínas

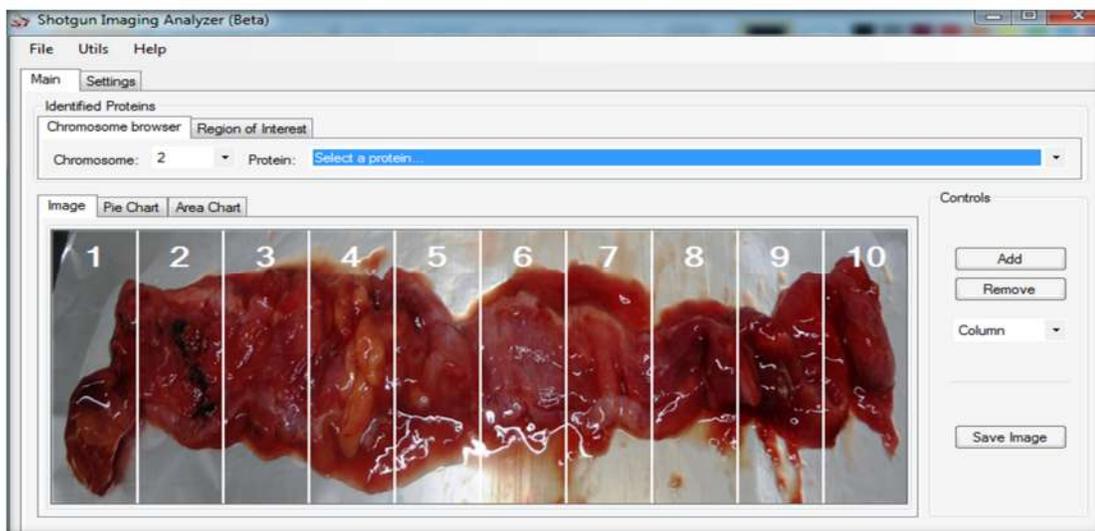
Os arquivos de MS1 foram extraídos com o módulo *RawReader* do PL e deconvoluídos usando as configurações padrão do Y.A.D.A. para a proteômica “*bottom-up*” (CARVALHO et al., 2009). O módulo de quantificação do *PatternLab* foi utilizado para a obtenção dos cromatogramas de íons extraídos (XICs), a partir dos arquivos contendo os perfis de massas (MS1) deconvoluídos e correlacionados com as sequências de peptídeos correspondentes (CARVALHO; FISCHER; XU; YATES; et al., 2012). Em seguida, o módulo *Regrouper* do *PatternLab* foi empregado para normalizar os dados quantitativos, de acordo com a abordagem de distribuição normalizada do fator de abundância do íon (dNIAF), o qual é uma adaptação da abordagem de distribuição normalizada de fatores de abundância espectrais (dNSAF) (ZHANG; WEN; WASHBURN; et al., 2010) para os XICs. Posteriormente, os arquivos *sparse matrix* e *index* do PL foram gerados (CARVALHO; FISCHER; XU; YATES; et al., 2012), servindo de entrada para a nossa ferramenta computacional de imagem, junto com as informações do NeXtProt e a imagem do tumor.

#### 3.4.7.4 O programa *Shotgun-Imaging Analyzer* (SIA)

O programa *Shotgun-Imaging Analyzer* (SIA) foi desenvolvido pelo nosso grupo para permitir uma avaliação visual dos resultados proteômicos. Ele está disponível para *download* através de (“<http://proteomics.fiocruz.br/software/shotgunImaging/>”). Nesse, o usuário pode gerar um mapa de calor de acordo com a abundância de proteínas nas áreas onde o tecido foi

seccionado. Assim, uma vez que a proteína é selecionada, a interface gráfica do programa (Figura 11) utilizará uma escala de cinza a azul claro para indicar as regiões onde a proteína não foi identificada, e tons diferentes de vermelho, intensificando com a abundância, para indicar áreas onde a mesma foi identificada.

Adicionalmente, o *Shotgun-Imaging Analyzer* possui funções que atendem aos objetivos do C-HPP, permitindo categorizar proteínas de acordo com o cromossomo. Vale ressaltar, que apenas proteínas identificadas nesse experimento foram listadas para os respectivos cromossomos. Outra opção disponível no *software* é a de apontar as *missing proteins* de acordo com o NeXtProt.

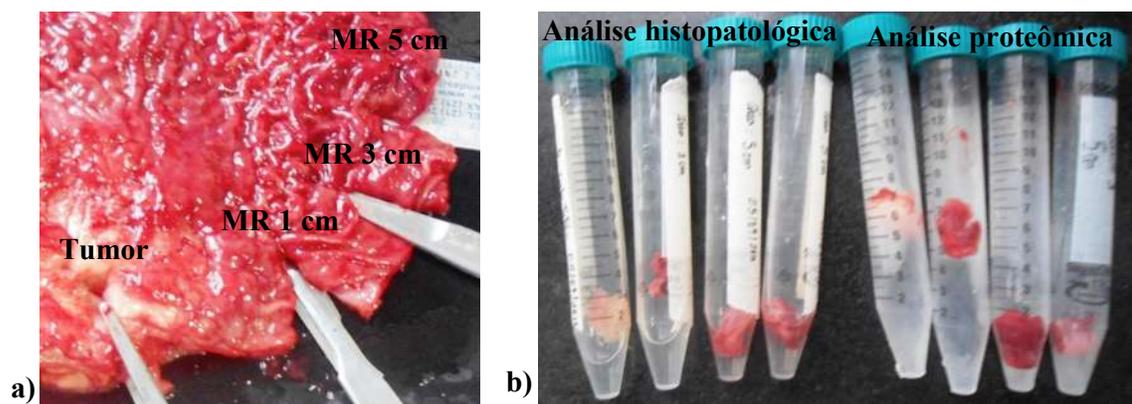


**Figura 11** - Interface gráfica do usuário (GUI) exibindo a imagem de um tecido com câncer gástrico analisado no trabalho.

Finalmente, pode-se também procurar automaticamente proteínas que estão mais abundantes em uma região de interesse (ROI), ou seja, proteínas identificadas em uma ordem decrescente de acordo com o índice de Golub (em outras palavras, este índice indica a diferença das médias dos valores de quantificação dividida pela soma dos seus desvios-padrão) na área tumoral, intermediária (tumor-margem) ou na margem de ressecção.

### 3.5 Etapa B - Análise histopatológica e proteômica, empregando a quantificação por marcação com isóbaros – iTRAQ, de tecidos coletados da mesma região (tumor e margem de ressecção).

Nesta segunda parte, foram coletados tecidos de dois pacientes (paciente 2 e 3) do Hospital Universitário Gaffrée e Guinle (HUGG). Estes compreenderam: um fragmento da região tumoral e três fragmentos da margem de ressecção a diferentes distâncias do tumor, ou seja, fragmento da margem de ressecção (MR) a 1 cm, 3 cm e 5 cm da neoplasia (Figura 12a) totalizando quatro fragmentos de cada paciente. Posteriormente, cada um desses fragmentos foi seccionado horizontalmente em duas partes, sendo que em uma parte foi feita a análise histopatológica e na outra, a análise proteômica (Figura 12b).



**Figura 12 - A)** Imagem do estômago de um paciente compreendendo os quatro fragmentos estudados (tumor, margem de ressecção - MR 1, 3 e 5 cm); **B)** Seção do estômago dividida horizontalmente em duas partes, uma para análise histopatológica e outra para a análise proteômica.

#### 3.5.1 Análise histopatológica

Os quatro fragmentos destinados às análises histopatológicas foram imediatamente inseridos em solução de formalina (formol) a 10%. Essa análise foi realizada no departamento de Anatomia Patológica do Hospital Universitário Gaffrée e Guinle (HUGG) sob responsabilidade do Dr. Rodrigo Panno Basílio de Oliveira, um dos colaboradores do trabalho. Resumidamente, os cortes histológicos dos quatro fragmentos de tecido de cada paciente foram emblocados em parafina e posteriormente cortados e transferidos para uma lâmina que foi corada com HE (hematoxilina e eosina) para a identificação do tipo histopatológico do tumor.

### 3.5.2 Análise proteômica

#### 3.5.2.1 Pulverização do tecido, extração e dosagem de proteínas

Os quatro fragmentos de tecido foram pulverizados e macerados com nitrogênio líquido, sendo adicionado RapiGest<sup>®</sup> (*Waters*) a uma concentração de 0,1% (p/v), preparado com o tampão bicarbonato de trietilamônio 50 mM (TEAB do inglês, *triethylammonium bicarbonate*). A normalização da concentração das amostras, a obtenção do sobrenadante e a quantificação de proteínas presente em cada fragmento foi realizada conforme descrito no item 3.4.1.

#### 3.5.2.2 Digestão triptica em solução

Cem microgramas de proteínas foram reduzidos com adição de cloridrato de tris (carboxietil) fosfina (TCEP do inglês, *tris (2-carboxyethyl) phosphine hydrochloride*) a 20 mM por trinta minutos a uma temperatura de 60°C. As amostras foram resfriadas à temperatura ambiente e posteriormente alquiladas com a adição de iodoacetamida (IAA) a 66,7 mM por trinta minutos à temperatura ambiente e no escuro. O pH das amostras foi medido e a digestão ocorreu mediante a adição da enzima tripsina (Promega) na proporção de 1/50 de enzima/substrato. Incubou-se a 37°C por 20 horas. Parou-se a reação com ácido trifluoroacético (TFA) a 10% (v/v) para uma concentração final de 0,5% (v/v) deixando por quarenta minutos a temperatura ambiente para a completa degradação do RapiGest<sup>®</sup>. Centrifugou-se a 6°C por trinta minutos a 20800 x g e retirou-se o sobrenadante.

#### 3.5.2.3 Dessalinização, concentração e quantificação dos peptídeos

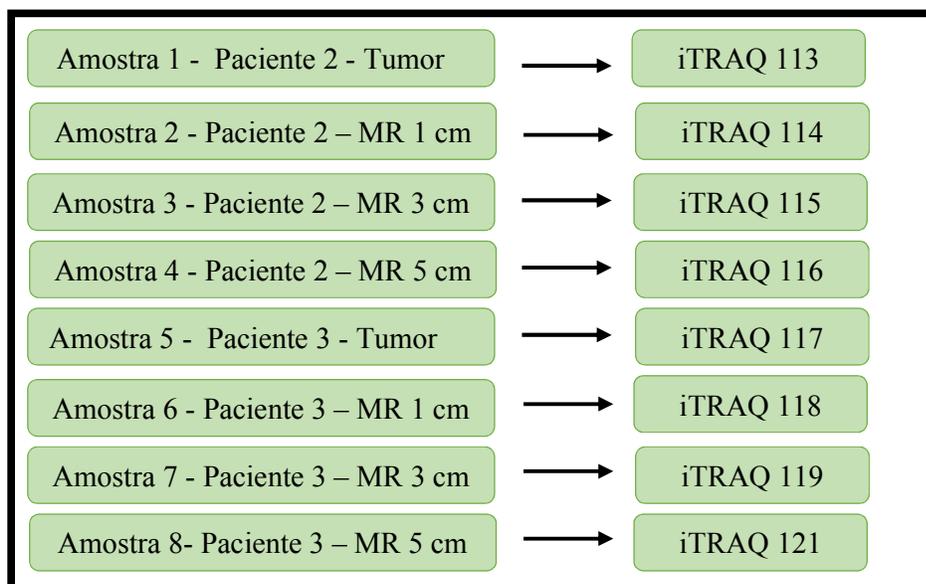
As amostras foram dessalinizadas e concentradas com colunas “*macro spin*” de C18 (*Harvard Apparatus*, nº de referência 744202) conforme recomendações do fabricante. Resumidamente, cada coluna foi ativada duas vezes com 500 µL de acetonitrila 100% por 10 minutos a temperatura ambiente e equilibrada três vezes com 300 µL de TFA 1%. As amostras resultantes da digestão que estavam em TFA 0,5% foram aplicadas nas respectivas colunas, sendo que se completou o volume com TEAB 50mM nas amostras que apresentaram um volume inferior a 100 µL. Lavou-se as colunas com TFA 0,1% e os peptídeos foram eluídos

com a solução de TFA 0,1% em acetonitrila 50% seguido de TFA 0,1% em acetonitrila 70%, ambos coletados no mesmo tubo.

Após cada passo (ativação, equilíbrio, aplicação da amostra, lavagem e eluição) foi realizada uma centrifugação a 2000 x g a 25°C por 3 minutos. Feito isso, os peptídeos foram concentrados no *Speed Vacuum (Thermo Fisher Scientific)*. Estes foram ressuspensos em 50 µL de TEAB 50 mM, e a quantificação foi feita pelo ensaio fluorimétrico – *Qubit 2.0® (Invitrogen)* de acordo com as instruções do fabricante (item 3.4.1).

#### 3.5.2.4 Marcação dos peptídeos com isóbaros (iTRAQ – 8 plex)

Os peptídeos obtidos no item 3.5.2.3 foram marcados com os reagentes do iTRAQ-8 plex seguindo as instruções do fabricante. Resumidamente, em cada marcador isóbaro (*tag*) foi adicionado 100 µL de isopropanol 100%, mantendo-se assim a concentração do solvente orgânico superior a 60%, quando fosse adicionado à amostra. Fez-se uma agitação em cada tubo e centrifugou-se a 15800 x g por 2 minutos. De acordo com os valores obtidos na quantificação (item 3.5.2.3), decidiu-se marcar 20 µg de peptídeos de cada amostra, totalizando ao final 160 µg (8 amostras x 20 µg). Para isso, retirou-se o volume correspondente a 20 µg e completou-se para um volume final de 50 µL com TEAB 50 mM. Feito isso, mediu-se o pH e as amostras foram incubadas a temperatura ambiente por duas horas. A marcação de cada amostra ocorreu como descrito na Figura 13.



**Figura 13** - Esquema da marcação com isóbaros (iTRAQ – 8 plex) nas amostras estudadas.

Passado esse tempo, as amostras foram reunidas em um único tubo, no qual foi realizado uma agitação e centrifugou-se a 15800 x g por 2 minutos. Os peptídeos marcados (160 µg) foram concentrados até um volume aproximado de 20 µL (*Speed Vacuum - Thermo Fisher Scientific*). Estes foram ressuspensos em 80 µL de TFA 0,1%, totalizando uma solução de 100 µL, a qual foi quantificada pelo ensaio fluorimétrico – *Qubit 2.0® (Invitrogen)* de acordo com as instruções do fabricante (item 3.4.1).

#### **3.5.2.4 Cromatografia de interação hidrofílica (HILIC)**

A partir da leitura da média da concentração dos peptídeos obtidos pelo ensaio fluorimétrico – *Qubit 2.0*, calculou-se o volume referente a 60 µg de peptídeos. Estes foram dessalinizados e concentrados mediante o uso de uma mini-coluna feita em um *tip* de 200 µL, empacotada com a resina Poros® R2 (*Applied Biosystems*). Esta coluna foi lavada com acetonitrila 100%, equilibrada com TFA 1% e a amostra que estava em TFA 0,1% foi aplicada na coluna. Lavou-se com TFA 0,1% e os peptídeos foram eluídos com a solução de TFA 0,1% em acetonitrila 90%.

Esta amostra dessalinizada foi introduzida em uma coluna de interação hidrofílica *TSKgel – Amide 80* (5 cm x 2 mm; 3 µm – tamanho da partícula; *Tosoh Bioscience*) usando um sistema de HPLC (*high performance liquid chromatograph*) da *Shimadzu* presente no Departamento de Bioquímica (UFRJ). Neste, foi aplicado um gradiente cromatográfico de fase B (100% de água em 0,1% de TFA) da seguinte maneira: 0-5% em 20 minutos, seguido de 5-22% em 40 minutos, 22-45% em 23 minutos, 45-60% em 5 minutos e retornando à condição inicial de 0% da fase B por 2 minutos. O fluxo foi de 0,2 mL/min, sendo monitorado as absorvâncias a 216 e 280 nm. Um total de 36 frações foram coletadas, as quais foram reunidas em oito frações. Essas foram concentradas no *Speed Vacuum (Thermo Fisher Scientific)*. Posteriormente, ressuspendeu-se em 20 µL da solução de ácido fórmico 0,1%, sendo a quantificação realizada através do ensaio fluorimétrico – *Qubit 2.0® (Invitrogen)* de acordo com as instruções do fabricante (item 3.4.1).

#### **3.5.2.5 Cromatografia de troca catiônica forte**

O restante da solução contendo os demais peptídeos provenientes da marcação com iTRAQ foram concentrados até um volume aproximado de 15 µL no *Speed Vacuum (Thermo Fisher Scientific)* e ressuspensos em 85 µL de solução B (10mM KH<sub>2</sub>PO<sub>4</sub> em 25% ACN, pH

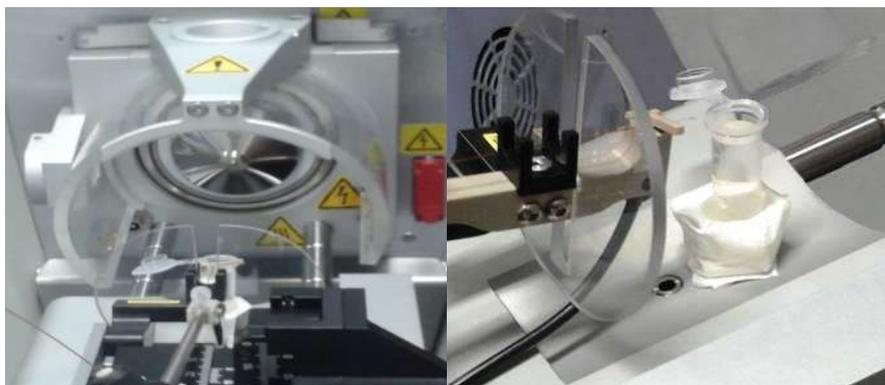
3), totalizando um volume de 100  $\mu\text{L}$ . Esses foram aplicados em “*spin columns*” de troca catiônica forte (*Harvard Apparatus*, nº de referência 744202) de acordo com as instruções do fabricante e descrito no item 3.5.2.3, sendo que a eluição dos peptídeos ocorreu por etapas, através da aplicação da solução B com concentrações crescentes de KCl de 75 mM, 150 mM, 250 mM e 400 mM. Estes resultaram em cinco frações (quatro passos salinos e o *flow through*), as quais foram concentradas no *Speed Vacuum (Thermo Fisher Scientific)*. A dessalinização, concentração e quantificação dessas amostras ocorreu conforme descrito no item 3.4.5.

### 3.5.2.6 Aquisição de dados de LC- MS/MS

Os peptídeos provenientes dos passos descritos acima (1  $\mu\text{g}$ ) foram carregados em um volume de 4  $\mu\text{L}$  numa pré-coluna de fase reversa empacotadas *in house*, com fase estacionária *ReposilPur C18 Acqua* (esferas de 5  $\mu\text{m}$  de diâmetro e tamanho de poro 120 Å, Dr. Maisch), com comprimento de 3 cm e diâmetro interno de 150  $\mu\text{m}$ . Posteriormente, estes peptídeos foram submetidos a uma coluna de fase reversa empacotadas também *in house*, com fase estacionária *ReposilPur C18 Acqua* (esferas de 3  $\mu\text{m}$  de diâmetro e tamanho de poro 120 Å, Dr. Maisch), com comprimento de 20 cm e diâmetro interno de 75  $\mu\text{m}$ , utilizando sistema de nanocromatografia, *Easy-nLC II (Proxeon)*; este processo foi realizado 3 vezes resultando em triplicata técnica. Os peptídeos foram eluídos da coluna com um gradiente cromatográfico de 5-40% da fase B (95% de acetonitrila em 0,1% de ácido fórmico) por 107 minutos. Posteriormente, os peptídeos foram eluídos com o gradiente de 40-95% da mesma fase móvel por 13 minutos, com fluxo de 200 nL/min, diretamente no espectrômetro de massas LTQ Orbitrap Velos (*Thermo Fisher Scientific*). Uma solução de hidróxido de amônio 5% foi posicionada na fonte de ionização entre a saída do capilar e a entrada no espectrômetro de massas (Figura 14).

Aquisição dos espectros no LTQ Orbitrap Velos foi realizada no modo *data dependent acquisition* (DDA), alternando automaticamente entre a aquisição do *full scan* MS e MS/MS, com uma lista de exclusão dinâmica de 60 s. Após a varredura inicial do MS ( $m/z$  350-2000) no sistema com resolução de 60,000 em 400  $m/z$ , os dez íons mais intensos com carga igual ou superior a 2+ foram isolados, um de cada vez, e fragmentados por dissociação por colisão em alta energia (HCD) usando energia de colisão normalizada de 35. Os parâmetros empregados para a operação do espectrômetro de massas foram: voltagem no *spray* de 1,8 kV, fluxo zero de *sheath* e *auxiliary gas*, temperatura de 200°C no capilar, a habilitação de *predictive automatic gain control* (AGC) e um nível de 69% de RF na *S-lens*, sendo as funções de *scan*

do espectrômetro de massas e os gradientes dos solventes no nLC controlados pelo *software Xcalibur 2.0 (Thermo Fisher Scientific)*.



**Figura 14** – Imagem da solução de hidróxido de amônio 5% posicionada na fonte de ionização entre a saída do capilar e a entrada no espectrômetro de massas.

### 3.5.3 Análise dos dados proteômicos

#### 3.5.3.1 *Peptide Spectrum Matching* – PSM

A análise de dados dessa etapa B foi realizada através de ferramentas do programa *PatternLab for proteomics (PL)* versão 4.0 (disponível em “<http://patternlabforproteomics.org>”). O banco de dados foi preparado usando a abordagem “*target-decoy*” para que pudéssemos realizar posteriormente uma análise estatística fornecendo um conjunto de identificações confiáveis. Este foi composto de sequências de *Homo sapiens* obtidas do UniProt (“<http://www.uniprot.org>”) em 04 de julho de 2014, incluindo também a versão reversa de cada sequência encontrada no banco de dados e as sequências de 127 contaminantes mais comuns na espectrometria de massas (exemplos: queratina, tripsina).

A ferramenta de busca, *Comet*, integrada ao PL 4.0 foi usada para comparar os espectros de massas em *tandem* obtidos experimentalmente com os espectros teóricos gerados a partir do banco de dados, selecionando os candidatos mais prováveis (ENG et al., 2013). Resumidamente, a busca foi limitada a candidatos trípticos considerando até duas falhas de clivagem (do inglês, *missed cleavage*); a carbamidometilação da cisteína e as modificações do iTRAQ- 8 *plex* (N-terminal e cadeia lateral de lisina) foram especificadas como modificações fixas. Para cada espectro analisado, foram considerados sequências oriundas do banco de dados com tolerância de 40 ppm de erro; e o *XCorr* foi utilizado como métrica de semelhança primária.

### 3.5.3.2 Validação do PSMs

A validação das identificações proteicas foi realizada pelo *Search Engine Processor* – SEPro (CARVALHO; FISCHER; XU; COCIORVA; et al., 2012) em cada arquivo resultante da cromatografia por HILIC e SCX, sendo que no final da análise foi realizada uma fusão de todos esses arquivos em um único através da ferramenta *SEPro Fusion*. Resumidamente, as identificações foram agrupadas pelo estado de carga do precursor ( $2+$  e  $\geq 3+$ ) resultando em dois subgrupos distintos. Para cada grupo, os valores do *XCorr*,  $\Delta CN$  e *Spectrum Count* resultantes do *Comet* foram usados para gerar um discriminador Bayesiano a fim de determinar uma função discriminante entre as identificações não-*decoys* e *decoys*. As identificações foram classificadas em uma ordem crescente de acordo com os valores do discriminador.

Os valores mínimos (*cut-off*) foram estabelecidos para que a lista final de identificações contivesse 1% de *decoys* (i.e., 1% FDR). Adicionalmente, aceitamos apenas sequências com seis ou mais aminoácidos. Os resultados foram pós-processados para aceitar PSMs com menos de 6 ppm, e um valor mínimo de *XCorr* de 1,90 foi imposto a proteínas que tivessem sido identificadas com um único espectro.

### 3.5.3.3 Quantificação relativa de proteínas

Os arquivos de MS1 foram extraídos com o *RawReader* e deconvoluídos usando as configurações padrão do Y.A.D.A. para a proteômica “*bottom-up*” (CARVALHO et al., 2009). Esse procedimento foi realizado para que se permita, em etapa adiante, a retirada dos espectros *multiplex*. Tais espectros produzem quantificações errôneas por misturar íons marcadores de peptídeos quase-isóbaros aos que co-fragmentam dentro da janela de isolamento no MS1. Assim, o Y.A.D.A aponta quais espectros devem ser descartados durante a análise quantitativa.

Posteriormente, o módulo *Isobaric Analyzer* do PL 4.0 foi utilizado para normalizar os sinais dos íons marcadores. Este procedimento é realizado independentemente, para cada análise de LC-MS/MS. Primeiramente, obtêm-se a soma dos sinais de cada íon marcador (i.e. 114,115,116,...); em seguida, para cada espectro, o valor normalizado é obtido dividindo o sinal de cada íon marcador pelo valor da respectiva somatória. A comparação proteômica foi realizada em três abordagens: tumor x margem de ressecção de 1 cm, tumor x margem de ressecção de 3 cm e tumor x margem de ressecção de 5 cm.

A análise comparativa considerou apenas proteínas identificadas com dois ou mais peptídeos que passassem no seguinte critério: mapear para apenas uma sequência proteica do

banco de dados, apresentar um *test-t* pareado com valor igual ou menor a 0,05 e um *ln fold change* com valor absoluto maior que 0,40. A partir disso, *p*-valores foram atribuídos a proteínas que estavam diferencialmente abundantes a partir de uma meta-análise dos *p*-valores dos peptídeos respectivamente mapeados de acordo com o método de *Stouffer* (STOUFFER et al., 1949). A equação de *Stouffer* permite ponderar *p*-valores; nós realizamos esse procedimento atribuindo um grau de confiança a cada peptídeo proporcional ao número de contagens espectrais que o mesmo obteve. O método de *Stouffer* encontra-se formalizado a seguir:

$$Z \sim \frac{\sum_{i=1}^k \omega_i Z_i}{\sqrt{\sum_{i=1}^k \omega_i^2}}$$

Onde *Z* (da estatística *Z-Score*) é aproximadamente igual ao somatório do *Z-Score* de cada peptídeo ponderado pelo número de contagens espectrais ( $\omega$ ) dividido pela raiz quadrada do somatório do quadrado das contagens espectrais.

Finalmente, após a obtenção dos dados de identificação e quantificação para as três abordagens de análises empregadas, foi realizada uma correlação entre os dados encontrados nessa segunda parte do trabalho com as informações clínicas e patológicas dos pacientes estudados.

#### 4. RESULTADOS E DISCUSSÃO

Neste trabalho, amostras de três pacientes foram estudadas, compreendendo a região tumoral e a margem de ressecção, em abordagens diferentes. As idades dos pacientes variaram de 41 a 82 anos, sendo dois do sexo masculino e um do feminino. Não foi detectado a presença da bactéria *H. Pylori* em nenhum. Dois pacientes foram submetidos a uma ressecção parcial e um à retirada total do estômago. O detalhamento das características clínicas encontra-se na Tabela 1. O paciente identificado como 1 forneceu amostras para a etapa A deste estudo enquanto as amostras dos pacientes 2 e 3 foram usadas na etapa B. As características histopatológicas estão descritas na Tabela 2, onde é possível verificar que tanto o local do câncer como o tipo histológico foram padronizados para todos os pacientes estudados, entre eles dois já vieram a óbito após o procedimento cirúrgico.

**Tabela 1** – Características clínicas dos pacientes estudados.

<b>Identificação</b>	<b>Idade</b>	<b>Sexo</b>	<b>Local da Coleta</b>	<b>Tipo de cirurgia</b>	<b>Histórico de carcinoma na família</b>	<b>Presença de <i>H. pylori</i></b>
Paciente 1	82	Masculino	Amazonas	Gastrectomia total	Sim. Câncer de intestino	Não
Paciente 2	41	Feminino	Rio de Janeiro	Gastrectomia subtotal	Não	Não
Paciente 3	66	Masculino	Rio de Janeiro	Gastrectomia subtotal	Sim. Câncer de fígado	Não

**Tabela 2** – Características histopatológicas dos pacientes estudados.

<b>Identificação</b>	<b>Local do câncer</b>	<b>Tipo histológico</b>	<b>Subtipo</b>	<b>Estadiamento</b>	<b>Óbito</b>
Paciente 1	Antro	Adenocarcinoma gástrico	Difuso	Borrmann III	Sim
Paciente 2	Antro	Adenocarcinoma gástrico	Difuso	Borrmann III	Sim
Paciente 3	Antro	Adenocarcinoma gástrico	Intestinal	Borrmann II	Não

#### 4.1 Etapa A – Análise do perfil proteico de tecidos das margens de ressecção e do tecido canceroso com a ferramenta computacional *Shotgun Imaging Analyzer* utilizando uma quantificação *label-free*.

Esta etapa compreende a avaliação do perfil proteico de um paciente o qual forneceu um fragmento de tecido, que foi verticalmente fracionada em dez partes, entre margem de ressecção e tecido canceroso, para análise pelo programa desenvolvido por nós, o *Shotgun Imaging Analyzer* (SIA). Este foi criado por nossa equipe possibilitando uma visualização do cenário proteômico dessas regiões estudadas correlacionada com identificação e quantificação de proteínas.

A leitura das concentrações de proteínas presente nos sobrenadantes das amostras estudadas encontra-se na Tabela 3. Já a leitura das concentrações dos peptídeos presente nas frações provenientes da cromatografia de troca catiônica forte encontra-se na Tabela 4.

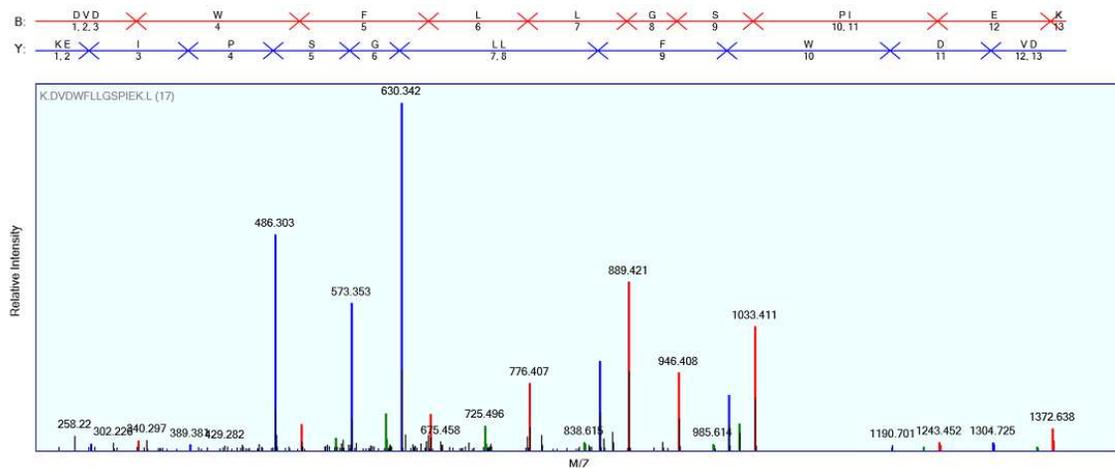
**Tabela 3** - Leitura da concentração de proteínas pelo ensaio fluorimétrico – *Qubit 2.0*® (*Invitrogen*).

Amostra	Concentração ( $\mu\text{g}/\mu\text{L}$ )	Diluição	Concentração do sobrenadante ( $\mu\text{g}/\mu\text{L}$ )
1	2,64	5	13,2
2	1,92	5	9,6
3	2,40	5	12,0
4	3,08	5	15,4
5	2,28	5	11,4
6	3,20	5	16,0
7	2,94	5	14,7
8	3,58	5	17,9
9	3,52	5	17,6
10	3,66	5	18,3

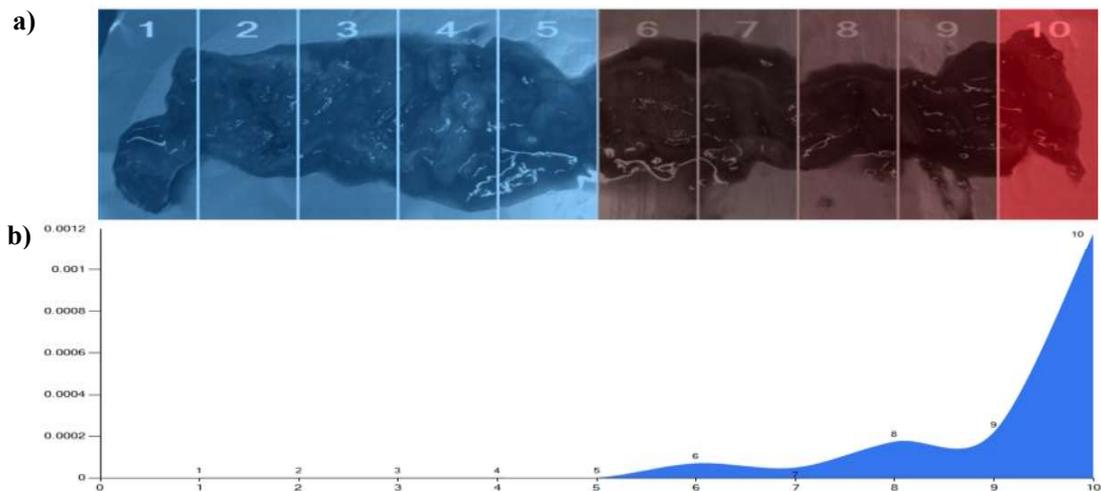
**Tabela 4** – Leitura da concentração dos peptídeos presente nas frações resultante dos passos salinos realizados.

Amostra	Fração 85 mM ( $\mu\text{g}/\mu\text{L}$ )	Fração 150 mM ( $\mu\text{g}/\mu\text{L}$ )	Fração 250 mM ( $\mu\text{g}/\mu\text{L}$ )	Fração 400 mM ( $\mu\text{g}/\mu\text{L}$ )
1	0,394	0,422	0,296	0,240
2	0,476	0,394	0,348	0,278
3	0,662	0,460	0,308	0,220
4	0,584	0,536	0,264	0,220
5	0,458	0,444	0,320	0,278
6	0,522	0,546	0,282	0,240
7	0,664	0,924	0,220	0,220
8	0,500	0,378	0,300	0,220
9	0,568	0,524	0,276	0,200
10	0,456	0,470	0,240	0,278

Em nossa análise foram identificadas 1936 proteínas (com redundância) e 997 por máxima parcimônia entre as dez seções, das quais apenas 651 possuíam 1 ou mais peptídeos únicos. Entre as proteínas identificadas, ressaltamos as *gastrokines*. Estas são descritas como proteínas abundantes especificamente no epitélio gástrico superficial. Em particular, apontamos a isoforma 2 da *gastrokine* (GKN2) codificada por um gene que foi mapeado no cromossomo 2, e que até o momento deste estudo não apresentava nenhuma evidência de existência por espectrometria de massas conforme apontado na ocasião pelo NeXtProt, sendo, portanto, uma “*missing protein*”. Nossos resultados contém a identificação de 4 peptídeos únicos para a GKN2, sendo um deles (*K.DVDWFLLGSPIEK.L*) exemplificado na Figura 15. Portanto, esse foi o primeiro relato da expressão da proteína GKN2 por espectrometria de massas; mais importante ainda, os dados quantitativos desta proteína permitiram a delimitação da região tumoral da não tumoral (Figura 16).



**Figura 15** - Espectro de massas sequencial do peptídeo único *K.DVDWFLGSPIEK.L* encontrado na sequência da proteína *gastrokine-2*.



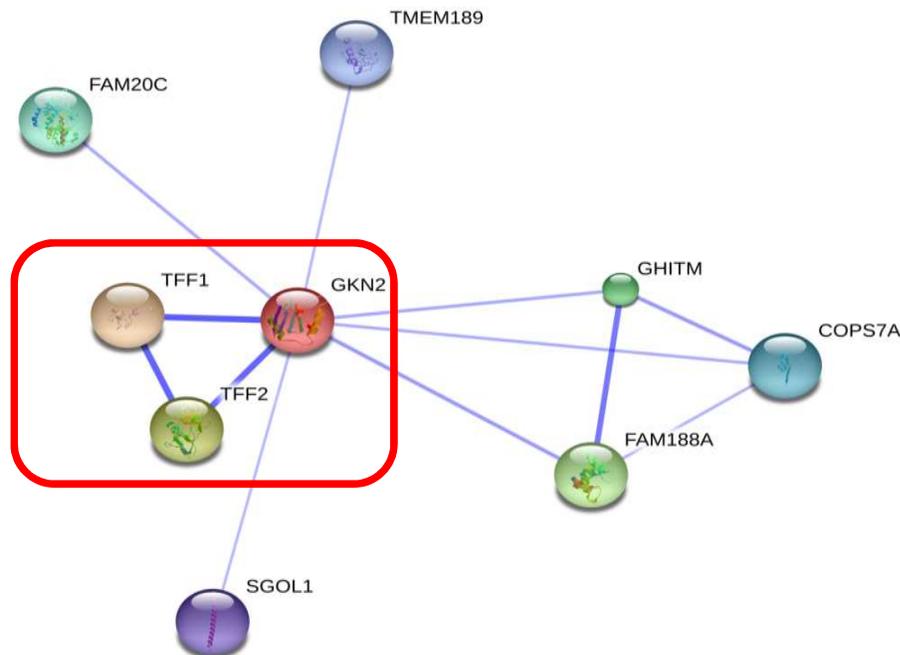
**Figura 16** – Resultado do *Shotgun Imaging Analyzer* para a proteína *NX\_Q86XP6*, isoforma 2 da *gastrokine*. **A)** Mapa de calor baseado na quantificação. **B)** O eixo y representa a distribuição normalizada de fatores de abundância espectrais (dNSAF) e o eixo x está alinhado com a imagem fornecida no *Shotgun Imaging* e os números correspondentes de cada seção do tecido. Um regressor foi usado para traçar o gráfico da área em azul. Seções 1-4 representa regiões tumorais, seções 5-7 foram classificadas como intermediárias e as seções 8-10 como margem de ressecção.

Na Figura 16 é possível demonstrar o aumento da expressão da GKN2 nas seções não pertencentes à região tumoral (5-10), o que está de acordo com os dados da literatura, que afirmam que as *gastrokines* têm uma menor abundância e algumas vezes estão até ausentes em amostras de câncer gástrico (OIEN et al., 2004). Em trabalho prévio, Moss e colaboradores utilizaram uma abordagem de perfil transcricional para descrever o decréscimo da expressão de GKN2 e sua relação com o prognóstico tumoral. Além disso, nesse mesmo trabalho reportaram a ausência da expressão destes transcritos em 85% dos cânceres do subtipo difuso e em 54% dos cânceres do subtipo intestinal ( $p$ -valor < 0,002) (MOSS et al., 2008). Outra isoforma da

*gastrokine* também encontrada no presente trabalho foi a GKN1, na qual o perfil de expressão, conforme apontado pelo *Shotgun imaging*, foi muito semelhante ao da GKN2 (Figura 16). Neste caso, a GKN1 foi identificada apenas na região não tumoral do tecido, mais especificamente nas seções 8-10. A função da isoforma GKN1, segundo a literatura, está relacionada em proteger e manter a integridade do epitélio gástrico, e uma deficiência dessa proteína pode resultar na instabilidade da mucosa gástrica o que conseqüentemente, pode contribuir para o desenvolvimento do câncer gástrico (MAO et al., 2012a; MOSS et al., 2008). Mais recentemente, Mao e colaboradores usando técnicas moleculares como, por exemplo, a reação em cadeia pela polimerase (PCR, do inglês, *polimerase chain reaction*), *Western blot*, e imunohistoquímica, verificaram um decréscimo progressivo da expressão da GKN1 a partir da mucosa normal até os tecidos gástricos com câncer, demonstrando que ela pode estar relacionada ao processo de carcinogênese (MAO et al., 2012a). No geral, a perda de expressão da GKN1 tem sido positivamente correlacionada com a perda da GKN2 (*p*-valor < 0,009), especialmente para o subtipo intestinal; e essa perda de expressão de ambas as proteínas tem sido associada a um prognóstico ruim (MOSS et al., 2008).

Outra proteína identificada e relacionada pela literatura com a expressão das *gastrokines*, é a isoforma 1 do fator *trefoil* (TFF1) codificada por um gene que foi mapeado no cromossomo 21 (NX\_P04155). Esta proteína é expressa principalmente nas células superficiais da mucosa gástrica, tendo sido encontrada em níveis elevados nessas células (MAO et al., 2012b). Estudos têm demonstrado a GKN2 em associação com a TFF1 como um heterodímero que provavelmente teria a função de manter a estabilidade da mucosa gástrica, contudo maiores detalhes do modo de ação deste heterodímero continuam desconhecidos (KOUZNETSOVA et al., 2007; MAO et al., 2012b; MENHENIOTT et al., 2013; WESTLEY et al., 2005). Além disso, ao realizar uma análise da interação da GKN2 através do programa STRING 10.0 (do inglês *Search Tool for the Retrieval of Interacting Genes/Proteins*, disponível em “<http://string-db.org/>”), é possível verificar a interação da GKN2 principalmente com as isoformas 1 e 2 do fator *trefoil* (TFF1 e TFF2, respectivamente) (Figura 17). No presente trabalho, a expressão da TFF1 foi encontrada em todas as seções do tecido, exceto na seção 3, que se situa dentro da região tumoral. Embora não seja possível chegar a uma conclusão estatística, ou seja, *p*-valor < 0,05, a média dos valores de dNIAF são maiores em média 1,7 na região não tumoral. Esse resultado encontrado para a expressão de TFF1 é coerente com a literatura, que afirma que esta deve estar ausente (ou menos abundante) na região tumoral; o que é bem similar a expressão das *gastrokines*, sendo que ambas proteínas têm sido descritas em alguns estudos como

possíveis supressores tumorais (FENG et al., 2014; KOUZNETSOVA et al., 2007; MAO et al., 2012b; MENHENIOTT et al., 2013; MOSS et al., 2008).

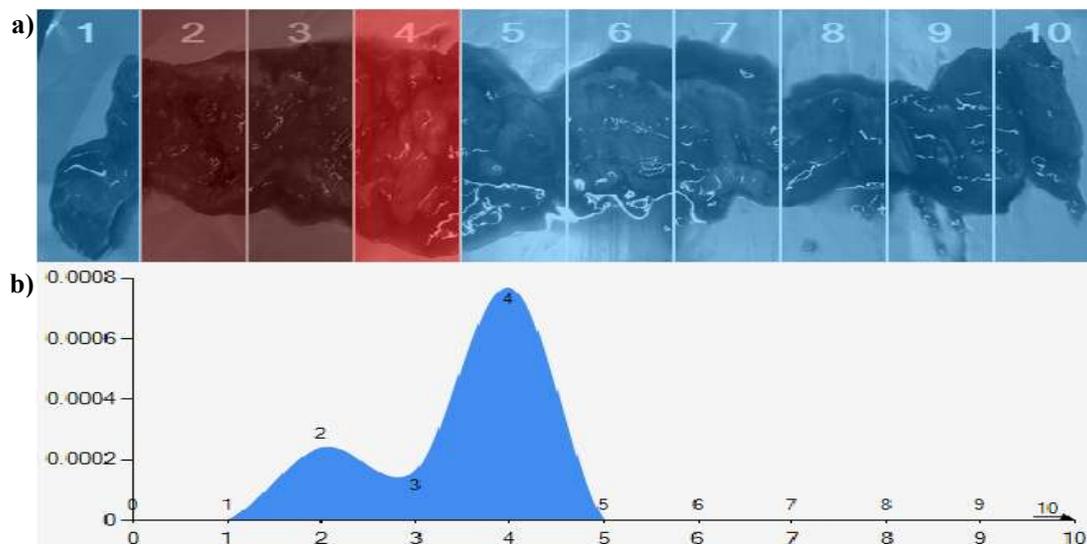


**Figura 17** – Rede de interações da proteína *gastrokine-2* (GKN2) através do programa STRING 10.0. As linhas roxas representam confiança de evidência na interação, sendo que quanto mais intenso é a cor e a linha mais alta é a confiança (*score* > 0,9). Fonte: “<http://string-db.org/>”.

Conforme descrito anteriormente, o *Shotgun Imaging Analyzer* pode ordenar as identificações proteicas de acordo com o índice *Golub*, comparando as diferenças de abundância de proteínas entre as regiões de interesse (ROIs). Dessa forma, é possível observar proteínas que possuem abundância predominante na região tumoral (seções 1-4) ou na margem. Entre as proteínas com os maiores índices de abundância na região tumoral, destacamos: a glutatona S-transferase Mu 3 e a lumican.

As glutatona S-transferases, cujo gene pertence ao cromossomo 1, constituem uma classe de proteínas que têm sido correlacionadas com o mecanismo de desintoxicação e proteção contra o desenvolvimento do câncer. Os genes desta família têm se mostrado altamente polimórficos, sendo expressos em alguns casos somente em 50% dos indivíduos, o que pode alterar a susceptibilidade a agentes cancerígenos e toxinas (KOČEVAR et al., 2012; MARTINS et al., 1998). Estudos prévios que buscaram correlacionar a expressão desses genes com o câncer gástrico foram realizados em populações com diferentes etnias, contudo não foi possível estabelecer uma associação direta entre ambos, necessitando de maiores investigações (KOČEVAR et al., 2012; YADAV et al., 2011). Com relação a expressão dessa proteína, ela tem sido encontrada com maior abundância em neuroblastomas e síndrome dos ovários

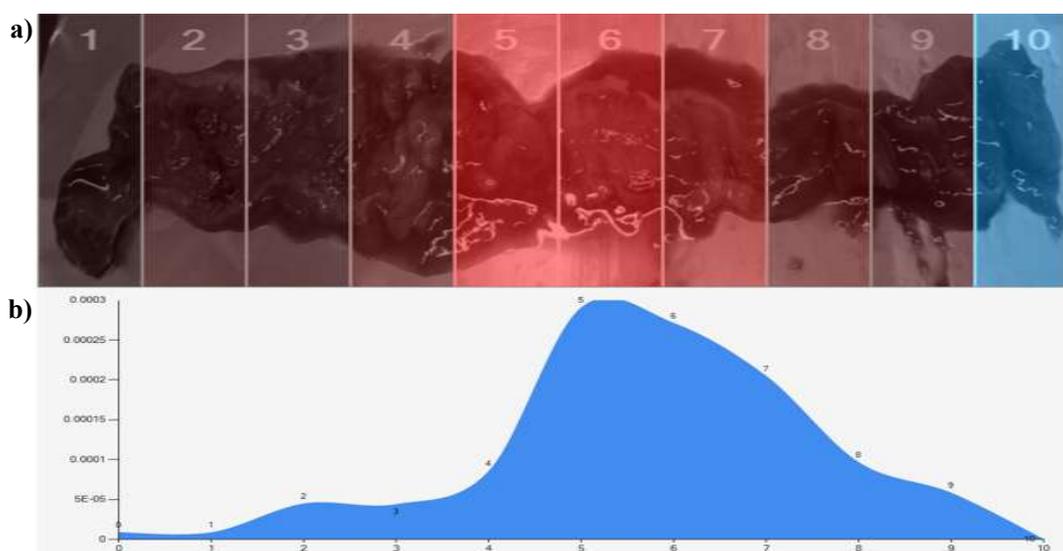
policísticos (CORTÓN et al., 2008; HE et al., 2005). Neste trabalho, a glutiona S-transferase Mu 3 foi encontrada com uma maior abundância na região tumoral (Figura 18).



**Figura 18** - Resultado do *Shotgun Imaging Analyzer* para a proteína NX\_P21266, isoforma 1 da glutiona S-transferase Mu 3. **A)** O mapa de calor baseado na quantificação pelo cromatograma extraído de íons e **B)** A representação gráfica da abundância dessa proteína através das seções do tecido.

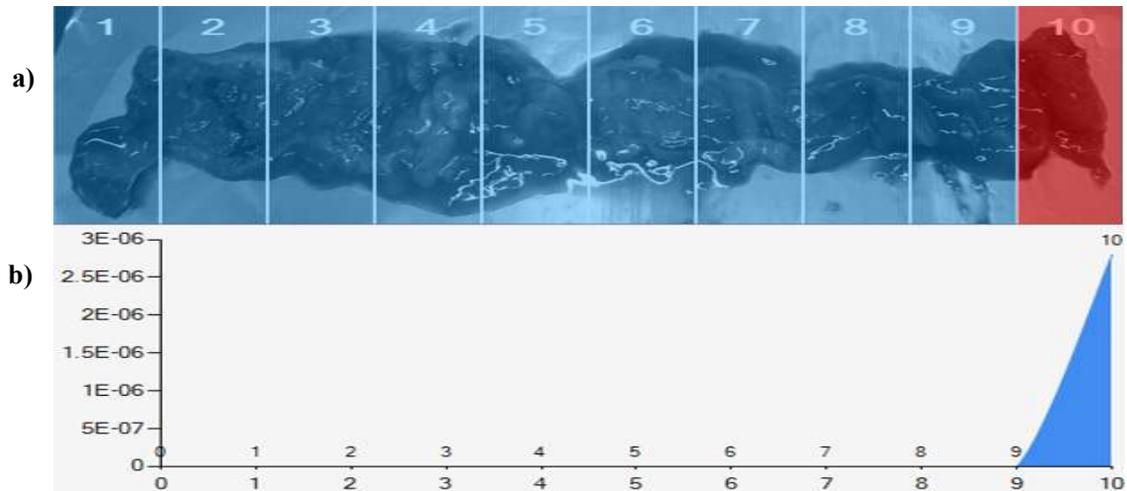
A lumican é uma das pequenas proteoglicanas ricas em leucina (SLRPs) com massa molecular em torno de 40 kDa (KAO et al., 2006). Trabalhos anteriores têm sugerido que ela desempenha um papel fundamental na adesão celular e que também pode servir como uma molécula reguladora de várias funções celulares, tais como migração celular e proliferação, apoptose, diferenciação e respostas inflamatórias (GU et al., 2011; KAO et al., 2006). Em cânceres, o aumento da abundância desta proteína tem sido correlacionado ao crescimento tumoral e a metástase em tumores pancreáticos e colorretal e, especialmente, em tumores da mama, nos quais a abundância dessa proteína também foi detectada em células estromais adjacentes às células tumorais (KAO et al., 2006; KELEMEN et al., 2008; NAITO, 2005). Segundo a literatura, níveis elevados na abundância dessa proteína podem ser associados com diferentes tipos patológicos de tumores (NAITO, 2005). A sua função na tumorigênese permanece indefinida, mas em estudo anterior verificou-se que o aumento na abundância da lumican suprime a transformação induzida pelas vias src e v-K-ras (YOSHIOKA et al., 2000). Tais observações podem ser explicadas pela possibilidade dos receptores de lumican mediarem diferentes vias de sinalização da transdução em tipos específicos de células, servindo de alicerce para possíveis infiltrações por macrófagos (FUNDERBURGH et al., 1997; NAITO, 2005). Em nossos resultados, a lumican foi identificada com maior abundância na região tumoral.

Exemplos de proteínas que apresentaram uma maior abundância na região intermediária (5-7) e na margem (8-10) são: a anexina e a E-caderina, respectivamente. A anexina é uma proteína de ligação de membrana, relacionada a uma grande variedade de funções celulares, como agregação de membrana, fagocitose, proliferação, apoptose e tumorigênese (CHENG et al., 2012). A baixa abundância dessa proteína tem sido associada a um prognóstico ruim em alguns tipos de câncer, como por exemplo, próstata, cabeça e pescoço, fígado, entre outros (DING et al., 2010; GARCIA PEDRERO et al., 2004; LIU et al., 2002). Nossos resultados demonstrados na Figura 19, estão de acordo com a literatura, onde resultados obtidos por abordagens complementares como, por exemplo, microchips de DNA, imunohistoquímica e “*immunoblot*”, enquadram a anexina como um potencial marcador para câncer gástrico (CHENG et al., 2012; EMOTO et al., 2001). Enquanto a função dessa proteína no câncer gástrico permanece desconhecida, existem cada vez mais evidências de que alterações na expressão da anexina e sua localização subcelular podem contribuir para o desenvolvimento e progressão da doença.



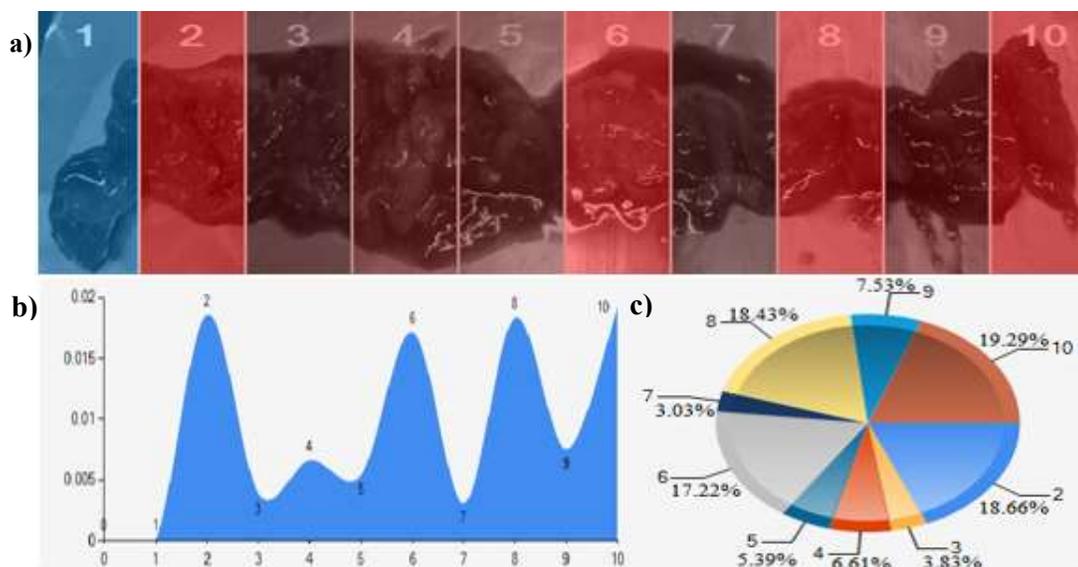
**Figura 19** - Resultado do *Shotgun Imaging Analyzer* da proteína NX\_P04083, isoforma 1 da anexina A1. **A)** O mapa de calor e **B)** Representação gráfica da abundância dessa proteína através das seções do tecido.

Com relação a E-caderina, essa foi identificada apenas na margem de ressecção (Figura 20), o que está alinhado com nossa publicação anterior (AQUINO et al., 2012). A literatura associa esta proteína com um papel importante na fase inicial do processo de tumorigênese, através da modulação da sinalização intracelular, podendo promover o crescimento tumoral e, eventualmente, a metástase (CHAN, 2006).



**Figura 20** - Resultado do *Shotgun Imaging Analyzer* da proteína NX\_P12830, isoforma 1 da E-caderina. **A)** O mapa de calor e **B)** Representação gráfica da abundância dessa proteína através das seções do tecido.

Entre as demais proteínas encontradas nessa etapa do estudo, uma classe de proteínas que chama a atenção pela sua abundância e que também já foram reportadas em trabalho anterior pelo nosso grupo, são as tropomiosinas (AQUINO et al., 2012). No presente trabalho, isoformas de tropomiosinas foram codificadas por genes mapeados nos cromossomos 1, 15, e 19, sendo que a isoforma 1 da tropomiosina cadeia alfa-4 apresentou expressão em aproximadamente todas as seções do tecido (Figura 21). De maneira geral, essa classe de proteínas é componente dos filamentos de actina que possuem um papel crítico na regulação da interação da actina e miosina, que, por sua vez, são moduladores importantes na dinâmica de adesão que determina a migração celular. Estudos anteriores têm demonstrado que diferentes isoformas de tropomiosinas desempenham funções fisiológicas distintas e seus perfis de expressão são profundamente alterados durante a progressão maligna de células cancerosas, visto que estas células modificam a organização do citoesqueleto e a morfologia celular (HE et al., 2004; STEHN et al., 2006). Em alguns tipos de cânceres como, por exemplo, o de mama, tem sido verificada uma baixa expressão dessas proteínas nos carcinomas primários; em contrapartida, a expressão é alta em tumores metastáticos (LEE et al., 2012). Resultados similares também foram reportados para o câncer gástrico na literatura (AQUINO et al., 2012; HE et al., 2004).



**Figura 21** - Resultado do *Shotgun Imaging Analyzer* para a proteína NX\_P67936, isoforma 1 da tropomiosina cadeia alpha-4. **A)** O mapa de calor, **B)** Representação gráfica e **C)** Percentagem da abundância dessa proteína nas seções do tecido.

Portanto, nesta primeira parte do trabalho, pode-se concluir algumas observações importantes sobre a metodologia usada. Primeiramente, sabe-se que o *MALDI-imaging*, se destaca pela sua resolução espacial notável. Por outro lado, abordagens por *electrospray* estão emergindo; como exemplo destacamos o trabalho liderado por L. Eberlin e R. Graham Cooks, o qual propõe o uso da dessorção de ionização por *electrospray* (*DESI*, do inglês, *Desorption electrospray ionization*) em uma abordagem lipidômica para o diagnóstico molecular intraoperativo de tumores cerebrais (EBERLIN et al., 2013). Este trabalho demonstra o potencial do *DESI-MS* para identificar o tipo histológico de tumores cerebrais, onde um classificador estatístico treinado é usado na análise dos dados por espectrometria de massas. Outro exemplo, é o trabalho liderado por Vladislav Petyuk e Richard Smith, no qual os autores utilizam a proteômica *shotgun* para gerar um mapa proteômico de cérebro de mamíferos, produzindo imagens de *shotgun* com um grau de identificação e resolução espacial similar a este trabalho (PETYUK et al., 2007, 2010).

Em conclusão, existem diversos relatos, similarmente aos supracitados, que demonstram o potencial dessas técnicas de panorâmica biomolecular e como podem impactar no tratamento médico. De maneira que nossa abordagem (*shotgun imaging*) pode trabalhar de maneira complementar para auxiliar no entendimento do câncer.

Os arquivos resultantes da análise pelo *Shotgun-Imaging Analyzer* (SIA) deste trabalho estão disponíveis em: “<http://proteomics.fiocruz.br/software/shotgunImaging/>”. Nele, é possível visualizar os mapas de calor para as proteínas identificadas nesse primeiro trabalho.

Os arquivos \*.RAW, juntamente com todos os arquivos do ProLuCID e do SEPro (\*.sqt) estão disponíveis em “<http://dm64.ioc.fiocruz.br/sia/>”.

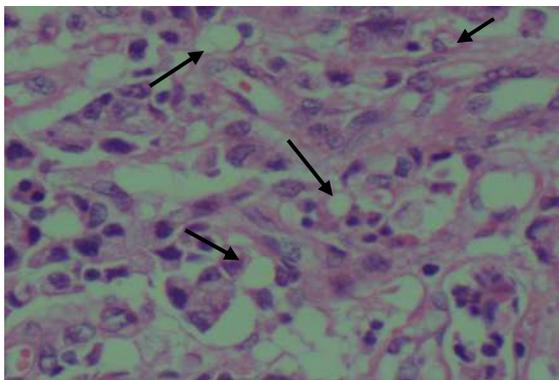
#### **4.2 Etapa B - Análise histopatológica e proteômica, empregando a quantificação por marcação com isóbaros – iTRAQ, de tecidos coletados da mesma região (tumor e margem de ressecção).**

Nesta segunda etapa realizamos um estudo com oito fragmentos de tecido (tumor e três margens de ressecção - 1, 3 e 5 cm) obtidos de dois pacientes diagnosticados com câncer gástrico. Cada fragmento foi avaliado por histopatologia e por proteômica.

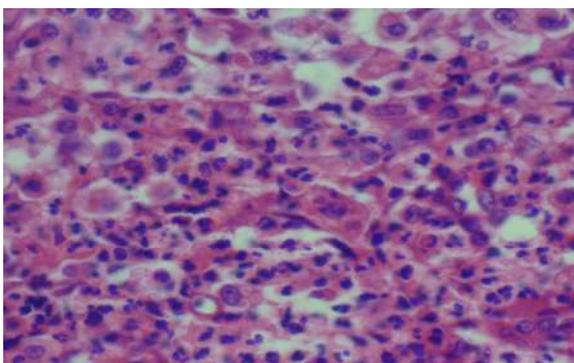
##### **4.2.1 Análise histopatológica**

A análise histopatológica foi realizada pelo serviço de anatomia patológica do Hospital Universitário Gaffrée e Guinle, sob a responsabilidade do Dr. Rodrigo Panno Basílio de Oliveira. Nesse, os quatro fragmentos dos dois pacientes foram analisados de forma independente do material coletado para o diagnóstico dos pacientes, não causando, portanto, prejuízo ao tratamento deles.

Os fragmentos foram submetidos a uma análise macroscópica e microscópica, respectivamente. Na análise macroscópica de ambos os pacientes, observou-se uma formação irregular, com características firme-elástica e pardo-clara, medindo aproximadamente 2,1 x 1,4 x 0,9 cm. Já na análise microscópica, para o paciente 2, no fragmento do tumor, observou-se um adenocarcinoma ulcerado tipo linite plástica pouco diferenciado com células em anel de sinete comprometendo toda a parede (Figura 22). Nas margens de ressecção desse paciente observou-se que a 1 cm havia ainda adenocarcinoma pouco diferenciado (Figura 23), porém nos fragmentos a 3 e 5 cm, estes apresentaram - se livre de neoplasia.

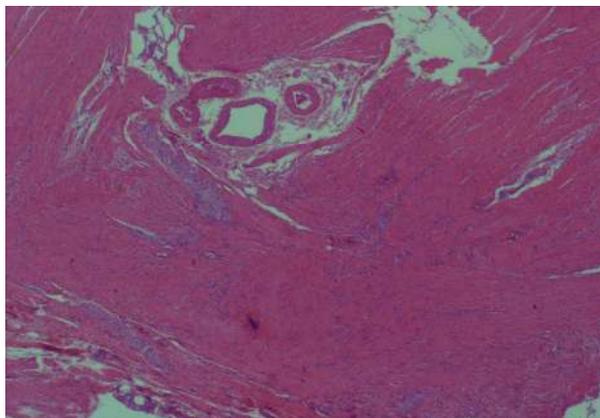


**Figura 22** - Corte histopatológico do fragmento tumoral do paciente dois corado com hematoxilina e eosina, onde é possível observar a presença de células em anéis de sinete indicados na figura.

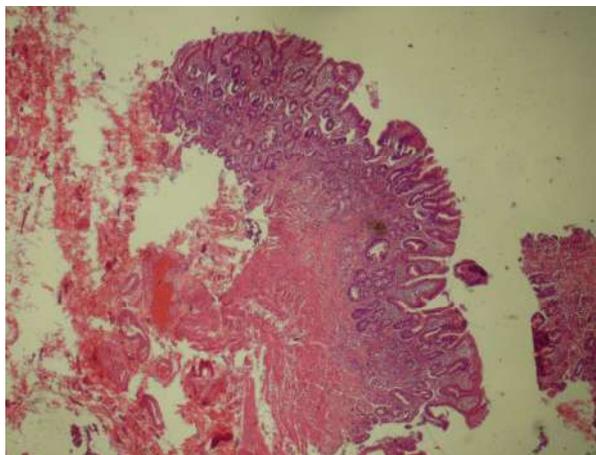


**Figura 23** - Corte histopatológico do fragmento da margem de ressecção a 1 cm do paciente dois corado com hematoxilina e eosina, onde verificou-se adenocarcinoma pouco diferenciado.

A análise microscópica do paciente 3 revelou no fragmento tumoral um adenocarcinoma pouco diferenciado infiltrando toda a espessura da parede. Nos fragmentos das margens de ressecção verificou-se que, a 1 cm ainda havia a presença de células que caracterizam o adenocarcinoma pouco diferenciado, porém a 3 cm (Figura 24) e 5 cm (Figura 25), não houve a observação de neoplasia, estando, portanto, estas livres de comprometimento.



**Figura 24** - Corte histopatológico do fragmento da margem de ressecção a 3 cm do paciente três corado com hematoxilina e eosina, livre de neoplasia.



**Figura 25** - Corte histopatológico do fragmento da margem de ressecção a 5 cm do paciente três corado com hematoxilina e eosina, livre de neoplasia.

Portanto, ambos os fragmentos tumorais foram classificados como adenocarcinoma pouco diferenciado. Os tecidos referentes a margem de ressecção a 1 cm estavam comprometidos com células cancerosas. Todavia, as margens de ressecção a 3 e 5 cm estavam livres de neoplasia, ou seja, eram margens de ressecção negativas de acordo com esta metodologia.

#### 4.2.2 Análise proteômica

A análise proteômica abordou os mesmos quatro fragmentos (tumor e três margens de ressecção – 1, 3 e 5 cm) dos pacientes da análise histopatológica descrita acima. Nesta, estes tecidos foram submetidos a uma quantificação por marcação com isóbaros, mediante a utilização do iTRAQ-8 *plex*. Após a marcação, dois tipos independentes de fracionamento cromatográfico foram realizados em paralelo, antes da análise de LC-MS/MS, onde originaram oito frações resultantes da cromatografia por HILIC e outras cinco da separação com SCX, totalizando treze frações cromatográficas ao final; cada fração foi posteriormente analisada por cromatografia em fase reversa diretamente acoplada ao LTQ Orbitrap Velos. A análise bioinformática considerou todas as frações em conjunto, permitindo obter um único resultado com FDR de 1%.

A leitura das concentrações de proteínas presente nas amostras estudadas nesta etapa encontra-se na Tabela 5. Já a leitura das concentrações dos peptídeos obtidos encontra-se descrita na Tabela 6.

**Tabela 5** – Leitura da concentração de proteínas pelo ensaio fluorimétrico – *Qubit 2.0*® (*Invitrogen*).

Amostra	Concentração ( $\mu\text{g}/\mu\text{L}$ )	Diluição	Concentração do sobrenadante ( $\mu\text{g}/\mu\text{L}$ )
1	3,36	5	16,80
2	4,48	1	4,48
3	3,40	1	3,40
4	3,26	1	3,26
5	3,06	5	15,30
6	3,14	1	3,14
7	3,30	5	16,50
8	5,20	1	5,20

**Tabela 6** - Leitura da concentração dos peptídeos antes da marcação pelo ensaio fluorimétrico – *Qubit 2.0*® (*Invitrogen*).

Amostra	Concentração ( $\mu\text{g}/\mu\text{L}$ )	Concentração ( $\mu\text{g}/\mu\text{L}$ )	Média ( $\mu\text{g}/\mu\text{L}$ )	Quantidade ( $\mu\text{g}$ )
	1° réplica	2° réplica		
1	0,572	0,610	0,591	28,37
2	1,470	1,440	1,455	69,84
3	1,810	1,862	1,836	88,13
4	1,780	1,730	1,755	84,24
5	0,342	0,498	0,42	20,16
6	1,920	1,860	1,89	90,72
7	0,768	0,792	0,78	37,44
8	1,530	1,550	1,54	73,92

Após a marcação dos peptídeos, a média da concentração na solução, inferida pelo ensaio fluorimétrico *Qubit 2.0*®, foi de 2,26  $\mu\text{g}/\mu\text{L}$ . Este resultado é um valor acima do esperado para a concentração final da amostra, o que pode ser resultante de algum contaminante que tenha permanecido na amostra, todavia este foi retirado posteriormente com os métodos cromatográficos utilizados nessa etapa (SCX e HILIC).

As leituras das concentrações dos peptídeos presentes nas frações resultantes do fracionamento cromatográfico por HILIC encontram-se na Tabela 7.

**Tabela 7** - Leitura da concentração dos peptídeos presente nas frações resultante da cromatografia líquida de interação hidrofílica (HILIC).

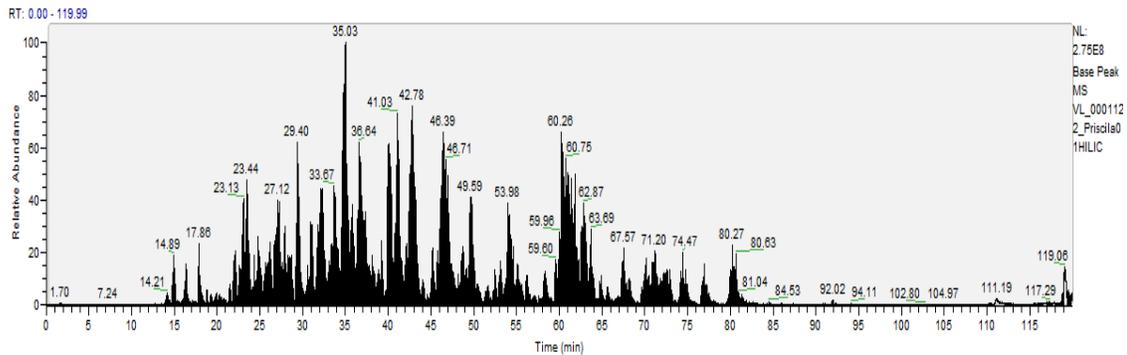
Fração	Concentração ( $\mu\text{g}/\mu\text{L}$ )
1	0,001
2	0,290
3	0,220
4	0,330
5	0,266
6	0,302
7	0,762
8	0,812

Em paralelo, também foi obtido as leituras das concentrações dos peptídeos presentes nas frações resultantes do fracionamento cromatográfico por SCX (Tabela 8).

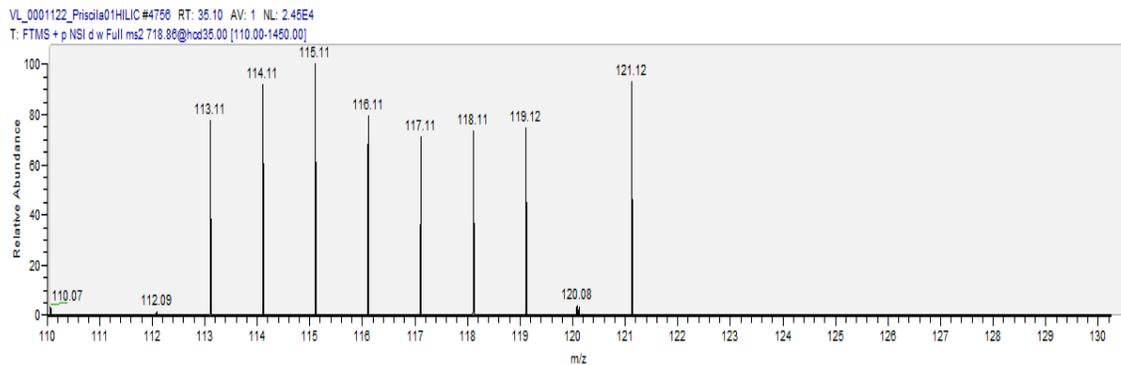
**Tabela 8** – Leitura da concentração dos peptídeos presente nas frações resultante da cromatografia líquida de troca catiônica forte (SCX).

Fração	Concentração ( $\mu\text{g}/\mu\text{L}$ )
9 – Fração 75 mM	1,38
10 – Fração 150 mM	1,95
11 – Fração 250 mM	0,620
12 – Fração 400 mM	0,220
13 - <i>Flow through</i>	0,220

Com relação aos cromatogramas de íons obtidos em nossa análise com o gradiente de 120 minutos, a Figura 26 ilustra um desses exemplos de uma das réplicas técnicas da amostra HILIC\_01. Já na Figura 27, os picos espectrais correspondentes aos íons marcadores no espectro de MS2, característico do iTRAQ-8 *plex*, são demonstrados; estes são utilizados para a posterior comparação relativa da abundância do peptídeo em questão, dentre as oito amostras biológicas.

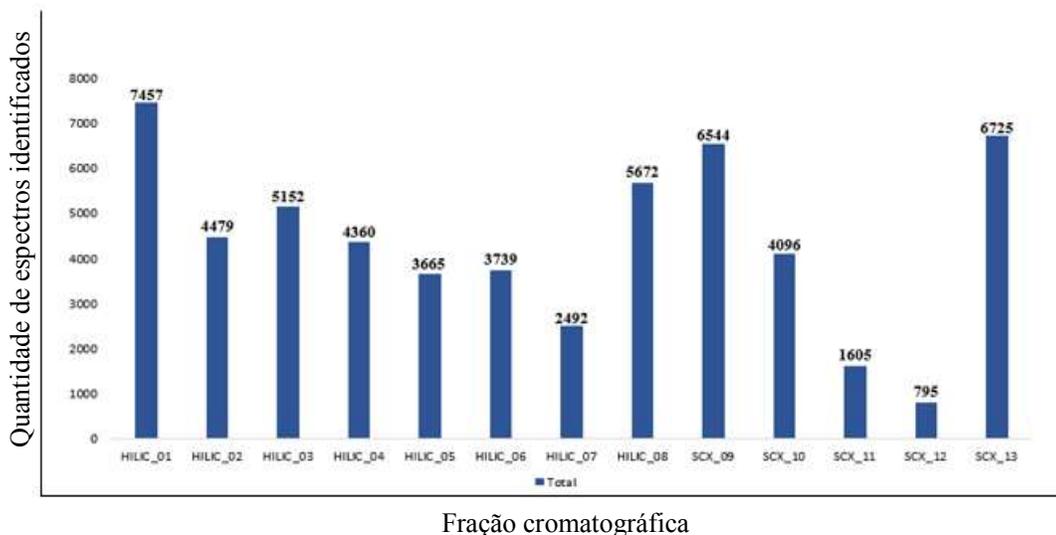


**Figura 26** – Perfil cromatográfico de uma das amostras analisadas (HILIC\_01) no estudo.



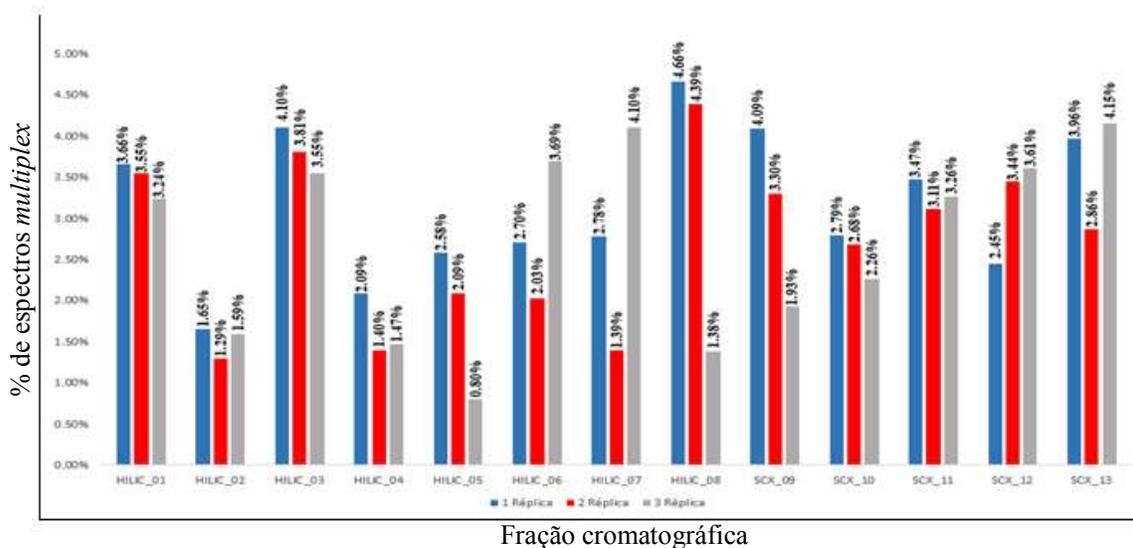
**Figura 27** – Um exemplo da região de baixa massa de um espectro de MS2 obtido neste estudo, enfatizando a presença dos íons marcadores característicos do iTRAQ-8 *plex*.

Nossos resultados listam 1314 proteínas por máxima parcimônia e 2468 (com redundância), das quais 920 possuíam pelo menos um peptídeo único. A Figura 28 demonstra a quantidade de espectros identificados de cada eluato; a primeira fração da cromatografia por HILIC e a 4ª fração da cromatografia por SCX foram as que apresentaram a maior e menor quantidade de espectros identificados, respectivamente.



**Figura 28** – Gráfico da quantidade de espectros identificados na triplicata técnica das treze frações cromatográficas.

Para permitir a comparação da quantificação relativa entre os peptídeos, utilizou-se o Y.A.D.A. para a eliminação de espectros *multiplex*. Ressaltamos que aproximadamente 5% dos espectros foram descartados devido a este procedimento (Figura 29).



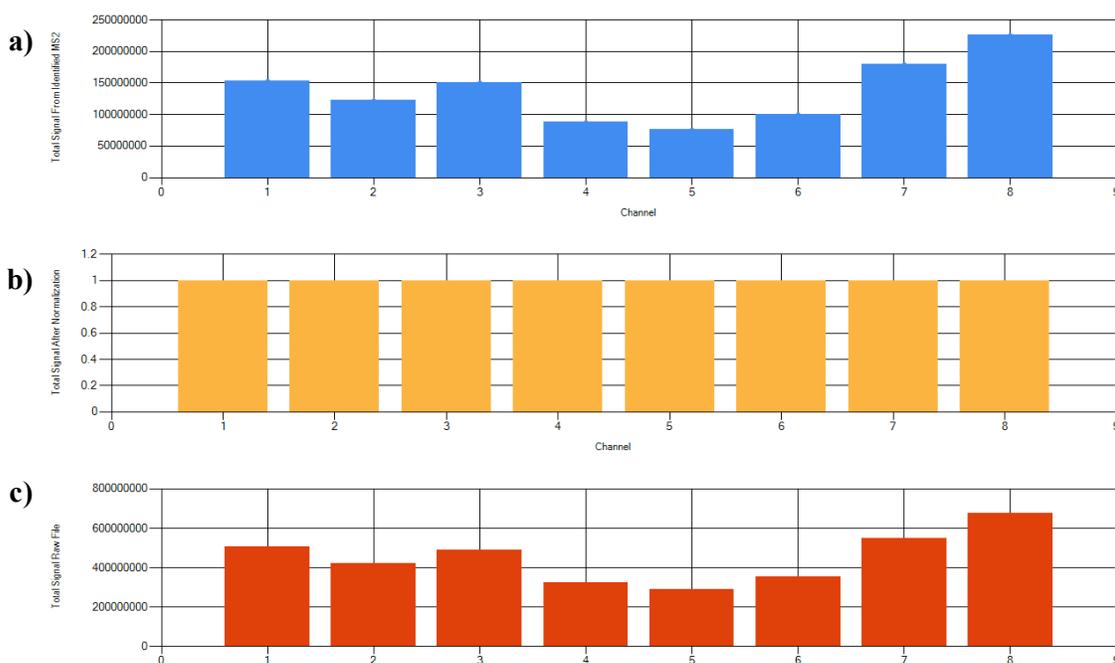
**Figura 29** – Representação gráfica da quantidade de espectro *multiplex* obtido em todas as amostras estudadas.

Após a deleção dos espectros *multiplex*, foi realizada a normalização da intensidade dos íons marcadores por amostra conforme descrito nos métodos. As comparações realizadas foram:

- Tumor x Margem de 1 cm

- Tumor x Margem de 3 cm
- Tumor x Margem de 5 cm

A normalização foi realizada dividindo o sinal do pico espectral de cada marcador pelo somatório do sinal para o marcador correspondente em uma determinada corrida. A Figura 30 exemplifica um gráfico de barras demonstrando o sinal para cada marcador antes e após a normalização (amarelo), considerando apenas os espectros identificados (azul) e todos os espectros (laranja). Observa-se que o somatório dos sinais de todos os espectros para uma determinada corrida (Figura 30c) possui um perfil bem semelhante quando comparado aos espectros identificados (Figura 30a). Vale ressaltar também que após a normalização, o somatório dos sinais para cada marcador é igual a 1 (Figura 30b).



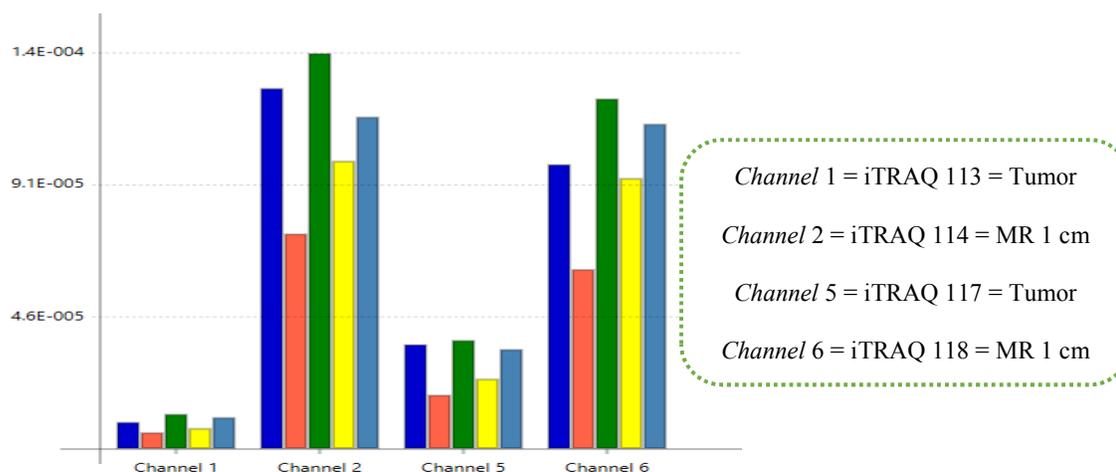
**Figura 30** – Representação gráfica da normalização realizada pelo módulo de quantificação para o iTRAQ do *PatternLab* (*Isobaric Analyzer*). **A)** Visualização dos sinais totais dos MS2 identificados nos oito canais; **B)** Visualização dos sinais totais depois da normalização nos oito canais; **C)** Visualização dos sinais totais do arquivo *RAW*.

A quantificação relativa dos peptídeos foi realizada de acordo com os parâmetros descritos no item 3.5.3.3 nas três abordagens descritas acima. Na Tabela 9, está descrito um resumo dos resultados obtidos para o total de peptídeos filtrados, a contagem espectral filtrada e o total de proteínas quantificadas em cada abordagem. Nessa, verifica-se que o total de proteínas diminui conforme a distância da margem de ressecção para a região tumoral aumenta.

**Tabela 9** – Resumo dos resultados obtidos da quantificação dos peptídeos para o iTRAQ do *PatternLab (Isobaric Analyzer)* nas três abordagens empregadas.

Abordagem	Total de peptídeos	Peptídeos filtrados	Contagem espectral filtrada	Total de proteínas filtradas
Tumor x MR1 cm	6342	846	13876	200
Tumor x MR3 cm	6342	888	11041	199
Tumor x MR5 cm	6342	817	12988	184

Como exemplo do cálculo de um dos *ln fold change* obtido em nossa análise, tem-se o peptídeo *R.IINEPTAAAIAYGLDKR.E* identificado mais de uma vez nas réplicas da amostra HILIC\_01. Neste, foram obtidos 5 espectros deste peptídeo. A Figura 31 permite observar, lado a lado, o sinal dos 4 marcadores decorrente dos 5 espectros. Claramente, observa-se que as barras correspondentes ao sinal dos íons marcadores oriundos da margem de ressecção 1 cm possuem uma intensidade maior do que os provenientes do tumor. O cálculo do *ln fold change* (-1,670), juntamente com os valores numéricos das intensidades para cada marcador estão disponíveis na Tabela 10. Tais valores obtidos do *ln fold change* também podem ser usados para calcular o *p-valor* desse peptídeo, aplicando o *test-t* pareado e considerando que a média esperada seria zero (Tabela 11).



**Figura 31** – Representação gráfica da abundância do peptídeo *R.IINEPTAAAIAYGLDKR.E* encontrado nas réplicas da amostra HILIC\_01, com 5 espectros entre os diferentes canais de marcação isóbara estudada. O sinal dos marcadores, decorrente de cada espectro, estão categorizados por cores.

**Tabela 10** – Dados referentes a quantificação do peptídeo *R.IINEPTAAAIAYGLDKR.E* encontrado mais de uma vez nas réplicas da amostra HILIC\_01 durante a eluição cromatográfica, os quais foram usados para calcular o valor do *ln fold change* (-1.670).

Amostra (ID)	iTRAQ 113	iTRAQ 114	iTRAQ 117	iTRAQ 118	Média $\frac{113}{117}$	Média $\frac{114}{118}$	Fold	ln Fold Change
HILIC_01 1º réplica	1.22E-05	1.37E-04	3.77E-05	1.21E-04	2.49E-05	1.29E-04	0.193	-1.64
HILIC_01 1º réplica	7.32E-06	9.95E-05	2.43E-05	9.36E-05	1.58E-05	9.66E-05	0.164	-1.81
HILIC_01 2º réplica	5.87E-06	7.44E-05	1.88E-05	6.22E-05	1.23E-05	6.83E-05	0.181	-1.71
HILIC_01 3º réplica	9.41E-06	1.25E-04	3.62E-05	9.83E-05	2.28E-05	1.11E-04	0.205	-1.59
HILIC_01 3º réplica	1.11E-05	1.15E-04	3.46E-05	1.12E-04	2.28E-05	1.14E-04	0.201	-1.60
<b>Total</b>								<b>-1.670</b>

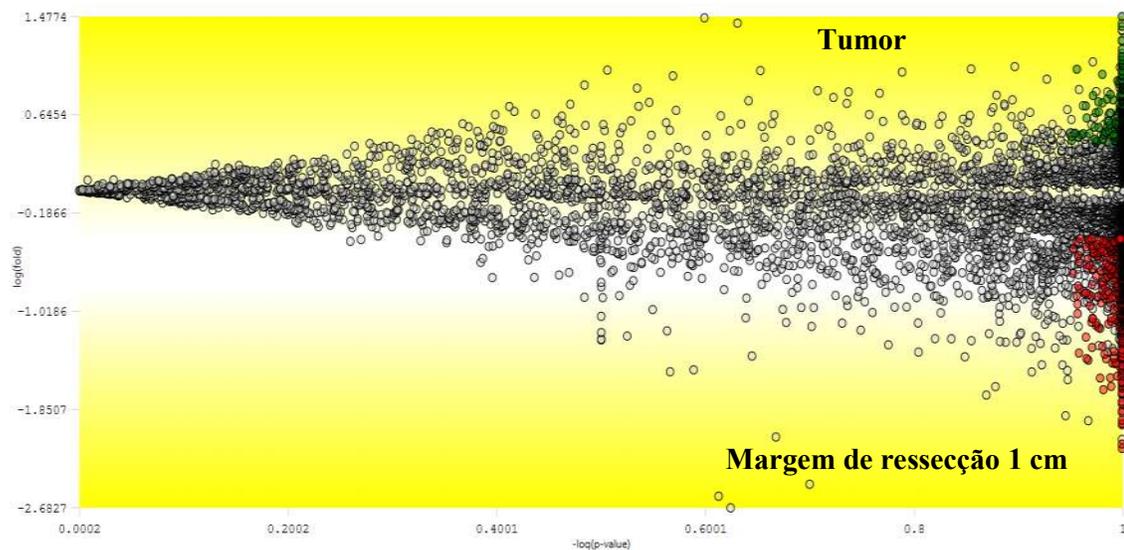
**Tabela 11** - Dados referentes aos valores de *ln fold change* obtidos na quantificação do peptídeo *R.IINEPTAAAIAYGLDKR.E* encontrado mais de uma vez nas réplicas da amostra HILIC\_01 durante a eluição cromatográfica, que foram usados para calcular o *p-valor* (0,001).

Amostra (ID)	ln Fold Change	p-valor
HILIC_01 1º réplica	-1.64	
HILIC_01 1º réplica	-1.81	
HILIC_01 2º réplica	-1.71	
HILIC_01 3º réplica	-1.59	
HILIC_01 3º réplica	-1.60	
<b>Total</b>		<b>0,001</b>

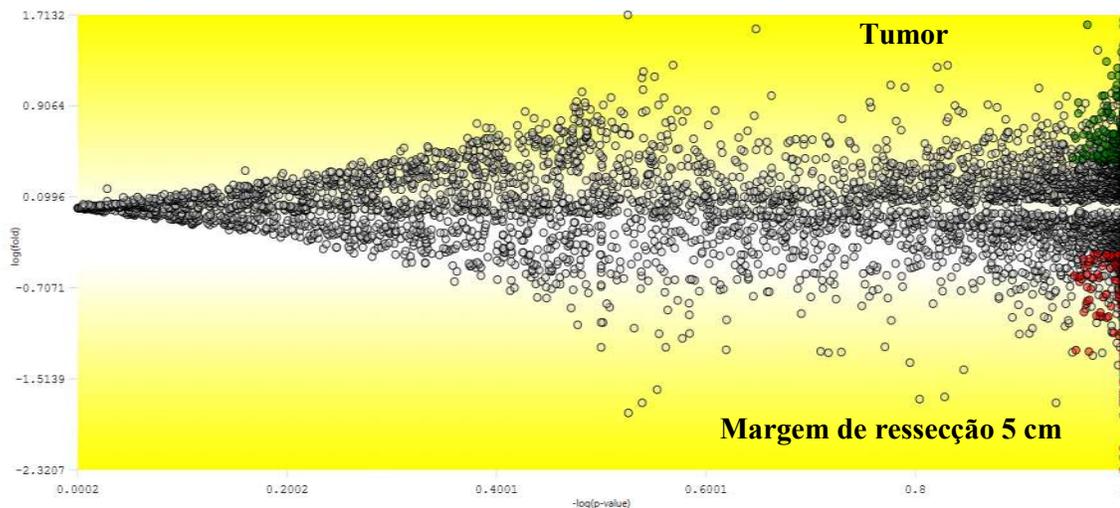
A Figura 32 demonstra a distribuição de peptídeos, de acordo com o *ln Fold Change* e probabilidade de estar diferencialmente abundante. Ao comparar a distribuição de peptídeos nos três cenários, verifica-se que, na margem de ressecção mais próxima à região tumoral (1 cm) há mais peptídeos com uma maior abundância na margem de ressecção do que na região

tumoral (Figura 32). Já na margem de ressecção com maior distância para a região tumoral (5 cm), observa-se que há menos peptídeos com maior abundância nessa mesma região (Figura 33). Um comportamento interessante é verificado ao comparar a margem de ressecção a 3 cm com a região tumoral (Figura 34), pois entre os três cenários, este é o que apresenta a maior diminuição na abundância de peptídeos na margem de ressecção; o que pode sinalizar uma região que apesar da distância para a região tumoral ser intermediária entre as estudadas, é a que apresenta o perfil de peptídeos mais diferencial.

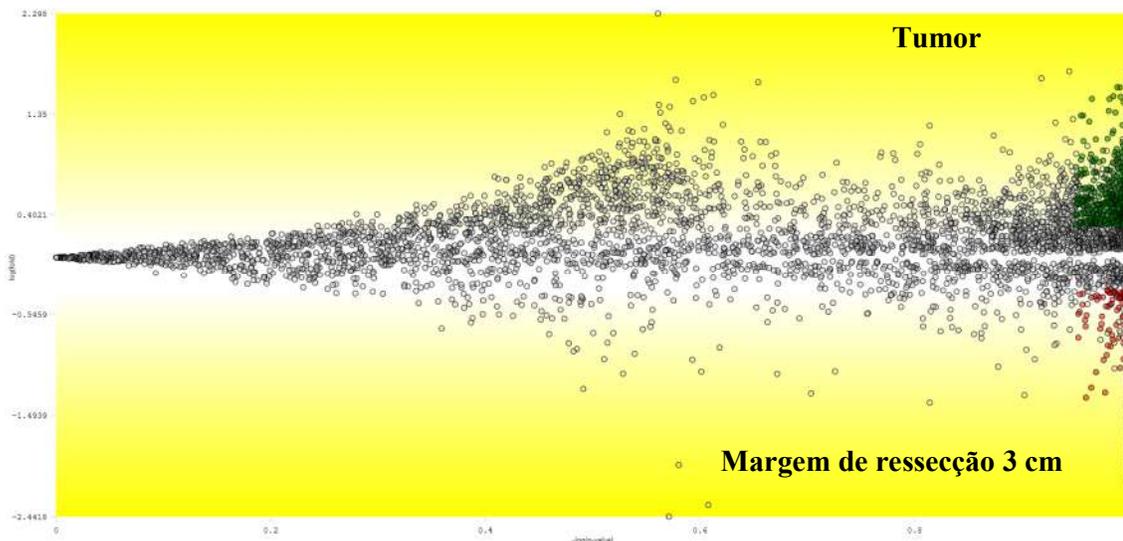
Para as proteínas, é observado algo semelhante ao comportamento dos peptídeos, no qual as proteínas da margem de ressecção mais próxima ao tumor (1 cm), estão mais abundantes na margem. Contudo, essa abundância diminuiu para a margem mais distante (5 cm) e um comportamento oposto foi observado, onde houve um aumento de proteínas mais abundantes no tumor. Além disso, o perfil diferencial da margem de ressecção a 3 cm perante as demais margens estudadas também foi visualizado a nível de proteínas; sendo a região com uma menor abundância na margem de ressecção comparando os três cenários, e também a que apresentou a maior abundância proteica na região tumoral.



**Figura 32** - Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 1 cm. Nessa, cada ponto representa um peptídeo que foi mapeado como um ponto no gráfico de acordo com  $-\text{Log}(p\text{-valor})$  (eixo das abscissas) e  $\text{Ln}(\text{Fold Change})$  (eixo das coordenadas). Os pontos em vermelho representam peptídeos que estão mais abundantes na margem de ressecção e que passaram no critério do  $p\text{-valor}$  e do  $\text{fold change}$ ; os pontos em verde demonstram os peptídeos que estão mais abundantes na região tumoral e que também passaram nos critérios estatísticos. Já os pontos em cinza significam peptídeos que não passaram pelos filtros estatísticos.



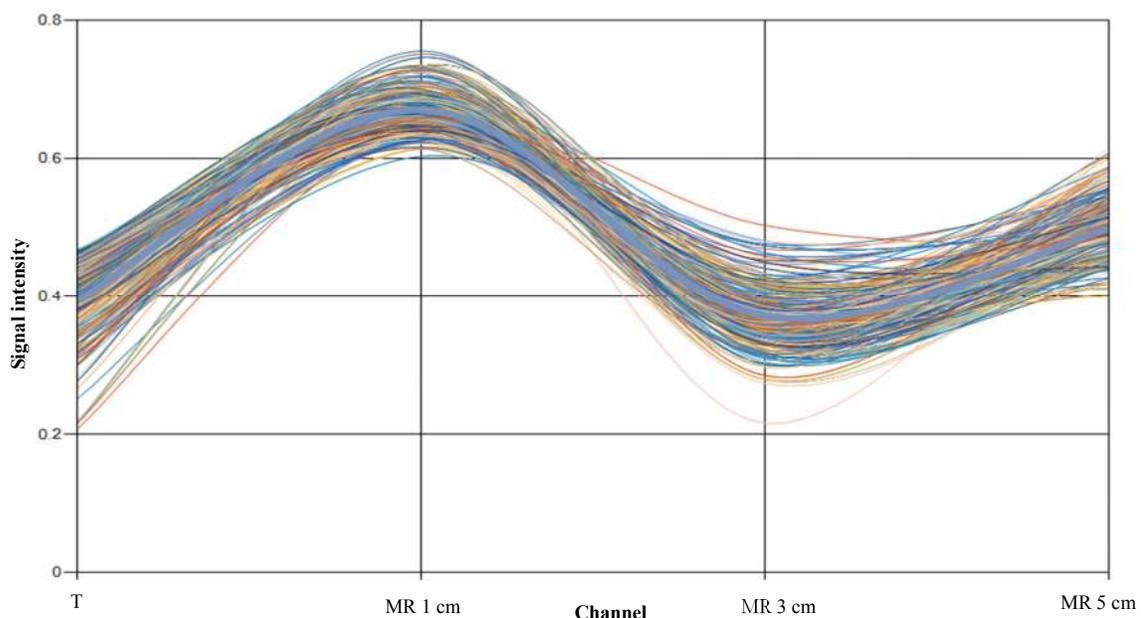
**Figura 33** – Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 5 cm.



**Figura 34** - Análise da distribuição de peptídeos observada na abordagem tumor x margem de ressecção a 3 cm.

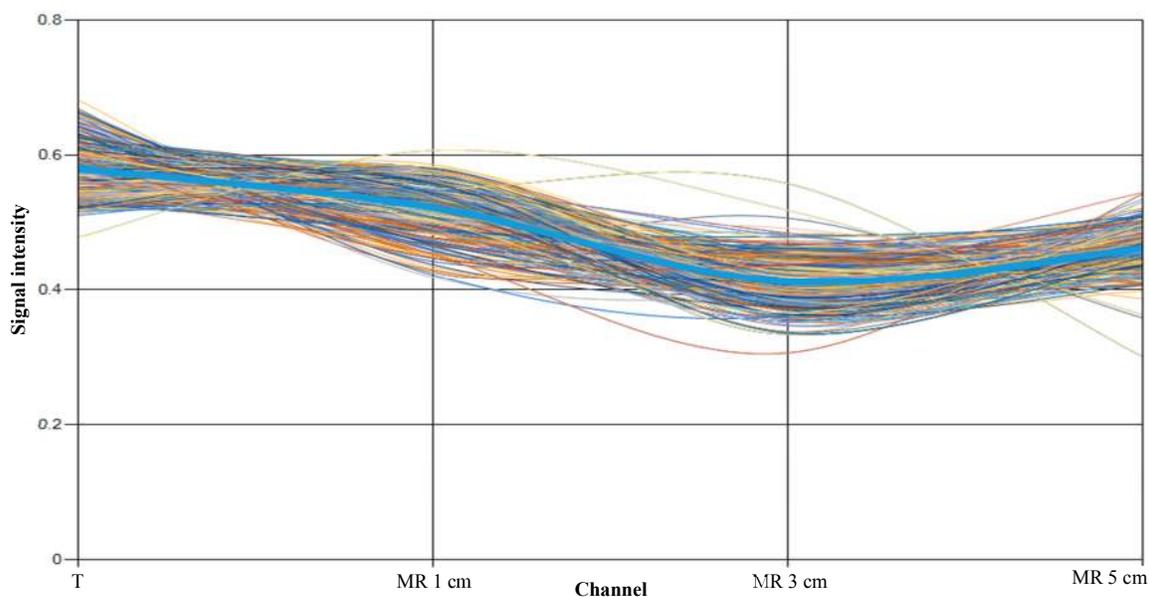
Além disso, uma análise levando em consideração a tendência de proteínas entre o tumor e as três margens de ressecção diferentes foi realizada com o módulo *TrendQuest* incluso no PL 4.0. Nessa análise, verificou-se que a margem de ressecção a 1 cm foi a que apresentou uma tendência mais diferencial entre os três tipos de margens de ressecção estudadas (Figura 35), o que é coerente com o resultado obtido pela análise histopatológica, na qual essa margem foi a única entre as demais que ainda apresentava a presença de células cancerosas. Portanto, ao nível molecular esperava-se dessa margem a 1 cm, uma diferença em seu perfil proteico perante as demais. Entre as proteínas identificadas na abordagem tumor x margem de 1 cm e que possuem correlação com processos de tumorigênese tem-se: a tropomiosina e a transgelina. A tropomiosina, em especial a tropomiosina cadeia alfa-4 e cadeia beta, é uma classe de proteínas

que também foi identificada e discutida na primeira etapa dessa tese (Etapa A) e em outro estudo do nosso grupo (AQUINO et al., 2012), onde a maior abundância dessa proteína tem sido correlacionada a processos cancerígenos, inclusive para o câncer gástrico (HE et al., 2004; LEE et al., 2012). A transgelina é uma proteína de ligação à actina que assim como a tropomiosina está relacionada à dinâmica do citoesqueleto por meio da estabilização dos filamentos de actina, e que tem sido descrita como envolvida no aumento do potencial metastático de um pequeno grupo de células tronco-tumorais (do inglês, *cancer stem cells* - CSCs), que podem ser entendidas como células capazes de iniciar tumores e promover a metástase (FISCHER et al., 2014a; LEE et al., 2010). Assim, esta proteína pode estar direta ou indiretamente envolvida em vários processos relacionados ao câncer, como por exemplo, a migração, proliferação, diferenciação ou apoptose (DVORAKOVA et al., 2014). Segundo a literatura, a transgelina foi inicialmente analisada como um supressor tumoral em alguns tipos de câncer (ASSINDER et al., 2009); contudo, em trabalhos mais recentes, a maior abundância dessa proteína tem sido associada a um papel pró-tumorigênese, em alguns tipos de câncer como de colorretal, pulmão, próstata, hepático e de mama (DVORAKOVA et al., 2014; LEE et al., 2010; RAO et al., 2015; WU et al., 2014). Nossos resultados estão alinhados aos trabalhos mais recentes por indicar esta proteína com maior abundância na região correlacionada a margem de ressecção positiva (1 cm).



**Figura 35** – Representação gráfica da análise da tendência de proteínas pelo *TrendQuest* entre a região tumoral e as três margens de ressecção estudadas. Nessa, cada linha representa uma proteína.

Proteínas com abundância relativamente uniforme nos quatro fragmentos estudados também foram encontradas, como por exemplo: as proteínas da família 14-3-3, que estão associadas a processos celulares relacionadas ao câncer como o ciclo celular e a apoptose, sendo considerados alvos terapêuticos importantes (FISCHER. et al., 2014b). Uma visualização da tendência dessas proteínas foi realizada também pelo o módulo *TrendQuest* e está mostrada na Figura 36.



**Figura 36** - Representação gráfica da tendência de proteínas relativamente uniforme nas quatro regiões estudadas pelo *TrendQuest*.

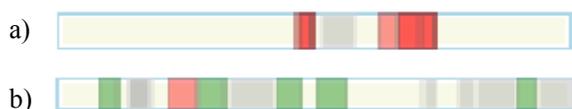
Uma heterogeneidade e abundância diferencial de certas proteínas na margem a 3 cm, também foi encontrada em nossos resultados como visualizado na distribuição de peptídeos acima (Figura 34), o que pode sinalizar que esta região seja uma área de transição do órgão, onde proteínas relacionadas a processos metabólicos e também a mobilidade celular se encontram alteradas. Alguns exemplos são: glutamato desidrogenase, quinona oxidoreductase, isocitrato desidrogenase, desmina, vinculina, calponina – 1, entre outros.

De maneira complementar, foi verificado também proteínas cuja abundância aumentava ou diminuía ao longo das três margens de ressecção estudadas. A profilina-1 e a caldesmona são exemplos de proteínas que à medida que houve o aumento da distância entre as margens de ressecção, a abundância dessas duas proteínas aumentaram na região tumoral. Interessantemente, ambas as proteínas assim como outras proteínas já citadas anteriormente, estão ligadas à mobilidade celular que possui um papel importante no desenvolvimento do câncer, onde o desencadeamento do processo de tumorigênese pode ser uma das consequências

da perda ou defeito de mecanismos que controlam a remodelação do citoesqueleto (LORENTE et al., 2014). Já com relação as proteínas cuja abundância aumentou na margem de ressecção com o aumento da distância para a região tumoral, tem-se a gastricsina e a lipase gástrica triacilglicerol; ambas proteínas relacionadas a processos digestivos na mucosa gástrica.

Outra característica interessante observada em nossos resultados é que as margens de ressecção mais distantes (3 e 5 cm), pelo resultado da análise patológica estavam livres de neoplasia. No entanto, conforme observado já em outro trabalho do nosso grupo (AQUINO et al., 2012), a nível molecular, estas margens já estão modificadas, onde proteínas correlacionadas a processos cancerígenos foram encontradas. Exemplos de tais proteínas são: a lumican e a glutathiona S-transferase (ambas já mencionadas e discutidas na etapa A), e a talina-1, que é uma proteína de adesão focal que regula as interações da integrina com a matriz extracelular, e tem sido encontrada com uma alta abundância em tumores de próstata, hepático e em melanoma gástrico primário (FISCHER et al., 2014a; KANAMORI et al., 2011).

A Figura 37a exemplifica a *gastrokine-1*, onde se observa uma diminuição na abundância de peptídeos na região tumoral. Entretanto, a Figura 37b exemplifica um caso particular (desmina), onde observa-se peptídeos aumentando, e outros diminuindo a abundância; todos pertencentes à mesma proteína. Na verdade, explicamos esta diminuição aparente por possivelmente ser alguma modificação pós-traducional no peptídeo; como a mesma altera a massa, o peptídeo deixa de ser identificado, resultando assim em uma falsa aparência de diminuição na abundância relativa. Esta metodologia, centrada em peptídeos, se destaca de forma pioneira por ser uma metodologia que pode apontar modificações pós-traducionais correlacionadas com a patologia.



**Figura 37** - Visualização da cobertura de peptídeos de proteínas *gastrokine-1* (A) e desmina (B). Nestas, as barras em vermelho indicam peptídeos que tiveram decréscimos em sua abundância; as barras verdes sinalizam peptídeos que tiveram um aumento na sua abundância e as barras cinzas representam peptídeos que não passaram nos critérios estatísticos.

Esta segunda parte do trabalho que combina análise histopatológica com a proteômica para o mesmo tecido, nos permite extrair várias conclusões inéditas. Primeiramente, as margens de ressecção mais distantes (3 e 5 cm), apesar de terem sido consideradas livres de neoplasias pela histopatologia, já apresentavam modificações moleculares de proteínas correlacionadas ao

processo de carcinogênese, reafirmando resultados já visualizados por nosso grupo em trabalho anterior com câncer gástrico (AQUINO et al., 2012). Além disso, ao analisar as três margens de ressecção diferentes, foi possível verificar proteínas cuja abundância aumentaram ou diminuíram à medida que houve um aumento na distância para a região tumoral, sinalizando possíveis processos metabólicos e de mobilidade celular que possam estar alterados em função da presença de células cancerosas no organismo desses pacientes. Verificou-se também as diferenças de abundância entre os peptídeos ao analisar a cobertura da proteína como um todo. Logo, diferenças moleculares significativas foram observadas tanto a nível de peptídeo como de proteínas entre os fragmentos analisados, fornecendo informações que podem ser úteis na compreensão da biologia do câncer gástrico e da alta recorrência dessa doença mesmo nos pacientes submetidos ao procedimento de ressecção do tumor.

Finalmente, as tabelas com as proteínas quantificadas e categorizadas por cromossomo nas três abordagens empregadas, seguindo os parâmetros descritos no item 3.5.3.3 para o módulo de quantificação para o iTRAQ do *PatternLab (Isobaric Analyzer)*, encontram-se nos anexos 5, 6 e 7.

## 5. CONCLUSÃO

A avaliação proteica das margens de ressecção de diferentes pacientes realizada no presente estudo apresentou características interessantes dessa região nas duas etapas realizadas neste trabalho.

Na primeira parte, verificou-se que a crescente importância do mapeamento biomolecular de alta performance correlacionado com uma disposição espacial tem demonstrado resultados que podem potencialmente impactar no tratamento médico. Neste contexto, nossa abordagem exibe uma resolução espacial quando comparada ao *MALDI-imaging*; porém, amparada pelo *MudPIT* e a quantificação por *label-free*, oferecendo como vantagem a identificação e quantificação de milhares de proteínas. O nosso algoritmo facilitou a identificação de uma *missing protein*, sendo esse trabalho o primeiro relato da expressão da *gastrokine-2* por espectrometria de massas (MS). Ainda com os dados de espectrometria de massas obtidos, foi possível demonstrar uma delimitação da expressão de proteínas ao longo das seções estudadas, onde pode verificar inclusive aquelas que estavam restritas a regiões de interesse (ROIs), mais especificamente a área tumoral e a margem de ressecção tais como, a glutathione S-transferase e a E-caderina, respectivamente. Além de que se possibilitou, de forma computacional, a rápida categorização dos dados em larga escala obtidos por proteômica aos devidos cromossomos alinhando aos interesses do projeto mundial C-HPP.

Já na segunda parte, demonstrou-se que as margens de ressecção mais distantes (3 e 5 cm), consideradas livres de neoplasias, apresentavam alterações moleculares de proteínas correlacionadas a oncogênese, o que está alinhado com resultados prévios publicados por nosso grupo em trabalho anterior com câncer gástrico. Tendências do comportamento proteico entre as diferentes regiões analisadas também foram estudadas, onde se observou proteínas cuja abundância aumentaram ou diminuíram à medida que houve um aumento na distância para a região tumoral, principalmente as proteínas correlacionadas a mobilidade celular; sinalizando que tais proteínas podem exercer papel importante no desenvolvimento do câncer. Demonstrou-se também que a nossa metodologia de análise de dados é pioneira no sentido de realizar análises centradas em peptídeos e, portanto, permitiu apontar modificações em peptídeos que possam estar correlacionados com a patologia em questão.

Ao comparar os tipos de abordagens proteômicas empregadas no trabalho com os dados clínicos e patológicos dos pacientes, verificaram-se diferenças importantes na abundância de certas proteínas ao analisar subtipos histológicos diferentes (*i.e.* difuso e o intestinal) assim como a localidade geográfica onde tais fragmentos foram coletados (*i.e.* Amazonas e Rio de

Janeiro). Logo, tais observações ressaltam a complexidade existente em estudar amostras de pacientes com câncer gástrico de regiões geográficas e subtipos histológicos distintos.

Portanto, com o presente trabalho, pode-se fazer uma análise mais detalhada da margem de ressecção, na qual modificações moleculares significativas foram observadas tanto a nível de peptídeo quanto de proteínas entre os diferentes tipos de tecidos estudados, fornecendo assim, informações que podem auxiliar na compreensão da biologia do câncer gástrico e também no processo de recorrência do tumor

## 6. COLABORAÇÕES E/ OU PRODUÇÃO CIENTÍFICA

O estudo proteômico com o câncer gástrico realizado durante o período do doutorado e apresentado nesta tese, resultou em participações em congressos nacionais e internacionais, assim como em um artigo publicado (Anexo 4), referente a primeira parte do trabalho (Parte A) e um artigo em andamento, relativo a segunda parte do estudo (Parte B), ambos como primeira autora.

Colaborei com o serviço de anatomia patológica do Hospital Universitário Clementino Fraga Filho (HUCFF-UFRJ), no qual auxiliei na análise proteômica de amostras de tecidos do tumor e as respectivas margens de ressecção cirúrgica, diferenciando-se do trabalho apresentado nessa tese, por investigar localizações anatômicas distintas do tumor no estômago e também por investigar a presença do vírus *Epstein-Barr* (EBV) nas amostras. Tal trabalho está em andamento.

Participei em diversos projetos envolvendo diferentes tipos de câncer, entre eles, o melanoma gástrico primário e o câncer colorretal. No trabalho com o melanoma gástrico primário e metástase meníngea resultante, em colaboração também com o serviço de anatomia patológica do Hospital Universitário Clementino Fraga Filho (HUCFF-UFRJ), auxiliei na análise proteômica de tecidos fixados em formalina e emblocados em parafina (FFPE) e na discussão dos dados que resultaram em duas publicações durante este período. Neste, foram identificadas diversas enzimas relacionadas a mudanças no metabolismo que estavam ligadas a produção de energia por elevadas taxas de glicólise através do efeito Wasburg (FISCHER et al., 2014) e demonstrou-se a eficiência da proteômica *shotgun* na avaliação da heterogeneidade celular de tecidos extraídos de lâminas, FFPE, com características patológicas diferentes (FISCHER et al., 2014b).

Já com relação ao trabalho com o câncer colorretal, colaborei com a Universidade Federal do Amazonas (UFAM), onde auxiliei com a análise proteômica associada à epigenética de tecidos da região tumoral e margens de ressecção, o qual também está em andamento.

Finalmente, tive a oportunidade de colaborar em projetos de proteômica computacional auxiliando no desenvolvimento de ferramentas pelo ponto de vista do usuário. Por exemplo, no *shotgun imaging analyzer* e no *PatternLab* 4.0, no qual o módulo *isobaric analyzer* foi desenvolvido durante este trabalho, expondo os primeiros resultados desta ferramenta e atendendo a diversas reivindicações do grupo durante o decorrer de ambos os projetos.

Tais projetos, assim como este de câncer gástrico, evidenciam os grandes desafios em analisar dados humanos devido à grande variabilidade entre os tipos diferentes de patologia estudada e as características únicas de cada paciente.

## 7. PERSPECTIVAS

Com os dados obtidos nas duas etapas dessa tese, alguns questionamentos sobre as possíveis mudanças moleculares que estão ocorrendo na região adjacente a região tumoral (margem de ressecção) foram investigadas, porém, novas perguntas surgiram. Portanto, como perspectivas desse trabalho, pretende-se aumentar o número de amostras que compreenda margens de ressecção a diferentes distâncias de subtipos histológicos diferentes, ou seja, estudar o subtipo difuso e intestinal a diferentes distâncias do tumor em um maior número de pacientes de regiões diferentes do Brasil; objetivando explorar as diferenças que podem ocorrer com a margem de ressecção devido ao subtipo histológico diferente e também a diferenciação na localização geográfica. Além disso, validar certas proteínas, como por exemplo: a *gastrokine* e a tropomiosina, que foram identificadas em ambas abordagens proteômicas utilizadas (*label-free* e marcação por isóbaros - iTRAQ) chamando a atenção pela sua abundância nas regiões estudadas.

Portanto, tais perspectivas visam ampliar o conhecimento sobre o perfil proteico da margem de ressecção para os diferentes subtipos histológicos de câncer gástrico e como essa região pode influenciar no processo de tumorigênese.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. AEBERSOLD, R.; BADER, G. D.; EDWARDS, A. M.; et al. The biology/disease-driven human proteome project (B/D-HPP): enabling protein research for the life sciences community. **Journal of Proteome Research**, v. 12, n. 1, p. 23–27. doi: 10.1021/pr301151m, 2013.
2. AEBERSOLD, R.; MANN, M. Mass spectrometry-based proteomics. **Nature**, v. 422, n. 6928, p. 198–207. doi: 10.1038/nature01511, 2003.
3. ALBERTS, B.; JOHNSON, A.; LEWIS, J.; et al. **Molecular Biology of the Cell**. 4th ed. Garland Science, 2002.
4. AQUINO, P. F.; FISCHER, J. S. G.; NEVES-FERREIRA, A. G. C.; et al. Are Gastric Cancer Resection Margin Proteomic Profiles More Similar to Those from Controls or Tumors? **Journal of Proteome Research**, v. 11, n. 12, p. 5836–5842. doi: 10.1021/pr300612x, 2012.
5. ASSINDER, S. J.; STANTON, J.-A. L.; PRASAD, P. D. Transgelin: an actin-binding protein and tumour suppressor. **The International Journal of Biochemistry & Cell Biology**, v. 41, n. 3, p. 482–486. doi: 10.1016/j.biocel.2008.02.011, 2009.
6. BANTSCHIEFF, M.; SCHIRLE, M.; SWEETMAN, G.; RICK, J.; KUSTER, B. Quantitative mass spectrometry in proteomics: a critical review. **Analytical and Bioanalytical Chemistry**, v. 389, n. 4, p. 1017–1031. doi: 10.1007/s00216-007-1486-6, 2007.
7. BONDARENKO, P. V.; CHELIUS, D.; SHALER, T. A. Identification and Relative Quantitation of Protein Mixtures by Enzymatic Digestion Followed by Capillary Reversed-Phase Liquid Chromatography–Tandem Mass Spectrometry. **Analytical Chemistry**, v. 74, n. 18, p. 4741–4749. doi: 10.1021/ac0256991, 2002.
8. CARVALHO, P. C.; FISCHER, J. S. G.; XU, T.; COCIORVA, D.; et al. Search Engine Processor: filtering and organizing PSMs. **Proteomics**, v. 12, n. 7, p. 944–949. doi: 10.1002/pmic.201100529, 2012.
9. CARVALHO, P. C.; FISCHER, J. S. G.; XU, T.; YATES, J. R.; BARBOSA, V. C. PatternLab: from mass spectra to label-free differential shotgun proteomics. **Current Protocols in Bioinformatics / Editorial Board, Andreas D. Baxevanis ... [et Al.]**, v. Chapter 13, p. Unit13.19. doi: 10.1002/0471250953.bi1319s40, 2012.

10. CARVALHO, P. C.; XU, T.; HAN, X.; et al. YADA: a tool for taking the most out of high-resolution spectra. **Bioinformatics (Oxford, England)**, v. 25, n. 20, p. 2734–2736. doi: 10.1093/bioinformatics/btp489, 2009.
11. CÉSAR, A. C. G.; SILVA, A. E.; TAJARA, E. H. Genetics and environmental factors in gastric carcinogenesis. **Arquivos de Gastroenterologia**, v. 39, n. 4, p. 253–259. doi: 10.1590/S0004-28032002000400009, 2002.
12. CHAN, A. O. O. E-cadherin in gastric cancer. **World Journal of Gastroenterology : WJG**, v. 12, n. 2, p. 199–203. doi: 10.3748/wjg.v12.i2.199, 2006.
13. CHENG, T.-Y.; WU, M.-S.; LIN, J.-T.; et al. Annexin A1 is associated with gastric cancer survival and promotes gastric cancer cell invasiveness through the formyl peptide receptor/extracellular signal-regulated kinase/integrin beta-1-binding protein 1 pathway. **Cancer**, v. 118, n. 23, p. 5757–5767. doi: 10.1002/cncr.27565, 2012.
14. CORTÓN, M.; BOTELLA-CARRETERO, J. I.; LÓPEZ, J. A.; et al. Proteomic analysis of human omental adipose tissue in the polycystic ovary syndrome using two-dimensional difference gel electrophoresis and mass spectrometry. **Human Reproduction (Oxford, England)**, v. 23, n. 3, p. 651–661. doi: 10.1093/humrep/dem380, 2008.
15. COTTRELL, J. S. Protein identification using MS/MS data. **Journal of Proteomics**, v. 74, n. 10, p. 1842–1851. doi: 10.1016/j.jprot.2011.05.014, 2011.
16. CRAIG, R.; BEAVIS, R. C. TANDEM: matching proteins with tandem mass spectra. **Bioinformatics (Oxford, England)**, v. 20, n. 9, p. 1466–1467. doi: 10.1093/bioinformatics/bth092, 2004.
17. DICKEN, B. J.; BIGAM, D. L.; CASS, C.; et al. Gastric adenocarcinoma: review and considerations for future directions. **Annals of Surgery**, v. 241, n. 1, p. 27–39, 2005.
18. DING, T.; YANG, L.; WANG, Y.; et al. Down-regulation of annexin II in prostate cancer is associated with Gleason score, recurrence, metastasis and poor prognosis. **Molecular Medicine Reports**, v. 3, n. 5, p. 781–787. doi: 10.3892/mmr.2010.332, 2010.
19. DVORAKOVA, M.; NENUTIL, R.; BOUCHAL, P. Transgelins, cytoskeletal proteins implicated in different aspects of cancer development. **Expert Review of Proteomics**, v. 11, n. 2, p. 149–165. doi: 10.1586/14789450.2014.860358, 2014.
20. EBERLIN, L. S.; NORTON, I.; ORRINGER, D.; et al. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. **Proceedings of the National Academy of Sciences**, v. 110, n. 5, p. 1611–1616. doi: 10.1073/pnas.1215687110, 2013.

21. EBERT, M. P. A.; NIEMEYER, D.; DEININGER, S. O.; et al. Identification and confirmation of increased fibrinopeptide a serum protein levels in gastric cancer sera by magnet bead assisted MALDI-TOF mass spectrometry. **Journal of Proteome Research**, v. 5, n. 9, p. 2152–2158. doi: 10.1021/pr060011c, 2006.
22. EIDHAMMER, I.; FLIKKA, K.; MARTENS, L.; MIKALSEN, S.-O. Protein, Proteome, and Proteomics. **Computational Methods for Mass Spectrometry Proteomics**. p.1–29. John Wiley & Sons, Ltd. Recuperado novembro 14, 2015, de <http://onlinelibrary.wiley.com/doi/10.1002/9780470724309.ch1/summary>, 2007.
23. EMOTO, K.; SAWADA, H.; YAMADA, Y.; et al. Annexin II overexpression is correlated with poor prognosis in human gastric carcinoma. **Anticancer Research**, v. 21, n. 2B, p. 1339–1345, 2001.
24. ENG, J. K.; JAHAN, T. A.; HOOPMANN, M. R. Comet: an open-source MS/MS sequence database search tool. **Proteomics**, v. 13, n. 1, p. 22–24. doi: 10.1002/pmic.201200439, 2013.
25. ENG, J. K.; MCCORMACK, A. L.; YATES, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. **Journal of the American Society for Mass Spectrometry**, v. 5, n. 11, p. 976–989. doi: 10.1016/1044-0305(94)80016-2, 1994.
26. Estimativa 2014 - Incidência de Câncer no Brasil. Recuperado agosto 2, 2014, de <http://www.inca.gov.br/estimativa/2014/estimativa-24042014.pdf>.
27. FENG, G.; ZHANG, Y.; YUAN, H.; et al. DNA methylation of trefoil factor 1 (TFF1) is associated with the tumorigenesis of gastric carcinoma. **Molecular Medicine Reports**, v. 9, n. 1, p. 109–117. doi: 10.3892/mmr.2013.1772, 2014.
28. FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M. Electrospray ionization for mass spectrometry of large biomolecules. **Science (New York, N.Y.)**, v. 246, n. 4926, p. 64–71, 1989.
29. FERLAY, J.; SOERJOMATARAM, I.; DIKSHIT, R.; et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. **International Journal of Cancer**, v. 136, n. 5, p. E359–386. doi: 10.1002/ijc.29210, 2015.
30. FERREIRA, C. R.; SARAIVA, S. A.; GARCIA, J. S. et al. Princípios e Aplicações da Espectrometria de Massas em Produção Animal. Recuperado novembro 14, 2015, de [http://invitrobrasil.com.br/pdf/artigos-publicados/22-Ferreira\\_CR\\_et\\_al\\_junho\\_2009.pdf](http://invitrobrasil.com.br/pdf/artigos-publicados/22-Ferreira_CR_et_al_junho_2009.pdf), 2009.

31. FISCHER, J. DE S. DA G.; CANEDO, N. H. S.; GONCALVES, K. M. DA S.; et al. Proteome analysis of formalin-fixed paraffin-embedded tissues from a primary gastric melanoma and its meningeal metastasis: a case report. **Current Topics in Medicinal Chemistry**, v. 14, n. 3, p. 382–387, 2014a.
32. FISCHER, J. S. G. **Proteoma diferencial do plasma de pacientes com glioblastoma recidivo e de células A172 em resposta ao álcool perílico**. Tese, Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2010.
33. FISCHER, J. S. G.; CARVALHO, P. C.; CANEDO, N. H. S.; et al. Applications of Formalin Fixed Paraffin-Embedded Tissue Proteomics in the Study of Cancer. **Austin Proteomics**, v. 1, n. 1, p. 1–7, 2014b.
34. FUNDERBURGH, J. L.; MITSCHLER, R. R.; FUNDERBURGH, M. L.; et al. Macrophage receptors for lumican. A corneal keratan sulfate proteoglycan. **Investigative Ophthalmology & Visual Science**, v. 38, n. 6, p. 1159–1167, 1997.
35. GARCIA PEDRERO, J. M.; FERNANDEZ, M. P.; MORGAN, R. O.; et al. Annexin A1 down-regulation in head and neck cancer is associated with epithelial differentiation status. **The American Journal of Pathology**, v. 164, n. 1, p. 73–79. doi: 10.1016/S0002-9440(10)63098-2, 2004.
36. GARTNER, L. P. **Tratado de Histologia Em Cores**. Elsevier Brasil, 2007.
37. GERBER, S. A.; RUSH, J.; STEMMAN, O.; KIRSCHNER, M. W.; GYGI, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. **Proceedings of the National Academy of Sciences of the United States of America**, v. 100, n. 12, p. 6940–6945. doi: 10.1073/pnas.0832254100, 2003.
38. GU, G.; CHENG, W.; YAO, C.; et al. Quantitative proteomics analysis by isobaric tags for relative and absolute quantitation identified Lumican as a potential marker for acute aortic dissection. **Journal of Biomedicine & Biotechnology**, v. 2011, p. 920763. doi: 10.1155/2011/920763, 2011.
39. HEBERT, A. S.; RICHARDS, A. L.; BAILEY, D. J.; et al. The one hour yeast proteome. **Molecular & cellular proteomics: MCP**, v. 13, n. 1, p. 339–347. doi: 10.1074/mcp.M113.034769, 2014.
40. HE, Q.-Y.; CHEUNG, Y. H.; LEUNG, S. Y.; et al. Diverse proteomic alterations in gastric adenocarcinoma. **Proteomics**, v. 4, n. 10, p. 3276–3287. doi: 10.1002/pmhc.200300916, 2004.

41. HE, Q.-Y.; ZHU, R.; REN, Y.; TAM, P. K. H.; CHIU, J.-F. Serological protein profiling of neuroblastoma by ProteinChip SELDI-TOF technology. **Journal of Cellular Biochemistry**, v. 95, n. 1, p. 165–172. doi: 10.1002/jcb.20417, 2005.
42. HILLENKAMP, F.; KARAS, M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. **Methods in Enzymology**, v. 193, p. 280–295, 1990.
43. HILLENKAMP, F.; KARAS, M.; BEAVIS, R. C.; CHAIT, B. T. Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry of Biopolymers. **Analytical Chemistry**, v. 63, n. 24, p. 1193A–1203A. doi: 10.1021/ac00024a716, 1991.
44. <http://dm64.ioc.fiocruz.br/sia/>. Recuperado novembro 16, 2015, de <http://dm64.ioc.fiocruz.br/sia/>.
45. <http://patternlabforproteomics.org/>. Recuperado novembro 16, 2015, de <http://patternlabforproteomics.org/>.
46. <http://proteomics.fiocruz.br/software/shotgunImaging/>. Recuperado novembro 16, 2015, de <http://proteomics.fiocruz.br/software/shotgunImaging/>.
47. <http://string-db.org/>. Recuperado novembro 16, 2015, de <http://string-db.org/>.
48. [http://tools.thermofisher.com/content/sfs/manuals/cms\\_049786.pdf](http://tools.thermofisher.com/content/sfs/manuals/cms_049786.pdf). Recuperado novembro 15, 2015, de [http://tools.thermofisher.com/content/sfs/manuals/cms\\_049786.pdf](http://tools.thermofisher.com/content/sfs/manuals/cms_049786.pdf).
49. <http://www.c-hpp.org/>. Recuperado maio 12, 2015, de <http://www.c-hpp.org/>.
50. <http://www.thermoscientific.com>. Recuperado maio 10, 2015, de <http://www.thermoscientific.com/en/home.html>.
51. <http://www.uniprot.org/>. Recuperado julho 4, 2015, de <http://www.uniprot.org/>.
52. HUBER, L. A. Is proteomics heading in the wrong direction? **Nature Reviews. Molecular Cell Biology**, v. 4, n. 1, p. 74–80. doi: 10.1038/nrm1007, 2003.
53. INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA. INCA - CÂNCER - Tipo - Estômago. Recuperado novembro 15, 2015a, de <http://www2.inca.gov.br/wps/wcm/connect/tiposdecancer/site/home/estomago/definicao>.

54. INSTITUTO NACIONAL DE CÂNCER JOSÉ ALENCAR GOMES DA SILVA. Estadiamento - INCA. Recuperado maio 8, 2015b, de [http://www.inca.gov.br/conteudo\\_view.asp?ID=54](http://www.inca.gov.br/conteudo_view.asp?ID=54).
55. INTERNATIONAL AGENCY FOR RESEARCH ON CANCER. Globocan 2012 - Home. Recuperado julho 3, 2015, de <http://globocan.iarc.fr/Default.aspx>.
56. JAPANESE GASTRIC CANCER ASSOCIATION. Japanese gastric cancer treatment guidelines 2010 (ver. 3). **Gastric Cancer: Official Journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association**, v. 14, n. 2, p. 113–123. doi: 10.1007/s10120-011-0042-4, 2011.
57. JIMENO, A.; HIDALGO, M. Molecular biomarkers: their increasing role in the diagnosis, characterization, and therapy guidance in pancreatic cancer. **Molecular Cancer Therapeutics**, v. 5, n. 4, p. 787–796. doi: 10.1158/1535-7163.MCT-06-0005, 2006.
58. KANAMORI, H.; KAWAKAMI, T.; EFFENDI, K.; et al. Identification by differential tissue proteome analysis of talin-1 as a novel molecular marker of progression of hepatocellular carcinoma. **Oncology**, v. 80, n. 5-6, p. 406–415. doi: 10.1159/000330734, 2011.
59. KANG, J.-Y.; SONG, S.-H.; YUN, J.; et al. Identification of long-range epigenetic silencing on chromosome 15q25 and its clinical implication in gastric cancer. **The American Journal of Pathology**, v. 185, n. 3, p. 666–678. doi: 10.1016/j.ajpath.2014.11.022, 2015.
60. KAO, W. W.-Y.; FUNDERBURGH, J. L.; XIA, Y.; LIU, C.-Y.; CONRAD, G. W. Focus on Molecules: Lumican. **Experimental eye research**, v. 82, n. 1, p. 3–4. doi: 10.1016/j.exer.2005.08.012, 2006.
61. KELEMEN, L. E.; COUCH, F. J.; AHMED, S.; et al. Genetic variation in stromal proteins decorin and lumican with breast cancer: investigations in two case-control studies. **Breast Cancer Research : BCR**, v. 10, n. 6, p. R98. doi: 10.1186/bcr2201, 2008.
62. KIM, M. G.; LEE, J.-H.; HA, T. K.; KWON, S. J. The distance of proximal resection margin dose not significantly influence on the prognosis of gastric cancer patients after curative resection. **Annals of Surgical Treatment and Research**, v. 87, n. 5, p. 223–231. doi: 10.4174/ast.2014.87.5.223, 2014.
63. KOČEVAR, N.; ODREMAN, F.; VINDIGNI, A.; GRAZIO, S. F.; KOMEL, R. Proteomic analysis of gastric cancer and immunoblot validation of potential biomarkers. **World Journal of Gastroenterology**, v. 18, n. 11, p. 1216–1228. doi: 10.3748/wjg.v18.i11.1216, 2012.

64. KOUZNETSOVA, I.; LAUBINGER, W.; KALBACHER, H.; et al. Biosynthesis of gastrokine-2 in the human gastric mucosa: restricted spatial expression along the antral gland axis and differential interaction with TFF1, TFF2 and mucins. **Cellular Physiology and Biochemistry: International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology**, v. 20, n. 6, p. 899–908. doi: 10.1159/000110450, 2007.
65. LANE, C. S. Mass spectrometry-based proteomics in the life sciences. **Cellular and molecular life sciences: CMLS**, v. 62, n. 7-8, p. 848–869. doi: 10.1007/s00018-005-5006-6, 2005.
66. LEE, E.-K.; HAN, G.-Y.; PARK, H. W.; SONG, Y.-J.; KIM, C.-W. Transgelin promotes migration and invasion of cancer stem cells. **Journal of Proteome Research**, v. 9, n. 10, p. 5108–5117. doi: 10.1021/pr100378z, 2010.
67. LEE, H.-H.; LIM, C.-A.; CHEONG, Y.-T.; SINGH, M.; GAM, L.-H. Comparison of Protein Expression Profiles of Different Stages of Lymph Nodes Metastasis in Breast Cancer. **International Journal of Biological Sciences**, v. 8, n. 3, p. 353–362. doi: 10.7150/ijbs.3157, 2012.
68. LEUNG, W. K.; WU, M.; KAKUGAWA, Y.; et al. Screening for gastric cancer in Asia: current evidence and practice. **The Lancet. Oncology**, v. 9, n. 3, p. 279–287. doi: 10.1016/S1470-2045(08)70072-X, 2008.
69. LIANG, Y.; DING, X.; WANG, X.; et al. Prognostic value of surgical margin status in gastric cancer patients. **ANZ journal of surgery**, v. 85, n. 9, p. 678–684. doi: 10.1111/ans.12515, 2015.
70. LIEBLER, D. C. **Introduction to Proteomics: Tools for the New Biology**. 2002 edition ed. Humana Press, 2001.
71. LIN, L.-L.; HUANG, H.-C.; JUAN, H.-F. Discovery of biomarkers for gastric cancer: a proteomics approach. **Journal of Proteomics**, v. 75, n. 11, p. 3081–3097. doi: 10.1016/j.jprot.2012.03.046, 2012.
72. LIU, S.-H.; LIN, C.-Y.; PENG, S.-Y.; et al. Down-regulation of annexin A10 in hepatocellular carcinoma is associated with vascular invasion, early recurrence, and poor prognosis in synergy with p53 mutation. **The American Journal of Pathology**, v. 160, n. 5, p. 1831–1837. doi: 10.1016/S0002-9440(10)61129-7, 2002.
73. LIU, W.; YANG, Q.; LIU, B.; ZHU, Z. Serum proteomics for gastric cancer. **Clinica Chimica Acta; International Journal of Clinical Chemistry**, v. 431, p. 179–184. doi: 10.1016/j.cca.2014.02.001, 2014.

74. LORENTE, G.; SYRIANI, E.; MORALES, M. Actin filaments at the leading edge of cancer cells are characterized by a high mobile fraction and turnover regulation by profilin I. **PloS One**, v. 9, n. 1, p. e85817. doi: 10.1371/journal.pone.0085817, 2014.
75. MACDONALD, F.; FORD, C.; CASSON, A. **Molecular Biology of Cancer**. 2 edition ed. London ; New York, N.Y: Taylor & Francis, 2004.
76. MAKAROV, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. **Analytical Chemistry**, v. 72, n. 6, p. 1156–1162. doi: 10.1021/ac991131p, 2000.
77. MAO, W.; CHEN, J.; PENG, T.-L.; et al. Downregulation of gastrokine-1 in gastric cancer tissues and restoration of its expression induced gastric cancer cells to apoptosis. **Journal of experimental & clinical cancer research: CR**, v. 31, p. 49. doi: 10.1186/1756-9966-31-49, 2012a.
78. MAO, W.; CHEN, J.; PENG, T.-L.; et al. Role of trefoil factor 1 in gastric cancer and relationship between trefoil factor 1 and gastrokine 1. **Oncology Reports**, v. 28, n. 4, p. 1257–1262. doi: 10.3892/or.2012.1939, 2012b.
79. MARRELLI, D.; DE STEFANO, A.; DE MANZONI, G.; et al. Prediction of recurrence after radical surgery for gastric cancer: a scoring system obtained from a prospective multicenter study. **Annals of Surgery**, v. 241, n. 2, p. 247–255, 2005.
80. MARTINS, G.; ALVES, M.; DIAS1, J.; et al. Glutathione S transferase mu polymorphism and gastric cancer in the Portuguese population. **Biomarkers**, v. 3, n. 6, p. 441–447. doi: 10.1080/135475098231084, 1998.
81. MENHENIOTT, T. R.; KURKLU, B.; GIRAUD, A. S. Gastrokines: stomach-specific proteins with putative homeostatic and tumor suppressor roles. **American Journal of Physiology. Gastrointestinal and Liver Physiology**, v. 304, n. 2, p. G109–121. doi: 10.1152/ajpgi.00374.2012, 2013.
82. MOSS, S. F.; LEE, J.-W.; SABO, E.; et al. Decreased expression of gastrokine 1 and the trefoil factor interacting protein TFIZ1/GKN2 in gastric cancer: influence of tumor histology and relationship to prognosis. **Clinical Cancer Research: An Official Journal of the American Association for Cancer Research**, v. 14, n. 13, p. 4161–4167. doi: 10.1158/1078-0432.CCR-07-4381, 2008.

83. MOTOYAMA, A.; YATES, J. R. Multidimensional LC Separations in Shotgun Proteomics. **Analytical Chemistry**, v. 80, n. 19, p. 7187–7193. doi: 10.1021/ac8013669, 2008.
84. NAITO, Z. Role of the small leucine-rich proteoglycan (SLRP) family in pathological lesions and cancer cell growth. **Journal of Nippon Medical School**, v. 72, n. 3, p. 137–145, 2005.
85. National Institute on Aging - Proteomics. Recuperado março 14, 2015, de <https://www.irp.nia.nih.gov/branches/lci/muproteomics.htm>.
86. NEOPTOLEMOS, J. P.; STOCKEN, D. D.; DUNN, J. A.; et al. Influence of resection margins on survival for patients with pancreatic cancer treated by adjuvant chemoradiation and/or chemotherapy in the ESPAC-1 randomized controlled trial. **Annals of Surgery**, v. 234, n. 6, p. 758–768, 2001.
87. NOGUEIRA, F. C. S. **Abordagens proteômicas no estudo do metabolismo de ácidos graxos e da deposição de alérgenos e toxinas em sementes maduras e em desenvolvimento de *Ricinus communis***. Tese, Rio de Janeiro: Universidade Federal do Rio de Janeiro, 2012.
88. NUSSBAUM, R. **Thompson & Thompson Genética Médica**. Elsevier Brasil, 2008.
89. OIEN, K. A.; MCGREGOR, F.; BUTLER, S.; et al. Gastroke 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation. **The Journal of Pathology**, v. 203, n. 3, p. 789–797. doi: 10.1002/path.1583, 2004.
90. OLSEN, J. V.; ONG, S.-E.; MANN, M. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. **Molecular & cellular proteomics: MCP**, v. 3, n. 6, p. 608–614. doi: 10.1074/mcp.T400003-MCP200, 2004.
91. OLSEN, J. V.; SCHWARTZ, J. C.; GRIEP-RAMING, J.; et al. A dual pressure linear ion trap Orbitrap instrument with very high sequencing speed. **Molecular & cellular proteomics: MCP**, v. 8, n. 12, p. 2759–2769. doi: 10.1074/mcp.M900375-MCP200, 2009.
92. ONG, S.-E.; MANN, M. Mass spectrometry-based proteomics turns quantitative. **Nature Chemical Biology**, v. 1, n. 5, p. 252–262. doi: 10.1038/nchembio736, 2005.
93. OPPENHEIMER, S. R.; MI, D.; SANDERS, M. E.; CAPRIOLI, R. M. Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. **Journal of Proteome Research**, v. 9, n. 5, p. 2182–2190. doi: 10.1021/pr900936z, 2010.

94. PERKINS, D. N.; PAPPIN, D. J.; CREASY, D. M.; COTTRELL, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. **Electrophoresis**, v. 20, n. 18, p. 3551–3567. doi: 10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2, 1999.
95. PETYUK, V. A.; QIAN, W.-J.; CHIN, M. H.; et al. Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography-mass spectrometry. **Genome Research**, v. 17, n. 3, p. 328–336. doi: 10.1101/gr.5799207, 2007.
96. PETYUK, V. A.; QIAN, W.-J.; SMITH, R. D.; SMITH, D. J. Mapping protein abundance patterns in the brain using voxelation combined with liquid chromatography and mass spectrometry. **Methods**, v. 50, n. 2, p. 77–84. doi: 10.1016/j.ymeth.2009.07.009, 2010.
97. RAO, D.; KIMLER, B. F.; NOTHNICK, W. B.; et al. Transgelin: a potentially useful diagnostic marker differentially expressed in triple-negative and non-triple-negative breast cancers. **Human Pathology**, v. 46, n. 6, p. 876–883. doi: 10.1016/j.humpath.2015.02.015, 2015.
98. ROSS, P. L.; HUANG, Y. N.; MARCHESE, J. N.; et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. **Molecular & cellular proteomics: MCP**, v. 3, n. 12, p. 1154–1169. doi: 10.1074/mcp.M400129-MCP200, 2004.
99. SCHULZ, W. A. **Molecular Biology of Human Cancers: An Advanced Student's Textbook**. 2007 edition ed. Dordrecht, The Netherlands ; Norwell, MA: Springer, 2005.
100. SHEN, J.-G.; CHEONG, J.-H.; HYUNG, W.-J.; et al. Influence of a microscopic positive proximal margin in the treatment of gastric adenocarcinoma of the cardia. **World Journal of Gastroenterology**, v. 12, n. 24, p. 3883–3886, 2006.
101. SHIN, D.; PARK, S.-S. Clinical importance and surgical decision-making regarding proximal resection margin for gastric cancer. **World Journal of Gastrointestinal Oncology**, v. 5, n. 1, p. 4–11. doi: 10.4251/wjgo.v5.i1.4, 2013.
102. SQUIRES III, M. H.; KOOBY, D. A.; POULTSIDES, G. A.; et al. Is it time to abandon the 5-cm margin rule during resection of distal gastric adenocarcinoma? A multi-institution study of the U.S. Gastric Cancer Collaborative. **Annals of Surgical Oncology**, v. 22, n. 4, p. 1243–1251. doi: 10.1245/s10434-014-4138-z, 2015.

103. STEHN, J. R.; SCHEVZOV, G.; O'NEILL, G. M.; GUNNING, P. W. Specialisation of the tropomyosin composition of actin filaments provides new potential targets for chemotherapy. **Current Cancer Drug Targets**, v. 6, n. 3, p. 245–256, 2006.
104. STOUFFER, S. A.; SUCHMAN, E. A.; DEVINNEY, L. C.; STAR, S. A.; WILLIAMS, R. M. **The American Soldier: Adjustment During Army Life**. Princeton University Press, 1949.
105. TANAKA, K.; WAKI, H.; IDO, Y.; et al. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. **Rapid Communications in Mass Spectrometry**, v. 2, n. 8, p. 151–153. doi: 10.1002/rcm.1290020802, 1988.
106. UCAR, E.; SEMERCI, E.; USTUN, H.; et al. Prognostic value of preoperative CEA, CA 19-9, CA 72-4, and AFP levels in gastric cancer. **Advances in Therapy**, v. 25, n. 10, p. 1075–1084. doi: 10.1007/s12325-008-0100-4, 2008.
107. WASHBURN, M. P.; WOLTERS, D.; YATES, J. R. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. **Nature Biotechnology**, v. 19, n. 3, p. 242–247. doi: 10.1038/85686, 2001.
108. WESTLEY, B. R.; GRIFFIN, S. M.; MAY, F. E. B. Interaction between TFF1, a gastric tumor suppressor trefoil protein, and TFIZ1, a brichos domain-containing protein with homology to SP-C. **Biochemistry**, v. 44, n. 22, p. 7967–7975. doi: 10.1021/bi047287n, 2005.
109. WILKINS, M. R.; PASQUALI, C.; APPEL, R. D.; et al. From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. **Bio/Technology (Nature Publishing Company)**, v. 14, n. 1, p. 61–65, 1996.
110. WU, W.; CHUNG, M. C. M. The gastric fluid proteome as a potential source of gastric cancer biomarkers. **Journal of Proteomics**, v. 90, p. 3–13. doi: 10.1016/j.jprot.2013.04.035, 2013.
111. WU, X.; DONG, L.; ZHANG, R.; YING, K.; SHEN, H. Transgelin overexpression in lung adenocarcinoma is associated with tumor progression. **International Journal of Molecular Medicine**, v. 34, n. 2, p. 585–591. doi: 10.3892/ijmm.2014.1805, 2014.
112. XU, T.; PARK, S. K.; VENABLE, J. D.; et al. ProLuCID: An improved SEQUEST-like algorithm with enhanced sensitivity and specificity. **Journal of Proteomics**, v. 129, p. 16–24. doi: 10.1016/j.jprot.2015.07.001, 2015.

113. YADAV, D.; CHANDRA, R.; SAXENA, R.; et al. Glutathione-S-transferase M1 and T1 genes and gastric cancer: a case control study in North Indian population. **Gene**, v. 487, n. 2, p. 166–169. doi: 10.1016/j.gene.2011.07.010, 2011.
114. YAKIREVICH, E.; RESNICK, M. B. Pathology of gastric cancer and its precursor lesions. **Gastroenterology Clinics of North America**, v. 42, n. 2, p. 261–284. doi: 10.1016/j.gtc.2013.01.004, 2013.
115. YANG, L. Incidence and mortality of gastric cancer in China. **World Journal of Gastroenterology**, v. 12, n. 1, p. 17–20, 2006.
116. YATES, J. R.; OSTERMAN, A. L. Introduction: Advances in Genomics and Proteomics. **Chemical Reviews**, v. 107, n. 8, p. 3363–3366. doi: 10.1021/cr068201u, 2007.
117. YATES, J. R.; RUSE, C. I.; NAKORCHEVSKY, A. Proteomics by mass spectrometry: approaches, advances, and applications. **Annual Review of Biomedical Engineering**, v. 11, p. 49–79. doi: 10.1146/annurev-bioeng-061008-124934, 2009.
118. YOSHIOKA, N.; INOUE, H.; NAKANISHI, K.; et al. Isolation of transformation suppressor genes by cDNA subtraction: lumican suppresses transformation induced by v-src and v-K-ras. **Journal of Virology**, v. 74, n. 2, p. 1008–1013, 2000.
119. ZHANG, X.; FANG, A.; RILEY, C. P.; et al. Multi-dimensional liquid chromatography in proteomics--a review. **Analytica Chimica Acta**, v. 664, n. 2, p. 101–113. doi: 10.1016/j.aca.2010.02.001, 2010.
120. ZHANG, Y.; FONSLow, B. R.; SHAN, B.; BAEK, M.-C.; YATES, J. R. Protein analysis by shotgun/bottom-up proteomics. **Chemical Reviews**, v. 113, n. 4, p. 2343–2394. doi: 10.1021/cr3003533, 2013.
121. ZHANG, Y.; WEN, Z.; WASHBURN, M. P.; FLORENS, L. Refinements to Label Free Proteome Quantitation: How to Deal with Peptides Shared by Multiple Proteins. **Analytical Chemistry**, v. 82, n. 6, p. 2272–2281. doi: 10.1021/ac9023999, 2010.

# ANEXOS

---

## ANEXOS

Anexo 1 - Parecer consubstanciado do comitê de ética (CEP) do Hospital Universitário Clementino Fraga Filho (HUCFF/UFRJ)

HOSPITAL UNIVERSITÁRIO  
CLEMENTINO FRAGA FILHO  
((HUCFF/ UFRJ))



## PARECER CONSUBSTANCIADO DO CEP

### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Estudo das proteínas presente em diversas regiões do câncer gástrico

**Pesquisador:** Priscila Ferreira de Aquino

**Área Temática:**

**Versão:** 4

**CAAE:** 10252913.5.0000.5257

**Instituição Proponente:** INSTITUTO DE QUIMICA

**Patrocinador Principal:** Financiamento Próprio

### DADOS DO PARECER

**Número do Parecer:** 632.982

**Data da Relatoria:** 08/05/2014

### Apresentação do Projeto:

Protocolo 074-123. Emenda recebida em 24.4.2014.

1. Informações básicas do Projeto de pesquisa, anexadas em 24.4.2014;
2. Carta de encaminhamento da emenda com justificativa, anexada em 24.4.2014;
3. Autorização de Instituição Co-participante, anexada em 18.4.2014;
4. Projeto de pesquisa em pdf contendo a emenda, anexado em 17.4.14;

### Objetivo da Pesquisa:

Não se aplica.

### Avaliação dos Riscos e Benefícios:

Não se aplica.

### Comentários e Considerações sobre a Pesquisa:

Não se aplica.

### Considerações sobre os Termos de apresentação obrigatória:

Não se Aplica.

### Recomendações:

Nenhuma.

**Endereço:** Rua Prof. Rodolpho Paulo Rocco N°255 Sala 01D-46

**Bairro:** Cidade Universitária **CEP:** 21.941-913

**UF:** RJ **Município:** RIO DE JANEIRO

**Telefone:** (21)3938-2480

**Fax:** (21)3938-2481

**E-mail:** cep@hucff.ufrj.br

HOSPITAL UNIVERSITÁRIO  
CLEMENTINO FRAGA FILHO  
((HUCFF/ UFRJ))



Continuação do Parecer: 632.982

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

**Considerações Finais a critério do CEP:**

RIO DE JANEIRO, 01 de Maio de 2014

---

**Assinador por:**  
**Carlos Alberto Guimarães**  
**(Coordenador)**

**Endereço:** Rua Prof. Rodolpho Paulo Rocco N°255 Sala 01D-46  
**Bairro:** Cidade Universitária **CEP:** 21.941-913  
**UF:** RJ **Município:** RIO DE JANEIRO  
**Telefone:** (21)3938-2480 **Fax:** (21)3938-2481 **E-mail:** cep@hucff.ufrj.br

Anexo 2 – Termo de Consentimento Livre e Esclarecido (TCLE) utilizado no estudo



Convidar-se você para participar do Projeto de Pesquisa com o título “**ESTUDO DAS PROTEÍNAS PRESENTES EM DIVERSAS REGIÕES DO CÂNCER GÁSTRICO**”. Este será realizado na Universidade Federal do Rio de Janeiro (UFRJ) em parcerias com outras instituições.

**Justificativa:** Tendo em vista o prognóstico pouco satisfatório do câncer gástrico em virtude de o diagnóstico quase sempre ser feito em estados avançados da doença, este projeto visa ajudar no desenvolvimento de novas metodologias para a detecção precoce do câncer, o que aumentaria a chance de curar esta doença.

**Obtenção do TCLE:** Em todos os casos o consentimento escrito dos voluntários (Termo de Consentimento Livre e Esclarecido) será obtido na enfermaria durante o período no qual o paciente estará internado para a realização do procedimento cirúrgico ao qual ele será posteriormente submetido sendo que a forma de abordagem será feita pela equipe de médicos participantes do projeto e a pesquisadora responsável mantendo a devida privacidade do mesmo. No Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA) a equipe médica responsável será o Dr. Marcus Vinicius Motta Valadão da Silva e o Dr. Antonio Carlos Accetta. No hospital Gaffrée e Guinle - UNIRIO será o Dr. Antonio Felipe Santa Maria Coquillard Ayres. Já na Fundação Centro de Controle em Oncologia do Estado do Amazonas – FCECON será o Dr. Sidney Raimundo Silva Chalub.

**Objetivos:** Identificar substâncias diferentes entre amostras da região tumoral e das regiões próximas a esta de pacientes com câncer gástrico, possibilitando assim uma maior compreensão do processo de cancerização das células.

**Procedimentos:** As análises em busca das substâncias referidas acima ocorrerão mediante técnicas apropriadas para este tipo de pesquisa, as quais entre elas têm-se o uso de ferramentas analíticas de alta resolução e precisão, e ferramentas de bioinformática para a identificação das proteínas encontradas entre as regiões estudadas.

**Possíveis riscos e/ou desconfortos:** O projeto proposto apresenta o risco inerente ao procedimento cirúrgico a qual o paciente já será submetido, assim como o risco de complicações pós-operatórias que o mesmo pode apresentar.

**Benefícios:** O presente projeto é uma das etapas iniciais do processo de descobertas, que visa no futuro auxiliar em um diagnóstico mais preciso nos estágios iniciais da doença.

**Informações e garantia de esclarecimento:** Durante todo o período de estudo você será acompanhado (a) pelo grupo de pesquisa que ficará à sua disposição para qualquer tipo de esclarecimentos que você necessite envolvendo informações mais aprofundadas do que está sendo feito no estudo. Para entrar em contato com os pesquisadores você poderá realizá-lo pelos telefones (92) 9148-3834 e 8803-9310 ou pelos e-mails: priscila\_aquino23@hotmail.com, gilberto@iq.ufrj.br e paulo@pcarvalho.com. Caso você tenha alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) do Hospital Universitário Clementino Fraga Filho/HUCFF/UFRJ – R. Prof. Rodolpho Paulo Rocco, n.º 255 – Cidade Universitária/Ilha do Fundão - Sala 01D-46/1º andar - pelo telefone (021) 2562-2480, de segunda a sexta-feira, das 8 às 15 horas, ou através do e-mail: cep@hucff.ufrj.br. E também com o Comitê de Ética de Pesquisa (CEP) do INCA, situado na Rua do Resende, 128 - sala 203, Centro. Telefones (021) 3207-4550 e 3207-4556, ou também pelo e-mail: cep@inca.gov.br.

**Compromisso de utilização do material biológico:** Ao autorizar esta coleta, suas amostras de tecido serão utilizadas nas análises descritas acima. O que não for usado não será guardado e sim descartado com o cuidado adequado para materiais biológicos humanos.

**Compromisso do pesquisador:** Os dados e os materiais coletados serão utilizados somente para esta pesquisa. Além disso, os resultados dos exames, bem como a avaliação de prontuário do paciente somente serão de responsabilidade dos pesquisadores envolvidos no projeto e dos profissionais que possam vir a ter relacionamento de atendimento e/ou de cuidados com o paciente e não será permitido acesso a terceiros (seguidores, empregadores, superiores hierárquicos), garantindo proteção contra qualquer tipo de discriminação e/ou preconceito. Assim como garante-se o direito a você (participante da pesquisa) de ser mantido atualizado sobre os resultados parciais da pesquisa, quando em estudos abertos ou de resultados que sejam do conhecimento dos pesquisadores.

**Transporte do material biológico:** O transporte seguirá as normas institucionais previamente estabelecidas.

**Questionário:** É assegurado ao participante (voluntário) da pesquisa o direito de recusar-se a responder as perguntas no questionário que venham a ocasionar constrangimentos de qualquer natureza.

**Autonomia, voluntariedade e retirada do estudo:** Você tem autonomia para decidir participar ou não do estudo, assim como também em retirar o seu consentimento de uso das informações coletadas em qualquer fase do estudo independente do motivo e sem prejuízo do atendimento que está recebendo.

**Privacidade e confidencialidade:** Os resultados da pesquisa serão analisados e publicados, mas sua identidade não será divulgada, sendo mantido sob sigilo absoluto. Informo também que todas as informações dos participantes serão utilizadas apenas para fins do presente estudo.

**ATENÇÃO:** A participação nesta pesquisa é voluntária, você não terá nenhuma despesa e também não receberá nenhum pagamento. Em caso de dano pessoal, diretamente causado pelos procedimentos ou tratamentos propostos neste estudo (nexo causal comprovado), o participante tem direito a tratamento médico na Instituição, bem como às indenizações legalmente estabelecidas.

### **Consentimento Pós-Infomação**

Acredito ter sido suficientemente informado a respeito das informações sobre o estudo acima citado que li ou que foram lidas para mim.

Eu discuti com o Dr. \_\_\_\_\_, sobre a minha decisão em participar nesse estudo. Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia de acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, sem penalidades ou prejuízos e sem a perda de atendimento nesta Instituição ou de qualquer benefício que eu possa ter adquirido. Eu receberei uma cópia desse Termo de Consentimento Livre e Esclarecido (TCLE) e a outra ficará com o pesquisador responsável por essa pesquisa. Além disso, estou ciente de que eu (ou meu representante legal) e o pesquisador responsável deveremos rubricar todas as folhas desse TCLE e assinar na última folha.

**Assinaturas:**

---

Nome do Sujeito da Pesquisa

---

Assinatura do Sujeito da Pesquisa

Data: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

Em caso de pacientes menores de 18 anos, analfabetos, semianalfabetos ou portadores de deficiência auditiva ou visual.

---

Nome do representante legal

---

Assinatura do representante legal

Data: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

**Pesquisadora responsável:**

---

Nome do Pesquisador Responsável

---

Assinatura do Pesquisador Responsável

Data: \_\_\_\_ / \_\_\_\_ / \_\_\_\_

**Anexo 3** – Questionário clínico utilizado no estudo

Nome:

Idade:

Sexo:

Histórico de doença na família:

Caso sim, qual a doença e relação de parentesco?

Possui alguma outra doença associada:

Caso sim, qual?

Está tomando medicamentos?

Caso sim, qual?

Passou por algum tratamento ou intervenção cirúrgica recentemente?

Caso sim, qual?

Possui dispepsia gástrica?

Estágio clínico do câncer:

Estágio cirúrgico do câncer:

Local do Câncer:

Tipo de Cirurgia:

Subtipo histológico:

Nº de inscrição:

Leito:

Telefone p/ contato:

Hepatite:

Diabetes:

Histórico de outro carcinoma? Qual?

Presença de *H. Pylori*:

Classificação de Lauren:

Diagnóstico da última endoscopia:

Local da biópsia:

Fuma:

Consumo de bebidas alcoólicas:

Total de amostras coletadas:

Paciente foi submetida à quimioterapia e/ou radioterapia:

Assinatura do Termo de Consentimento (TLCE):

Observações:

Data e Local:

# Exploring the Proteomic Landscape of a Gastric Cancer Biopsy with the Shotgun Imaging Analyzer

Priscila Ferreira Aquino,<sup>†,‡,¶</sup> Diogo Borges Lima,<sup>‡,¶</sup> Juliana de Saldanha da Gama Fischer,<sup>‡</sup> Rafael Donadelli Melani,<sup>†</sup> Fabio C. S. Nogueira,<sup>†</sup> Sidney R. S. Chalub,<sup>§</sup> Elzalina R. Soares,<sup>⊥</sup> Valmir C. Barbosa,<sup>||</sup> Gilberto B. Domont,<sup>†,\*</sup> and Paulo C. Carvalho<sup>‡,\*</sup>

<sup>†</sup>Proteomics Unit, Rio de Janeiro Proteomics Network, Department of Biochemistry, Federal University of Rio de Janeiro, Rio de Janeiro 21941-909, Brazil

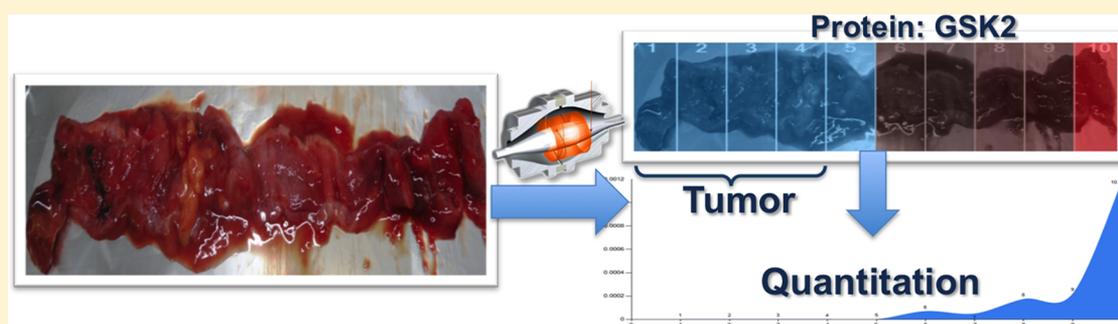
<sup>‡</sup>Laboratory for Proteomics and Protein Engineering, Carlos Chagas Institute, Fiocruz Paraná 81350-010, Brazil

<sup>§</sup>Departamento de Cirurgia Digestiva, Universidade Federal do Amazonas, Amazonas, Brazil

<sup>⊥</sup>Departamento de Química, Universidade Federal do Amazonas, Amazonas, Brazil

<sup>||</sup>Systems Engineering and Computer Science Program, COPPE, Federal University of Rio de Janeiro, Rio de Janeiro 21941-909, Brazil

## Supporting Information



**ABSTRACT:** Accessing localized proteomic profiles has emerged as a fundamental strategy to understand the biology of diseases, as recently demonstrated, for example, in the context of determining cancer resection margins with improved precision. Here, we analyze a gastric cancer biopsy sectioned into 10 parts, each one subjected to MudPIT analysis. We introduce a software tool, named Shotgun Imaging Analyzer and inspired in MALDI imaging, to enable the overlaying of a protein's expression heat map on a tissue picture. The software is tightly integrated with the NeXtProt database, so it enables the browsing of identified proteins according to chromosomes, quickly listing human proteins never identified by mass spectrometry (i.e., the so-called missing proteins), and the automatic search for proteins that are more expressed over a specific region of interest on the biopsy, all of which constitute goals that are clearly well-aligned with those of the C-HPP. Our software has been able to highlight an intense expression of proteins previously known to be correlated with cancers (e.g., glutathione S-transferase Mu 3), and in particular, we draw attention to Gastroke-2, a “missing protein” identified in this work of which we were able to clearly delineate the tumoral region from the “healthy” with our approach. Data are available via ProteomeXchange with identifier PXD000584.

**KEYWORDS:** chromosome 15, shotgun imaging, gastric cancer, Gastroke

## INTRODUCTION

The study of proteomic landscapes has surfaced as a key technology, with applications in diagnosing and in better understanding pathologies. One of the pioneers of this field, Richard Caprioli, has continually used MALDI (matrix-assisted laser desorption ionization) imaging to pinpoint molecular changes previously undetected by immunohistochemistry or morphological assessments, and thus diagnosed as “normal” but having nevertheless many characteristics in common with tumors.<sup>1,2</sup> The widely adopted MALDI imaging approach is capable of profiling tissues by describing their landscape as a

group of mass spectral peaks with corresponding intensities. Software applications are then used to perform image analyses to allow the visualization of tissue areas where some mass spectral peak appears that is more commonly found in tumors. One of the limitations of this strategy is that protein identification is fairly compromised, given the limited availability of samples (thin tissue slices) and the challenges

**Special Issue:** Chromosome-centric Human Proteome Project

**Received:** September 10, 2013

**Published:** November 22, 2013

associated with locally trypsinizing and identifying in-depth a variety of components, although efforts have been made to overcome these issues compared to that of shotgun proteomic strategies. Shotgun proteomics is a “bottom-up” approach that enables large-scale protein analysis of almost complete proteomes;<sup>3</sup> here we employ a shotgun approach for imaging, yielding as a disadvantage the lack of spatial resolution that MALDI imaging does. Thus, advantages and disadvantages of each methodology make them complementary. Nevertheless, we show shotgun imaging to be an important research tool perfectly suited to collect and supply qualitative and quantitative protein data that are unavailable by other techniques; MALDI and shotgun complement each other.

The multidimensional protein identification technology (MudPIT) embodies a powerful shotgun proteomic strategy that employs two-dimensional liquid chromatography (LC/LC) online with tandem mass spectrometry (MS/MS).<sup>4</sup> We previously employed this approach to compare gastric cancer biopsies with both resection margins and healthy tissues and concluded that resection margins “looked more proteomically alike” to cancer biopsies than to healthy tissues.<sup>5</sup> Among the conclusions, we postulated that resection margins play a key role in Paget’s “soil to seed” hypothesis, which states that cancer cells require a special microenvironment to nourish and that understanding it could ultimately lead to more effective treatments. With this as motivation, here we take an additional step by further exploring the spatial proteomic landscape of gastric cancer biopsies. To this end, we sliced a biopsy into 10 sections and then meticulously analyzed each piece by MudPIT. While the widely adopted approach would be to use MALDI imaging to characterize a profile for the tissue at hand, we believe that this would have the drawback of not leading to as large-scale protein identification as that of MudPIT.<sup>6</sup> Although our proposed data analysis strategy does not offer the spatial resolution from MALDI imaging approaches, it is nonetheless backed up by an in-depth quantitative MudPIT analysis that, in all, we have termed shotgun imaging.

Another motivation for this work stems from the chromosome-centric human proteome project (C-HPP) initiated by the Human Proteome Organization (HUPO), whose goal is to map the entire human protein set, which lacks any mass spectrometric evidence, the so-called “missing proteins”. One of the goals of NeXtProt has therefore been to keep track of which proteins have mass spectrometric identifications. NeXtProt is then helping to pave the way along the C-HPP roadmap by filling in as a human protein knowledgebase that includes data on protein expression from tissues and cells and by tracking the progress in identifying the remaining 3844 missing proteins. With this as motivation, we have taken the foundations of HUPO (C-HPP and B/D-HPP) as core guidelines for our shotgun imaging pipeline and exemplify it with the proteomic profiles of gastric cancer to search for missing proteins and provide information linked to this disease. As a result, we have tailored our imaging analysis software to be able to join MudPIT results with images and to extract information from NeXtProt, aiming to address C-HPP goals such as browsing proteins by chromosomes, easily listing the identifications of missing proteins, and helping to track the ones whose expressions correlate with diseased regions. The entire workflow described herein is integrated into the PatternLab for Proteomics pipeline, which offers an arsenal of tools for analyzing shotgun proteomic data, ranging from

protein quantitation, to differential proteomics, to gene ontology analysis, to name a few.<sup>7</sup>

In summary, here we present a software application that allows for overlaying quantitative shotgun proteomic data onto images. As described above, there are well-established approaches for achieving this goal, but they are not backed up by in-depth protein identification nor easily integrated with a software. Regardless, existing strategies are complementary, and therefore, the purpose determines what methods are best.

## ■ MATERIALS AND METHODS

### Subjects

This study was approved by the Ethics Committee of the Clementino Fraga Filho University Hospital of the Federal University of Rio de Janeiro (HUCFF/UFRJ: MEMO, No. 10252913.5.0000.5257, CAAE). The tumor biopsy was collected at the Oncology Control Foundation Center of the Amazonas State (FCECON) after the signing of informed consent. The biopsy was acquired from an area along the stomach that included tumor and resection margin, during the operation procedures on a male patient. Briefly, the tumor was located in the gastric antrum, and the resection margin was macroscopically defined during the surgery as a 10 cm rim of healthy-looking tissue surrounding the tumor. The biopsy was then subtyped, and the clinical stage of the disease was determined according to the tumor, node, and metastasis (TNM) classification of the American Joint Committee on Cancer (AJCC). The histological type was determined to be adenocarcinoma and classified as T4. Subsequently, the biopsy was vertically divided into 10 sections of equal length to be further analyzed by MudPIT.

### Protein Solubilization with RapiGest and Digestion with Trypsin

All biopsy sections were pulverized with liquid nitrogen. Each protein pellet was resuspended independently with RapiGest SF according to the manufacturer’s instructions to a final concentration of 0.1%. The samples were quantified using the Qubit fluorimetric quantitation (Invitrogen) as described in the manufacturer’s instructions. One hundred micrograms of each sample was reduced with 20 mM dithiothreitol (DTT) at 60 °C for 30 min. The samples were cooled to room temperature and incubated in the dark with 66 mM of iodoacetamide (IAA) for 30 min. Afterward, all samples were digested overnight with trypsin (Promega) at the ratio of 1/50 w/w (E/S) at 37 °C. Following digestion, all reactions were acidified with 10% formic acid (1% final concentration) to stop the proteolysis. The samples were centrifuged for 30 min at 20 800 rcf to remove insoluble material.

### LC/LC/MS/MS Data Acquisition

Sixty micrograms of the digested peptide mixture was desalted using a reversed-phase column manually packed in a tip using the Poros R2 resin (Applied Biosystems). The desalted peptides were resuspended in solution A (5 mM KH<sub>2</sub>PO<sub>4</sub> + 25% ACN pH 3) and loaded into a strong cation exchange microspin column from Harvard Apparatus. Peptides were eluted from the column in a stepwise manner by applying solution A with increasing KCl concentrations of 85, 150, 250, and 400 mM. Each fraction was desalted once again and analyzed on a reversed-phase column coupled to an Orbitrap Velos mass spectrometer (Thermo, San Jose, CA). The peptide mixtures were separated in a 20 cm analytical column (75 μm

inner diameter) that was packed in-house with 3  $\mu\text{m}$  C18 beads (Reprosil-AQ Pur, Dr. Maisch). The flow rate at the tip of the reverse column was 200 nL/min when the mobile phase composition was 5% acetonitrile in 0.1% formic acid. We then applied a 120 min gradient: using first 5–50% acetonitrile in 0.1% formic acid for 100 min, then 50–95% acetonitrile in 0.1% formic acid for 20 min. The effluent from the nLC column was directly electrosprayed into the mass spectrometer.

The LTQ Orbitrap Velos instrument was operated in data-dependent acquisition mode to automatically switch between full scan MS and MS/MS acquisition with a dynamic exclusion of 90 s. For the HCD/CID top 6 method, survey full scan MS spectra (from  $m/z$  350–2000) were acquired in the Orbitrap system with resolution of 60 000 at  $m/z$  of 400. The six most intense peptide ions with charge states of 2+ or 3+ were sequentially isolated and fragmented in the CID and HCD collision cells using normalized collision energies of 35 and 50, respectively. The resulting fragments were detected in the Orbitrap system with a resolution of 7500. Standard mass spectrometric conditions for all experiments were as follows: 2.5 kV spray voltage, no sheath and auxiliary gas flow, heated capillary temperature of 235  $^{\circ}\text{C}$ , predictive automatic gain control (AGC) enabled, and an S-lens RF level of 70%. Mass spectrometer scan functions and nLC solvent gradients were controlled by the Xcalibur data system (Thermo, San Jose, CA).

### Peptide Spectrum Matching

Mass spectra were extracted to the MS2 format using PatternLab's RawReader (available at: <http://proteomics.fiocruz.br/Softwares.aspx>). The NeXtProt database release 2013-07-15 was downloaded. A target decoy database was then generated using PatternLab to include a reversed version of each sequence found in the database plus those from 127 common mass spectrometry contaminants. The ProLuCID search engine (v 1.3) was used to compare experimental tandem mass spectra against those theoretically generated from our sequence database and to select the most likely peptide spectrum matches (PSMs).<sup>8</sup> Briefly, the search was limited to fully and semitryptic peptide candidates and imposed carbamidomethylation and oxidation of methionine as fixed and variable modification, respectively. The search engine accepted peptide candidates within a 50 ppm tolerance from the measured precursor  $m/z$  and used the XCorr and Z-Score as the primary and secondary search engine scores, respectively.

### Assessment of PSMs

The validity of the PSMs was assessed using the Search Engine Processor (SEPro)<sup>9</sup> (v 2.1.0.23). Identifications were grouped by charge state (+2 and  $\geq$ 3) and then by tryptic status (fully tryptic, semitryptic), resulting in four distinct subgroups. For each result, the ProLuCID XCorr, DeltaCN, and Z-Score values were used to generate a Bayesian discriminator. The identifications were sorted in nondecreasing order according to the discriminator score. A cutoff score was established to accept a false-discovery rate (FDR) of 1% based on the number of labeled decoys. This procedure was independently performed on each data subset, resulting in an FDR that was independent of tryptic status or charge state. Additionally, a minimum sequence length of six amino acid residues was required. Results were postprocessed to only accept PSMs with less than 5 ppm and proteins supported by two or more independent evidence (e.g., identification of a peptide with different charge states, a modified and a nonmodified version of

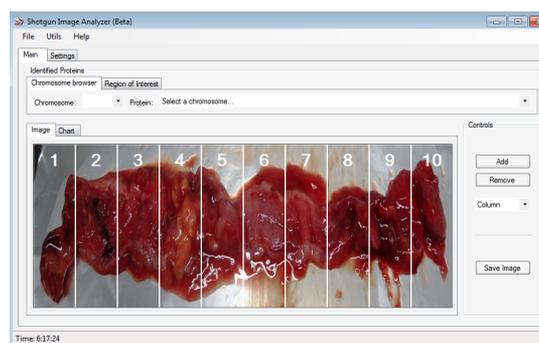
the same peptide, or two different peptides). This last filter led to a 0% FDR in all search results. For our quantitative shotgun imaging analysis, we only considered proteins having at least one unique peptide identified.

### Relative Protein Quantitation

The MS1 was extracted with RawReader and deconvoluted using YADA's default settings for bottom-up shotgun proteomics.<sup>10</sup> Extracted ion chromatograms (XICs) were obtained utilizing SEPro's Quantitation module as previously described.<sup>7</sup> In what followed, PatternLab's Regrouper module normalized the quantitative data according to the distributed normalized ion abundance factor (dNIAF)<sup>5</sup> approach, which is an adaptation of dNSAF for XICs.<sup>11</sup> Finally, PatternLab's index and sparse matrix files were generated;<sup>7</sup> briefly, these are text files that summarize all the results from all MudPIT runs (i.e., identification and quantification) and serve as input to our imaging software together with NeXtProt and the tumor's image.

### Shotgun Imaging Software

We have developed a shotgun imaging software to allow the visual assessment of proteomic results. We henceforth refer to it as the shotgun imaging analyzer (SIA). SIA was implemented in C# NetFramework 4.5 for the Windows operating system (7 or later) and is available for download at <http://proteomics.fiocruz.br/software/shotgunImaging>. Its graphical user interface (GUI) displaying an image of the biopsy used in this work is shown in Figure 1. The GUI contains two main tabs: one for



**Figure 1.** Shotgun imaging graphical user interface displaying an image of the gastric biopsy analyzed in this work.

viewing the results and the other for entering the index and sparse matrix files generated with PatternLab for Proteomics,<sup>7</sup> NeXtProt, and a path to an image file of the biopsy to overlay an expression heat map. From there on, the user can then generate the heat map's wireframe, indicating the locations where the biopsy was sectioned. Technically, this translates into each section corresponding to a class label of the sparse matrix that should be generated according to a previously published bioinformatics protocol.<sup>7</sup>

SIA has been tailored to take advantage of NeXtProt and therefore fits well in the context of C-HPP. A pull-down menu enables the browsing of identified proteins according to a chromosome of interest. By proceeding in this way, only proteins identified in the experiment at hand for that respective chromosome will be listed. Once a protein is selected, the GUI will overlay a heat map using a gray/light-bluish scale to indicate regions (lanes) where the protein was not identified and different tones of red to indicate its relative abundance in the lanes where it was identified. Quantitative values are made

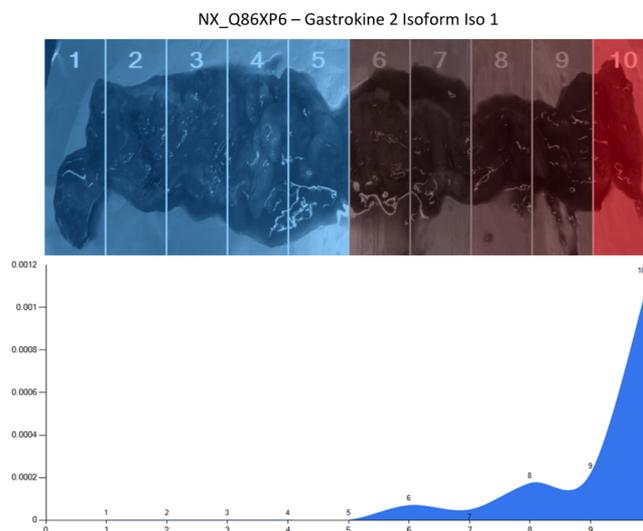
available in a table located in the Utils menu. Another option is that this menu allows a report to be displayed containing the identified proteins that were marked as missing by NeXtProt.

Once all the data have been loaded through the GUI and the biopsy sections have been determined using the controls on the right, the user can save all information into a unified file for sharing with colleagues by means of the Save option available in the File menu; this file has an “.sip” extension for shotgun imaging project. Finally, SIA offers an option to automatically search for proteins that present a higher expression in a region of interest (ROI). For the experiment at hand, this translates into informing the software the numbers of the lanes that correspond to the tumor in order to enable the sorting by SIA of the identified proteins in a nonincreasing order according to the Golub index (i.e., the difference of the means of the quantitation values divided by sum of their standard deviations).

## RESULTS AND DISCUSSION

We have described SIA, a software for shotgun imaging that enables the correlation of high-throughput identification and quantitation of proteins with a visual assessment of their locations. SIA follows a chromosome-centric approach that enables the listing of identified proteins according to the chromosomes to which they correspond. Moreover, it can automatically pinpoint proteins that show a higher expression in a given ROI (e.g., the diseased area) and is aligned with some of the goals of C-HPP.<sup>12,13</sup> Our analysis has identified up to 1936 proteins (with redundancy), of which 657 have at least one unique peptide. Our identifications are made available through the \*.sepr files and can be viewed with the Search Engine Processor viewer; more on their availability in the end of this article; Supporting Information Table 1 lists the proteins identified in each of the 10 biopsy sections. A desired feature in the quest for novel biomarker candidates is to pinpoint tissue-specific proteins.

Gastrokinins have been described as abundant proteins that are specifically expressed in the superficial gastric epithelium and have high evolutionary conservation. In particular, we draw attention to the identification of gastrokine isoform 2 (GKN2), a chromosome 2 protein that lacks mass spectrometric evidence of expression according to NeXtProt and therefore belongs to the list of the so-called missing proteins. Most importantly, GKN2 poses as a key protein for gastric cancer prognosis according to complementary approaches (e.g., western, PCR, microarray).<sup>14–16</sup> That said, as far as we know, this is the first report to demonstrate the protein expression of GKN2 and its use for delineating the tumoral region using mass spectrometric evidence. Interestingly, GKN2, an abundant protein for the tissue at hand, is down-regulated and many times described as absent in gastric cancer.<sup>17</sup> In a previous report, Moss and collaborators used a transcriptional profiling approach to describe a decreased expression of GKN2 and its relation to tumor prognosis; the authors reported an expression loss of GKN2 in 85% of diffuse intestinal-type cancers and correlated this with a significantly worse outcome ( $p < 0.03$ ).<sup>16</sup> More recently, Mao and collaborators corroborated these results using RT-PCR, Western blot, and immunohistochemistry for GKN1.<sup>14,15</sup> They further discuss that this isoform function is to protect and maintain the integrity of the gastric epithelium, and that a restoration of GKN1 expression can suppress the gastric cancer cell viability. Our shotgun imaging result for GKN2 is found in Figure 2; we note that the shotgun image



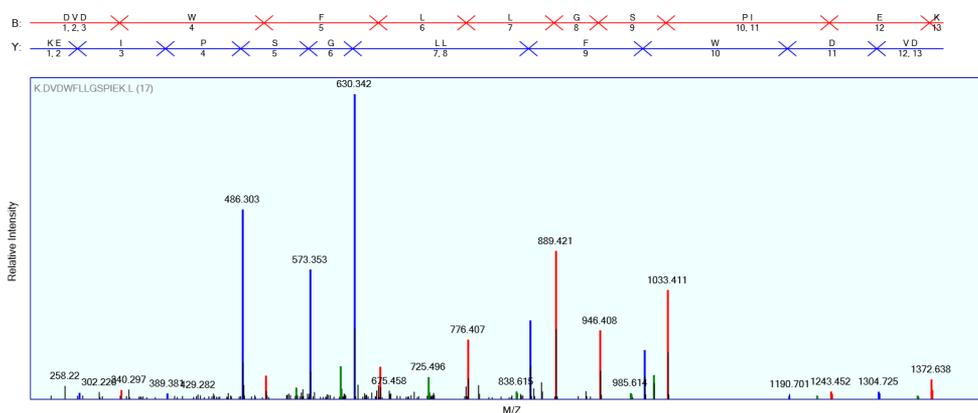
**Figure 2.** Shotgun imaging analyzer result for protein NX\_Q86XP6, Gastrokine 2 isoform. The top panel refers to the overlaid quantitation heat map. The bluish colors indicate absence of NX\_Q86XP6, darker red (e.g., as found in lane 10) is correlated with a higher expression of this protein by XIC. In the panel below, the y-axis represents the distributed normalized spectral abundance factor values (i.e., quantitation values), and the x-axis was aligned to the shotgun image to correspond to the slice numbers. A spline regressor was used to plot the area chart in blue. Lanes 1–4 are the tumoral regions, and regions 6–10 were classified as nontumoral.

identification of GKN1 isoform is very similar to that of Figure 2 (data not shown). Our identification of GKN2 contains four unique peptides, of which we exemplify the mass spectrum of one of them in Figure 3. Results clearly demonstrate increasing expressions in slices further apart from the tumoral region (1–4).

It is worthy to note that we also identified trefoil factor isoform 1 (TFF1) from chromosome 21 (NX\_P04155), a protein that has been found in high levels in the upper gastric mucosal cells and known to interact with GKN2. Their modes of action remain unsolved; however, there was a recent demonstration of the existence of a GKN2–trefoil factor 1 heterodimer.<sup>18</sup> TFF1 expression was found in all tissue sections except in lane 3, which lies within the tumoral region. Although we cannot make a statistical claim (i.e.,  $p > 0.05$ ), the average values of DNIAF are 1.7 higher in the nontumoral region. This result is supported by the literature that states that TFF1 should be absent (or under-expressed) in the tumor area.<sup>16</sup>

As previously described, SIA can sort all protein identifications according to the Golub index when comparing differences in ROIs, which for the case at hand are sections 1–4 (the tumor ROIs). Among the top-ranking proteins in the results of the automatic search, we point out annexin, glutathione S-transferases Mu 3, and lumican. We discuss these proteins next and then take further note of the role played by tropomyosins in tumor development and of their presence in the biopsy in question.

Annexin is one of the calcium and membrane-binding proteins. The literature correlates this protein with a wide variety of cellular functions, such as membrane aggregation, phagocytosis, proliferation, apoptosis, and even tumorigenesis.<sup>19</sup> A lower abundance of annexin has been associated with a poor prognosis in prostate cancer,<sup>20</sup> head and neck cancer,<sup>21</sup> sinonasal adenocarcinomas,<sup>22</sup> and hepatocellular cancer.<sup>23</sup>



**Figure 3.** Tandem mass spectrum of the tryptic peptide K.DVDWFLGSPIEK.L uniquely found in the GKN-2 sequence.

Interestingly, at the molecular level, our results have shown annexin A1 (chromosome 9, NX\_P04083) to be less abundant in cancer ROIs, which is in good agreement with previous work that, by complementary approaches such as DNA microarrays, immunoblotting, and immunohistochemistry,<sup>19,24,25</sup> pinpoint the annexin gene as a potential marker for gastric cancer. While the role of this protein in gastric cancer is still unknown, there is increasing evidence that changes in annexin expression and its subcellular localization can contribute to the development and progression of the disease.

Glutathione S-transferases (chromosome 1) constitute a class of proteins that have been correlated with detoxification and a protective mechanism against the development of cancer. The gene family for this class has been shown to be highly polymorphic, being expressed in only 50% of individuals.<sup>26–28</sup> A previous study correlated an elevated expression of this protein's mRNA with a poor prognosis.<sup>29</sup> Our results indicate a higher abundance of glutathione S-transferases Mu 3 in the tumoral ROIs.

Lumican is one of the small leucine-rich proteoglycans (SLRPs), with molecular weight of about 40 kDa comprised in four domains.<sup>30</sup> Previous works suggest that it plays a key role in cell–cell adhesion, in regulating stromal collagen matrix assembly, and it may also serve as a regulatory molecule of several cellular functions, such as cell migration and proliferation, apoptosis, differentiation, and inflammatory responses.<sup>30–32</sup> In cancers, lumican expression has been correlated with tumoral growth and metastasis in colorectal and pancreatic tumors, benign prostatic hyperplasia, and especially in breast tumors, where it has been detected in stromal cells adjacent to tumor cells.<sup>33,30,34,35</sup> Furthermore, high expression levels of this protein can be associated with high pathological tumor grades.<sup>35</sup> Its role in tumorigenesis remains elusive, but recent studies demonstrate that overexpression of lumican suppresses transformation induced by v-src and v-K-ras.<sup>30,34</sup> These observations may be explained by the possibility that the lumican receptors mediate different signaling transduction pathways in a cell-type-specific manner, which in turn provides a scaffold for macrophage infiltration.<sup>30,35,36</sup> In our imaging results, lumican expression is mostly found in the tumoral area.

Tropomyosins are components of actin filaments that play a critical role in regulating the interaction of actin and myosin. They are, therefore, important modulators of the adhesion dynamics that determines cell migration.<sup>37</sup> Previous studies demonstrate that different tropomyosin isoforms perform

distinct physiological roles, and that their expression profile is profoundly changed during the malignant progression of cancer cells since such cells change the organization of cytoskeletal and cellular morphology.<sup>38,39</sup> In some types of cancers, such as breast cancer, a low expression of these proteins in the primary carcinoma has been detected, while on the other hand, high levels have been observed in metastatic tumors.<sup>40</sup> Similar results are reported for gastric cancer in the literature.<sup>39</sup> In our results, tropomyosins were found with elevated and approximately uniform expression levels throughout all biopsy sections.

We would also like to point out that in this work we identified E-cadherin only in the resection margin; this result is aligned with our previous report.<sup>5</sup> In the literature, this protein has been associated with an important role in the early stage of tumorigenesis by modulating intracellular signaling which can promote tumor growth and be associated with metastasis.<sup>41</sup>

Finally, we wish to conclude with some remarks concerning the methodology we have used. First of all, we fully concur with the wider community in recognizing that MALDI imaging, with its remarkable spatial resolution, is arguably the current gold standard for exploring proteomic landscapes. On the other hand, we do also recognize that electrospray approaches are emerging as competitors. An example comes from the recently published work headed by L. Eberlin and R. Graham Cooks, in which they propose the use of ambient mass spectrometry (DESI) in a lipidomic approach to the intraoperative molecular diagnosis of brain tumors.<sup>42</sup> Moreover, their work demonstrates the potential of DESI-MS to identify the histological type of brain tumors. These authors' approach includes a statistical classifier trained from the same mass spectrometry imaging used for histopathology diagnosis. Another example is a work led by Vladislav Petyuk and Richard Smith, where the authors utilized shotgun proteomics with electrospray ionization to profile a mammalian brain; they were able to produce shotgun images with in-depth identification but bound by similar spatial resolution as in our work.<sup>43,44</sup> Reports such as this demonstrate the ever-growing importance of in-depth, high-throughput biomolecular mapping correlated with spatial disposition in potentially impacting medical treatment. While our approach exhibits poor spatial resolution, it is backed up by MudPIT, itself the gold standard in terms of number of identifications, which enabled us to identify a key missing protein and its partner related to gastric cancer. Taken together, these three strategies (MALDI, DESI, and shotgun imaging) can work in complementary ways to aid in developing more effective treatments for diseases.

## Availability

SIA and the project file analyzed in this work are available through our site, <http://proteomics.fiocruz.br/software/shotgunImaging>. By downloading the project file into SIA, one can view shotgun imaging heat maps for any of our protein identifications. The RAW data, together with all ProLuCID (\*.sqt) and SEPro files, are available at <http://dm64.ioc.fiocruz.br/sia/>. The mass spectrometry proteomics data have also been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository<sup>45</sup> with the data set identifier PXD000584.

## ■ ASSOCIATED CONTENT

### Supporting Information

Supplementary Table 1. A list of identified proteins for each of the 10 tumor sections. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: paulo@pcarvalho.com. Phone: +55(41)3316-3230. Fax: +55(41)3316-3267.

\*E-mail: gilberto@iq.ufjr.br. Phone: +55(21) 2562-7353.

### Author Contributions

<sup>¶</sup>P.F.A. and D.B.L. contributed equally to this work.

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

P.C.C. was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and RPT02H PDTIS/Carlos Chagas Institute-Fiocruz Parana. D.B.L. was funded by Instituto de Biologia Molecular do Paraná. V.C.B. acknowledges partial support from CNPq, CAPES, and a FAPERJ BBP grant. G.B.D. acknowledges support from CNPq (BP 308819/2011) and FAPERJ (E-26/110.138/2013). P.F.A. acknowledges scholarship support from CAPES.

## ■ REFERENCES

- (1) Chaurand, P.; Norris, J. L.; Cornett, D. S.; Mobley, J. A.; Caprioli, R. M. New developments in profiling and imaging of proteins from tissue sections by MALDI mass spectrometry. *J. Proteome Res.* **2006**, *5* (11), 2889–2900.
- (2) Oppenheimer, S. R.; Mi, D.; Sanders, M. E.; Caprioli, R. M. Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. *J. Proteome Res.* **2010**, *9* (5), 2182–2190.
- (3) Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., III. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **2013**, *113* (4), 2343–2394.
- (4) Washburn, M. P.; Ulaszek, R.; Deciu, C.; Schieltz, D. M.; Yates, J. R., III. Analysis of quantitative proteomic data generated via multidimensional protein identification technology. *Anal. Chem.* **2002**, *74* (7), 1650–1657.
- (5) Aquino, P. F.; Fischer, J. S.; Neves-Ferreira, A. G.; Perales, J.; Domont, G. B.; Araujo, G. D.; Barbosa, V. C.; Viana, J.; Chalub, S. R.; Lima de Souza, A. Q.; Carvalho, M. G.; Leao de Souza, A. D.; Carvalho, P. C. Are gastric cancer resection margin proteomic profiles more similar to those from controls or tumors? *J. Proteome Res.* **2012**, *11* (12), 5836–5842.
- (6) Kim, H. K.; Reyzer, M. L.; Choi, I. J.; Kim, C. G.; Kim, H. S.; Oshima, A.; Chertov, O.; Colantonio, S.; Fisher, R. J.; Allen, J. L.; Caprioli, R. M.; Green, J. E. Gastric cancer-specific protein profile

identified using endoscopic biopsy samples via MALDI mass spectrometry. *J. Proteome Res.* **2010**, *9* (8), 4123–4130.

- (7) Carvalho, P. C.; Fischer, J. S.; Xu, T.; Yates, J. R., III; Barbosa, V. C. PatternLab: From Mass Spectra to Label-Free Differential Shotgun Proteomics. In *Current Protocols in Bioinformatics*; Wiley: New York, 2012; Chapter 13, Unit 13-19.

- (8) Xu, T.; Venable, J. D.; Park, S. K.; Cociorva, D.; Lu, B.; Liao, L.; Wohlschlegel, J.; Hewel, J.; Yates, J. R., III. ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Mol. Cell. Proteomics* **2006**, *5*, S174.

- (9) Carvalho, P. C.; Fischer, J. S.; Xu, T.; Cociorva, D.; Balbuena, T. S.; Valente, R. H.; Perales, J.; Yates, J. R., III; Barbosa, V. C. Search engine processor: Filtering and organizing peptide spectrum matches. *Proteomics* **2012**, *12* (7), 944–949.

- (10) Carvalho, P. C.; Xu, T.; Han, X.; Cociorva, D.; Barbosa, V. C.; Yates, J. R., III. YADA: A tool for taking the most out of high-resolution spectra. *Bioinformatics* **2009**, *25* (20), 2734–2736.

- (11) Zhang, Y.; Wen, Z.; Washburn, M. P.; Florens, L. Refinements to label free proteome quantitation: how to deal with peptides shared by multiple proteins. *Anal. Chem.* **2010**, *82* (6), 2272–2281.

- (12) Aebbersold, R.; Bader, G. D.; Edwards, A. M.; van Eyk, J. E.; Kussmann, M.; Qin, J.; Omenn, G. S. The biology/disease-driven human proteome project (B/D-HPP): enabling protein research for the life sciences community. *J. Proteome Res.* **2013**, *12* (1), 23–27.

- (13) Huhmer, A. F.; Paulus, A.; Martin, L. B.; Millis, K.; Agreste, T.; Saba, J.; Lill, J. R.; Fischer, S. M.; Dracup, W.; Lavery, P. The chromosome-centric human proteome project: a call to action. *J. Proteome Res.* **2013**, *12* (1), 28–32.

- (14) Mao, W.; Chen, J.; Peng, T. L.; Yin, X. F.; Chen, L. Z.; Chen, M. H. Downregulation of gastrokine-1 in gastric cancer tissues and restoration of its expression induced gastric cancer cells to apoptosis. *J. Exp. Clin. Cancer Res.* **2012**, *31* (1), 49.

- (15) Mao, W.; Chen, J.; Peng, T. L.; Yin, X. F.; Chen, L. Z.; Chen, M. H. *Helicobacter pylori* infection and administration of non-steroidal anti-inflammatory drugs down-regulate the expression of gastrokine-1 in gastric mucosa. *Turk. J. Gastroenterol.* **2012**, *23* (3), 212–219.

- (16) Moss, S. F.; Lee, J. W.; Sabo, E.; Rubin, A. K.; Rommel, J.; Westley, B. R.; May, F. E.; Gao, J.; Meitner, P. A.; Tavares, R.; Resnick, M. B. Decreased expression of gastrokine 1 and the trefoil factor interacting protein TFIZ1/GKN2 in gastric cancer: influence of tumor histology and relationship to prognosis. *Clin. Cancer Res.* **2008**, *14* (13), 4161–4167.

- (17) Oien, K. A.; McGregor, F.; Butler, S.; Ferrier, R. K.; Downie, I.; Bryce, S.; Burns, S.; Keith, W. N. Gastrokine 1 is abundantly and specifically expressed in superficial gastric epithelium, down-regulated in gastric carcinoma, and shows high evolutionary conservation. *J. Pathol.* **2004**, *203* (3), 789–797.

- (18) Menhennott, T. R.; Kurklu, B.; Giraud, A. S. Gastrokines: stomach-specific proteins with putative homeostatic and tumor suppressor roles. *Am. J. Physiol. Gastrointest. Liver Physiol.* **2013**, *304* (2), G109–G121.

- (19) Cheng, T. Y.; Wu, M. S.; Lin, J. T.; Lin, M. T.; Shun, C. T.; Huang, H. Y.; Hua, K. T.; Kuo, M. L. Annexin A1 is associated with gastric cancer survival and promotes gastric cancer cell invasiveness through the formyl peptide receptor/extracellular signal-regulated kinase/integrin beta-1-binding protein 1 pathway. *Cancer* **2012**, *118* (23), 5757–5767.

- (20) Ding, T.; Yang, L.; Wang, Y.; Yuan, J.; Chen, T.; Cai, X. Down-regulation of annexin II in prostate cancer is associated with Gleason score, recurrence, metastasis and poor prognosis. *Mol. Med. Rep.* **2010**, *3* (5), 781–787.

- (21) Garcia Pedrero, J. M.; Fernandez, M. P.; Morgan, R. O.; Herrero, Z. A.; Gonzalez, M. V.; Suarez, N. C.; Rodrigo, J. P. Annexin A1 down-regulation in head and neck cancer is associated with epithelial differentiation status. *Am. J. Pathol.* **2004**, *164* (1), 73–79.

- (22) Rodrigo, J. P.; Garcia-Pedrero, J. M.; Llorente, J. L.; Fresno, M. F.; Allonca, E.; Suarez, C.; Hermsen, M. Down-regulation of annexin A1 and A2 protein expression in intestinal-type sinonasal adenocarcinomas. *Hum. Pathol.* **2011**, *42* (1), 88–94.

- (23) Liu, S. H.; Lin, C. Y.; Peng, S. Y.; Jeng, Y. M.; Pan, H. W.; Lai, P. L.; Liu, C. L.; Hsu, H. C. Down-regulation of annexin A10 in hepatocellular carcinoma is associated with vascular invasion, early recurrence, and poor prognosis in synergy with p53 mutation. *Am. J. Pathol.* **2002**, *160* (5), 1831–1837.
- (24) Emoto, K.; Sawada, H.; Yamada, Y.; Fujimoto, H.; Takahama, Y.; Ueno, M.; Takayama, T.; Uchida, H.; Kamada, K.; Naito, A.; Hirao, S.; Nakajima, Y. Annexin II overexpression is correlated with poor prognosis in human gastric carcinoma. *Anticancer Res.* **2001**, *21* (2B), 1339–1345.
- (25) Mussunoor, S.; Murray, G. I. The role of annexins in tumour development and progression. *J. Pathol.* **2008**, *216* (2), 131–140.
- (26) Alves, G. M. Glutathione S transferase mu polymorphism and gastric cancer in the Portuguese population. *Biomarkers* **1998**, *3* (6), 441–447.
- (27) Martinez, C.; Martin, F.; Fernandez, J. M.; Garcia-Martin, E.; Sastre, J.; Diaz-Rubio, M.; Agundez, J. A.; Ladero, J. M. Glutathione S-transferases mu 1, theta 1, pi 1, alpha 1 and mu 3 genetic polymorphisms and the risk of colorectal and gastric cancers in humans. *Pharmacogenomics* **2006**, *7* (5), 711–718.
- (28) Kocevar, N.; Odreman, F.; Vindigni, A.; Grazio, S. F.; Komel, R. Proteomic analysis of gastric cancer and immunoblot validation of potential biomarkers. *World J. Gastroenterol.* **2012**, *18* (11), 1216–1228.
- (29) Kearns, P. R.; Chrzanoska-Lightowlers, Z. M.; Pieters, R.; Veerman, A.; Hall, A. G. Mu class glutathione S-transferase mRNA isoform expression in acute lymphoblastic leukaemia. *Br. J. Haematol.* **2003**, *120* (1), 80–88.
- (30) Kao, W. W.; Funderburgh, J. L.; Xia, Y.; Liu, C. Y.; Conrad, G. W. Focus on molecules: lumican. *Exp. Eye Res.* **2006**, *82* (1), 3–4.
- (31) Kao, W. W.; Liu, C. Y. Roles of lumican and keratocan on corneal transparency. *Glycoconjugate J.* **2002**, *19* (4–5), 275–285.
- (32) Gu, G.; Cheng, W.; Yao, C.; Yin, J.; Tong, C.; Rao, A.; Yen, L.; Ku, M.; Rao, J. Quantitative proteomics analysis by isobaric tags for relative and absolute quantitation identified lumican as a potential marker for acute aortic dissection. *J. Biomed. Biotechnol.* **2011**, 920763.
- (33) Kelemen, L. E.; Couch, F. J.; Ahmed, S.; Dunning, A. M.; Pharoah, P. D.; Easton, D. F.; Fredericksen, Z. S.; Vierkant, R. A.; Pankratz, V. S.; Goode, E. L.; Scott, C. G.; Rider, D. N.; Wang, X.; Cerhan, J. R.; Vachon, C. M. Genetic variation in stromal proteins decorin and lumican with breast cancer: investigations in two case-control studies. *Breast Cancer Res.* **2008**, *10* (6), R98.
- (34) Yoshioka, N.; Inoue, H.; Nakanishi, K.; Oka, K.; Yutsudo, M.; Yamashita, A.; Hakura, A.; Nojima, H. Isolation of transformation suppressor genes by cDNA subtraction: lumican suppresses transformation induced by v-src and v-K-ras. *J. Virol.* **2000**, *74* (2), 1008–1013.
- (35) Naito, Z. Role of the small leucine-rich proteoglycan (SLRP) family in pathological lesions and cancer cell growth. *J. Nippon Med. Sch.* **2005**, *72* (3), 137–145.
- (36) Funderburgh, J. L.; Mitschler, R. R.; Funderburgh, M. L.; Roth, M. R.; Chapes, S. K.; Conrad, G. W. Macrophage receptors for lumican. A corneal keratan sulfate proteoglycan. *Invest. Ophthalmol. Visual Sci.* **1997**, *38* (6), 1159–1167.
- (37) Calmettes, C. Medullary cancer of the thyroid. Apropos of 20 years' experience in France. *Ann. Endocrinol. (Paris)* **1988**, *49* (1), 10–16.
- (38) Stehn, J. R.; Schevzov, G.; O'Neill, G. M.; Gunning, P. W. Specialisation of the tropomyosin composition of actin filaments provides new potential targets for chemotherapy. *Curr. Cancer Drug Targets* **2006**, *6* (3), 245–256.
- (39) He, Q. Y.; Cheung, Y. H.; Leung, S. Y.; Yuen, S. T.; Chu, K. M.; Chiu, J. F. Diverse proteomic alterations in gastric adenocarcinoma. *Proteomics* **2004**, *4* (10), 3276–3287.
- (40) Lee, H. H.; Lim, C. A.; Cheong, Y. T.; Singh, M.; Gam, L. H. Comparison of protein expression profiles of different stages of lymph nodes metastasis in breast cancer. *Int. J. Biol. Sci.* **2012**, *8* (3), 353–362.
- (41) Chan, A. O. E-cadherin in gastric cancer. *World J. Gastroenterol.* **2006**, *12* (2), 199–203.
- (42) Eberlin, L. S.; Norton, I.; Orringer, D.; Dunn, I. F.; Liu, X.; Ide, J. L.; Jarmusch, A. K.; Ligon, K. L.; Jolesz, F. A.; Golby, A. J.; Santagata, S.; Agar, N. Y.; Cooks, R. G. Ambient mass spectrometry for the intraoperative molecular diagnosis of human brain tumors. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (5), 1611–1616.
- (43) Petyuk, V. A.; Qian, W. J.; Smith, R. D.; Smith, D. J. Mapping protein abundance patterns in the brain using voxelation combined with liquid chromatography and mass spectrometry. *Methods* **2010**, *50* (2), 77–84.
- (44) Petyuk, V. A.; Qian, W. J.; Chin, M. H.; Wang, H.; Livesay, E. A.; Monroe, M. E.; Adkins, J. N.; Jaitly, N.; Anderson, D. J.; Camp, D. G.; Smith, D. J.; Smith, R. D. Spatial mapping of protein abundances in the mouse brain by voxelation integrated with high-throughput liquid chromatography-mass spectrometry. *Genome Res.* **2007**, *17* (3), 328–336.
- (45) Vizcaino, J. A.; Cote, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; O'Kelly, G.; Schoenegger, A.; Ovelleiro, D.; Perez-Riverol, Y.; Reisinger, F.; Rios, D.; Wang, R.; Hermjakob, H. The PRoteomics IDentifications (PRIDE) database and associated tools: status in 2013. *Nucleic Acids Res.* **2013**, No. 41, D1063–D1069.

**Anexo 5** – Lista de proteínas identificadas e quantificadas na abordagem tumor x margem de ressecção 1 cm.

<b>Média <i>ln Fold change</i></b>	<b><i>Stouffers p-valor</i></b>	<b>Descrição</b>	<b>Cromossomo</b>
-1.582	0.002802894	<i>AGR2</i>	7
-1.316	0.001454961	<i>UMP-CMP kinase</i>	1
-1.224	0.001771352	<i>Mucin-5AC (Fragments)</i>	11
-1.145	8.80599E-05	<i>Aldo-keto reductase family 1 member B10</i>	7
-1.053	4.57298E-07	<i>Carbonic anhydrase 2</i>	8
-1.008	0.000578145	<i>Hemoglobin subunit alpha</i>	16
-1.007	0.001937121	<i>Nucleolin</i>	2
-1.006	0.003421906	<i>GDP-mannose 4,6 dehydratase</i>	6
-0.996	0.002571323	<i>Gastricsin</i>	6
-0.972	0.002431069	<i>Peroxiredoxin-2</i>	19
-0.94	0.000140681	<i>Hemoglobin subunit beta</i>	11
-0.915	0.001728043	<i>Calreticulin</i>	19
-0.891	0.000444631	<i>Gastrokine-1</i>	2
-0.883	0.001076268	<i>Gastric triacylglycerol lipase</i>	10
-0.88	5.15894E-05	<i>Retinal dehydrogenase 1</i>	9
-0.877	0.001846368	<i>Histone H4</i>	1
-0.874	0.000123498	<i>78 kDa glucose-regulated protein</i>	9
-0.864	0.000116621	<i>Protein disulfide-isomerase A4</i>	7
-0.827	0.000708631	<i>Protein disulfide-isomerase A3</i>	15
-0.815	2.02703E-05	<i>Apolipoprotein A-I</i>	11
-0.799	0.001222021	<i>Aldehyde dehydrogenase, mitochondrial</i>	12
-0.793	0.003299947	<i>Vigilin</i>	2
-0.785	0.000378983	<i>Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1</i>	2
-0.758	0.001385043	<i>Fumarate hydratase, mitochondrial</i>	1
-0.752	0.00268739	<i>Hemoglobin subunit delta</i>	11
-0.741	0.00355895	<i>Marginal zone B- and B1-cell-specific protein</i>	5
-0.731	2.13397E-07	<i>Endoplasmic</i>	12
-0.73	0.000163694	<i>Ig mu chain C region</i>	14
-0.726	0.000295813	<i>Carbonic anhydrase 1</i>	8

-0.722	0.00012933	<i>Filamin-B</i>	3
-0.717	0.001260775	<i>Peroxiredoxin-5, mitochondrial</i>	11
-0.713	0.002836167	<i>Ribosome-binding protein 1</i>	20
-0.71	0.003504548	<i>GTP:AMP phosphotransferase AK3, mitochondrial</i>	9
-0.704	0.003994198	<i>Cytosolic non-specific dipeptidase</i>	18
-0.668	3.32475E-06	<i>Alpha-1-antitrypsin =2</i>	14
-0.667	0.001282391	<i>Isocitrate dehydrogenase [NADP], mitochondrial</i>	15
-0.666	3.01246E-05	<i>Apolipoprotein B-100</i>	2
-0.655	8.06086E-06	<i>Complement C3</i>	19
-0.654	0.003072743	<i>Very long-chain specific acyl-CoA dehydrogenase, mitochondrial</i>	17
-0.652	0.003834406	<i>Heterogeneous nuclear ribonucleoprotein U</i>	1
-0.65	7.64943E-05	<i>Ras GTPase-activating-like protein IQGAP1</i>	15
-0.646	0.002208117	<i>Glutathione S-transferase P</i>	11
-0.646	0.003716903	<i>Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit</i>	20
-0.634	0.00083367	<i>Hemopexin</i>	11
-0.624	0.000550793	<i>Malate dehydrogenase, mitochondrial</i>	7
-0.622	2.15012E-05	<i>Clathrin heavy chain 1</i>	17
-0.606	0.001372795	<i>Cytoplasmic dynein 1 heavy chain 1</i>	14
-0.605	0.000124743	<i>60 kDa heat shock protein, mitochondrial</i>	2
-0.597	2.16381E-07	<i>Serum albumin</i>	14
-0.597	0.000489989	<i>Ig alpha-1 chain C region</i>	4
-0.591	0.002312271	<i>Glycogen phosphorylase, brain form</i>	20
-0.581	0.003300855	<i>N(G),N(G)-dimethylarginine dimethylaminohydrolase 1</i>	1
-0.567	3.89584E-07	<i>Alpha-2-macroglobulin</i>	12
-0.551	0.002348465	<i>Bifunctional purine biosynthesis protein PURH</i>	2
-0.528	0.003880072	<i>Fibrinogen beta chain</i>	4
-0.486	0.00214128	<i>Importin subunit beta-1 2</i>	17

-0.454	0.001635475	<i>Fibrinogen alpha chain</i>	4
-0.448	0.003539532	<i>Trifunctional enzyme subunit alpha, mitochondrial</i>	2
-0.442	0.001561467	<i>Heat shock cognate 71 kDa protein</i>	11
-0.395	0.002302959	<i>Ig kappa chain C region</i>	2
-0.373	0.003732332	<i>Ig gamma-1 chain C region</i>	14
-0.372	0.003486882	<i>Plectin</i>	8
-0.314	8.35572E-05	<i>Serotransferrin</i>	3
0.095	0.002353966	<i>Peptidyl-prolyl cis-trans isomerase A</i>	7
0.149	0.000364645	<i>Talin-1</i>	9
0.294	0.002424884	<i>Profilin-1</i>	17
0.329	0.002874584	<i>EH domain-containing protein 2</i>	19
0.377	0.000416274	<i>Alpha-enolase</i>	1
0.391	0.001121188	<i>Destrin</i>	20
0.447	0.002511114	<i>Myosin light chain kinase, smooth muscle</i>	3
0.474	0.000340715	<i>Transgelin</i>	11
0.48	4.31824E-05	<i>Caldesmon</i>	7
0.502	5.66214E-15	<i>Filamin-A</i>	X
0.524	1.60044E-09	<i>Vinculin</i>	10
0.547	0.002316982	<i>Moesin</i>	1
0.547	0.003150159	<i>Protein enabled homolog</i>	X
0.554	0.000458138	<i>Transketolase PE=1 SV=1</i>	3
0.559	0.002461713	<i>Tropomyosin alpha-4</i>	19
0.562	0.002262157	<i>PDZ and LIM domain protein 7</i>	5
0.596	4.85642E-07	<i>Filamin-C</i>	7
0.625	0.000269301	<i>Dihydropyrimidinase-related protein 3</i>	5
0.632	0.003911889	<i>PDZ and LIM domain protein 3</i>	4
0.725	1.0643E-05	<i>Calponin-1</i>	19
0.764	0.00326279	<i>Hepatoma-derived growth factor</i>	1
0.965	0.000937725	<i>Tropomyosin beta chain</i>	9

**Anexo 6** – Lista de proteínas identificadas e quantificadas na abordagem tumor x margem de ressecção 3 cm.

<b>Média ln Fold change</b>	<b>Stouffers p-valor</b>	<b>Descrição</b>	<b>Cromossomo</b>
-1.027	0.008653622	<i>Phosphoglycerate mutase 1</i>	10
-1.017	0.000306271	<i>Hemoglobin subunit alpha</i>	16
-0.954	0.000984088	<i>Hemoglobin subunit beta</i>	11
-0.879	0.001076268	<i>Gastric triacylglycerol lipase</i>	10
-0.765	0.002334934	<i>Hemoglobin subunit delta</i>	11
-0.748	0.000190476	<i>Desmin</i>	2
-0.694	0.000774512	<i>Carbonic anhydrase 1</i>	8
-0.678	0.006388655	<i>Apolipoprotein A-II</i>	1
-0.648	0.007817257	<i>Peroxiredoxin-2</i>	19
-0.623	0.002725654	<i>Apolipoprotein B-100</i>	2
-0.61	0.001620264	<i>Creatine kinase B-type</i>	14
-0.561	5.26E-07	<i>Serum albumin</i>	4
-0.527	0.002907488	<i>Ig alpha-1 chain C region</i>	14
-0.516	0.005085202	<i>Retinal dehydrogenase 1</i>	9
-0.495	0.003108948	<i>Synemin</i>	15
-0.332	0.000705565	<i>Apolipoprotein A-I</i>	11
-0.319	0.00099896	<i>Carbonic anhydrase 2</i>	8
-0.292	0.006704129	<i>Ig mu chain C region</i>	14
0.035	0.005261984	<i>Superoxide dismutase [Cu-Zn]</i>	21
0.065	0.008294681	<i>Glutamate dehydrogenase</i>	10
0.142	0.005768361	<i>Phosphatidylethanolamine-binding protein 1</i>	12
0.158	0.003416386	<i>Glutamate dehydrogenase 1, mitochondrial</i>	10
0.333	0.001292439	<i>Calreticulin</i>	19
0.37	0.005589614	<i>Neutral alpha-glucosidase AB</i>	11
0.398	0.000595881	<i>Triosephosphate isomerase</i>	12
0.448	0.006611393	<i>Tropomyosin beta chain</i>	9
0.45	0.007946127	<i>Sulfide:quinone oxidoreductase, mitochondrial</i>	15
0.451	6.86E-06	<i>Talin-1</i>	9
0.454	0.003042598	<i>Coronin-1C</i>	12

0.456	0.005075753	<i>Aldo-keto reductase family 1 member B10</i>	7
0.458	0.003770332	<i>Adenosylhomocysteinase</i>	20
0.463	4.99E-05	<i>78 kDa glucose-regulated protein</i>	9
0.464	0.001162201	<i>Endoplasmin</i>	12
0.473	0.005120236	<i>Malate dehydrogenase, cytoplasmic</i>	2
0.477	0.008596488	<i>Cysteine and glycine-rich protein 1</i>	1
0.478	0.005387878	<i>Heterogeneous nuclear ribonucleoprotein U</i>	1
0.478	0.00728256	<i>tRNA-splicing ligase RtcB homolog</i>	22
0.479	0.003098225	<i>Phosphoglucomutase-1</i>	1
0.48	0.005177736	<i>Cytosolic non-specific dipeptidase</i>	18
0.483	0.001426995	<i>Electron transfer flavoprotein subunit beta</i>	19
0.485	0.002873458	<i>Dihydropyrimidinase-related protein 2</i>	8
0.487	0.005813556	<i>Heat shock 70 kDa protein 1A/1B</i>	6
0.489	0.000106081	<i>Gelsolin</i>	9
0.494	0.003438697	<i>Glutathione S-transferase P</i>	11
0.497	0.002433573	<i>Alcohol dehydrogenase [NADP(+)]</i>	1
0.5	0.005630234	<i>Quinone oxidoreductase</i>	1
0.505	0.005177736	<i>T-complex protein 1 subunit alpha</i>	6
0.508	0.001005518	<i>Heat shock protein HSP 90-alpha</i>	14
0.509	2.75E-05	<i>Filamin-C</i>	10
0.509	1.27E-05	<i>Neuroblast differentiation-associated protein AHNAK</i>	7
0.509	0.005075753	<i>Ribosome maturation protein SBDS</i>	11
0.509	6.10E-08	<i>Vinculin</i>	7
0.514	0.000370555	<i>WD repeat-containing protein 1</i>	4
0.531	0.007998209	<i>Leiomodin-1</i>	1
0.531	0.000202707	<i>Trifunctional enzyme subunit alpha, mitochondrial</i>	2
0.532	0.000415083	<i>Alanine--tRNA ligase, cytoplasmic</i>	16
0.536	0.006558577	<i>Importin subunit beta-1</i>	17
0.544	0.004461481	<i>Palladin</i>	4
0.546	0.00335178	<i>Catalase</i>	10
0.546	6.40E-07	<i>Vimentin</i>	11

0.548	0.0070169	<i>Septin-2</i>	2
0.55	0.002231262	<i>4-trimethylaminobutyraldehyde dehydrogenase</i>	1
0.554	0.000806164	<i>Clathrin heavy chain 1</i>	17
0.559	0.007385518	<i>PDZ and LIM domain protein 5</i>	4
0.561	0.002192938	<i>Chloride intracellular channel protein 4</i>	1
0.563	6.77E-06	<i>Phosphoglycerate kinase 1</i>	X
0.566	1.70E-07	<i>Pyruvate kinase PKM</i>	15
0.568	2.61E-05	<i>Dihydropyrimidinase-related protein 3</i>	2
0.568	0.001896981	<i>Fatty acid synthase</i>	17
0.568	0.006808435	<i>Heat shock 70 kDa protein 4</i>	5
0.568	0.006865538	<i>Poly(rC)-binding protein 1</i>	5
0.577	0.00214128	<i>Aspartate aminotransferase, mitochondrial</i>	16
0.579	3.77E-05	<i>Elongation factor 2</i>	19
0.579	0.006954519	<i>Interleukin enhancer-binding factor 2</i>	1
0.581	0.001719898	<i>T-complex protein 1 subunit theta</i>	21
0.582	0.00838573	<i>Annexin A2</i>	15
0.584	1.30E-05	<i>Caldesmon</i>	7
0.585	2.93E-05	<i>Ribosome-binding protein 1</i>	20
0.587	0.002763416	<i>UV excision repair protein RAD23 homolog B</i>	9
0.589	0.008252081	<i>Phosphoglucomutase-2 OS=Homo sapiens GN=PGM2 PE=1 SV=4 : F5H6V2 Phosphoglucomutase-2</i>	4
0.591	5.40E-05	<i>Heat shock cognate 71 kDa protein</i>	11
0.593	0.005507603	<i>Galectin-3</i>	14
0.594	0.007323972	<i>Peptidyl-prolyl cis-trans isomerase</i>	11
0.595	1.08E-05	<i>Alpha-actinin-4</i>	5
0.595	0.00511706	<i>Nucleophosmin</i>	19
0.6	1.25E-05	<i>Moesin</i>	X
0.602	0.000120212	<i>L-lactate dehydrogenase A chain</i>	11
0.604	0.00091081	<i>14-3-3 protein epsilon</i>	17
0.605	3.42E-05	<i>14-3-3 protein gamma</i>	7
0.605	0.005231349	<i>Protein disulfide-isomerase A4</i>	7
0.61	4.11E-05	<i>Calponin-1</i>	1

0.61	0.004470938	<i>Protein enabled homolog</i>	19
0.615	1.25E-05	<i>Adenylyl cyclase-associated protein 1</i>	1
0.62	0.002367066	<i>Filamin-B</i>	3
0.621	2.90E-05	<i>60 kDa heat shock protein, mitochondrial</i>	14
0.621	0.001114598	<i>Cytoplasmic dynein 1 heavy chain 1</i>	2
0.637	0.005177736	<i>Lamin-B1</i>	5
0.637	0.002635192	<i>N(G),N(G)-dimethylarginine dimethylaminohydrolase 1</i>	1
0.638	0.00125322	<i>Glutamine--fructose-6-phosphate aminotransferase [isomerizing] 1</i>	2
0.64	0.000138103	<i>Lumican</i>	17
0.64	0.002896551	<i>Profilin-1</i>	12
0.646	0.003460243	<i>Protein DJ-1</i>	1
0.649	0.005447141	<i>Ig gamma-4 chain C region</i>	14
0.654	3.16E-06	<i>Spectrin alpha chain, non-erythrocytic 1</i>	9
0.657	0.004117825	<i>Ras GTPase-activating-like protein IQGAP1</i>	15
0.658	0.000559887	<i>Ribonuclease inhibitor</i>	11
0.659	0.004317599	<i>Inorganic pyrophosphatase</i>	5
0.659	0.005387878	<i>Leukocyte elastase inhibitor</i>	6
0.659	0.000643562	<i>Stress-70 protein, mitochondrial</i>	10
0.66	0.001064326	<i>14-3-3 protein eta</i>	6
0.66	0.005275581	<i>ATP-binding cassette sub-family F member 1</i>	22
0.662	5.91E-06	<i>Spectrin alpha chain, non-erythrocytic 1</i>	1
0.665	8.32E-09	<i>Plectin</i>	8
0.665	0.002256279	<i>Thioredoxin</i>	9
0.667	0.000601278	<i>X-ray repair cross-complementing protein 5</i>	2
0.668	0.002578807	<i>14-3-3 protein theta</i>	2
0.669	0.005786854	<i>Leucine-rich PPR motif-containing protein, mitochondrial</i>	2

0.672	0.002673495	<i>Myristoylated alanine-rich C-kinase substrate</i>	6
0.682	0.001325427	<i>Cytosol aminopeptidase</i>	4
0.684	0.008402029	<i>Non-POU domain-containing octamer-binding protein (Fragment)</i>	X
0.686	0.002016529	<i>Fructose-bisphosphate aldolase A</i>	15
0.686	0.002461713	<i>Peptidyl-prolyl cis-trans isomerase B</i>	16
0.688	0.005075753	<i>Testin</i>	7
0.694	0.005075753	<i>Hsp90 co-chaperone Cdc37</i>	19
0.699	5.73E-07	<i>Spectrin beta chain, non-erythrocytic 1</i>	2
0.71	0.006699959	<i>Fibulin-1</i>	22
0.713	0.008454974	<i>Interferon-induced guanylate-binding protein 1</i>	1
0.714	0.002584602	<i>GTP-binding protein SAR1a</i>	10
0.715	7.76E-06	<i>Alpha-enolase</i>	1
0.721	0.006591718	<i>Hypoxia up-regulated protein 1</i>	11
0.722	6.13E-07	<i>Prelamin-A/C</i>	1
0.724	0.001647092	<i>14-3-3 protein beta/alpha</i>	20
0.729	0.005275581	<i>Calnexin</i>	5
0.739	0.002487011	<i>Transitional endoplasmic reticulum ATPase</i>	9
0.744	0.000351263	<i>Heat shock protein HSP 90-beta</i>	6
0.747	8.69E-05	<i>Isocitrate dehydrogenase [NADP] cytoplasmic</i>	12
0.747	0.001183469	<i>Rho GDP-dissociation inhibitor 2</i>	2
0.758	0.000116956	<i>Protein disulfide-isomerase A3</i>	15
0.766	0.005779373	<i>Galectin-1</i>	22
0.775	0.006642728	<i>Purine nucleoside phosphorylase</i>	14
0.78	0.000320512	<i>6-phosphogluconate dehydrogenase, decarboxylating</i>	19
0.78	0.006352045	<i>Glucose-6-phosphate isomerase</i>	1
0.799	6.85E-06	<i>Transketolase</i>	3
0.801	0.005214548	<i>Heterogeneous nuclear ribonucleoprotein Q</i>	6
0.802	0.000114369	<i>Chloride intracellular channel protein 1</i>	6

0.806	0.002715736	<i>Splicing factor, proline- and glutamine-rich</i>	1
0.813	0.002457611	<i>LIM and SH3 domain protein 1</i>	17
0.829	0.003316001	<i>Calpain-1 catalytic subunit</i>	11
0.836	0.008582921	<i>Coronin-1A</i>	16
0.843	0.002035664	<i>Peroxiredoxin-1</i>	1
0.876	0.004416929	<i>Glucose-6-phosphate isomerase</i>	19
0.887	0.005387878	<i>Protein S100-A11</i>	1
0.888	0.000230883	<i>Peptidyl-prolyl cis-trans isomerase A</i>	7
0.898	0.001369469	<i>Transgelin-2</i>	1
0.9	0.00511706	<i>Protein S100-A6</i>	1
0.903	0.000608326	<i>Heterogeneous nuclear ribonucleoproteins A2/B1</i>	7
0.921	0.006991614	<i>Proteasome subunit alpha type-1</i>	11
0.951	0.005538124	<i>Protein S100-P</i>	4
0.974	0.000979338	<i>Tropomyosin alpha-4 chain</i>	19
0.977	0.006320721	<i>Eukaryotic initiation factor 4A-I</i>	17
1.02	0.007937994	<i>Cystatin-B</i>	21
1.035	0.007531761	<i>Cathepsin D</i>	11
1.042	0.005137454	<i>Protein disulfide-isomerase</i>	17
1.042	0.005137454	<i>Protein disulfide-isomerase</i>	17
1.078	0.000583713	<i>Plastin-2</i>	13
1.087	0.00171249	<i>Macrophage-capping protein</i>	2
1.263	0.003422247	<i>Xaa-Pro dipeptidase</i>	19
1.265	0.000898748	<i>Hepatoma-derived growth factor</i>	1
1.397	0.005177736	<i>Thymosin beta-10</i>	2

Anexo 7 – Lista de proteínas identificadas e quantificadas na abordagem tumor x margem de ressecção 5 cm.

<b>Média ln Fold change</b>	<b>Stouffers p-valor</b>	<b>Descrição</b>	<b>Cromossomo</b>
-1.315	0.002571323	<i>Gastricsin</i>	6
-1.275	0.001076268	<i>Gastric triacylglycerol lipase</i>	10
-1.062	0.000118054	<i>Hemoglobin subunit alpha</i>	16
-0.885	0.000140681	<i>Hemoglobin subunit beta</i>	11
-0.861	0.002500117	<i>Peroxiredoxin-2</i>	19
-0.854	0.001689638	<i>Ig mu chain C region</i>	14
-0.809	0.000295813	<i>Carbonic anhydrase 1</i>	8
-0.775	0.002802894	<i>AGR2</i>	7
-0.769	0.001056468	<i>Hemoglobin subunit delta</i>	11
-0.733	3.72E-05	<i>Retinal dehydrogenase 1</i>	9
-0.709	9.26E-05	<i>Carbonic anhydrase 2</i>	8
-0.692	0.002373761	<i>Haptoglobin</i>	16
-0.662	0.003580337	<i>Glyceraldehyde-3-phosphate dehydrogenase</i>	12
-0.659	0.001246518	<i>Peroxiredoxin-5, mitochondrial</i>	11
-0.637	0.002968795	<i>Malate dehydrogenase, mitochondrial</i>	7
-0.633	0.000288746	<i>Creatine kinase B-type</i>	14
-0.608	1.91E-06	<i>Apolipoprotein B-100</i>	2
-0.6	0.001020378	<i>Isocitrate dehydrogenase [NADP], mitochondrial</i>	15
-0.57	4.37E-05	<i>Alpha-2-macroglobulin</i>	12
-0.567	2.94E-05	<i>Alpha-1-antitrypsin</i>	14
-0.562	0.003637882	<i>Histone H4</i>	1
-0.555	0.001139054	<i>Ig alpha-1 chain C region</i>	14
-0.554	0.000103767	<i>Apolipoprotein A-I</i>	11
-0.533	2.01E-07	<i>Serum albumin</i>	4
-0.515	0.001283491	<i>Fibrinogen beta chain</i>	4
-0.486	0.001850825	<i>Fibrinogen alpha chain</i>	4
-0.41	0.000192109	<i>Complement C3</i>	19
-0.372	0.001927326	<i>Desmin</i>	2
-0.357	0.000305373	<i>Protein disulfide-isomerase A4</i>	7
-0.217	0.000273213	<i>Serotransferrin</i>	3
-0.217	0.0014754	<i>Endoplasmin</i>	12

-0.205	0.004139311	<i>Ig kappa chain C region</i>	2
0.2	1.49E-05	<i>Talin-1</i>	9
0.214	0.001337167	<i>Transgelin</i>	11
0.222	0.000868379	<i>Myosin light chain kinase, smooth muscle</i>	3
0.288	0.00100796	<i>Destrin</i>	11
0.288	0.002290946	<i>Heat shock cognate 71 kDa protein</i>	20
0.305	0.002005195	<i>Tropomyosin 1 (Alpha), isoform CRA_f</i>	15
0.331	2.16E-05	<i>Pyruvate kinase PKM</i>	15
0.351	1.96E-05	<i>Vimentin</i>	10
0.374	0.000293717	<i>Gelsolin</i>	9
0.406	3.76E-06	<i>Spectrin beta chain, non-erythrocytic 1</i>	2
0.411	0.00012002	<i>L-lactate dehydrogenase A chain</i>	11
0.439	0.002972857	<i>Phosphoglucomutase-like protein 5</i>	9
0.449	0.003457008	<i>Palladin</i>	4
0.478	2.22E-16	<i>Filamin-A</i>	X
0.485	6.32E-07	<i>Filamin-C</i>	7
0.498	0.003046152	<i>Ribonuclease inhibitor</i>	11
0.5	0.003062905	<i>Trifunctional enzyme subunit alpha, mitochondrial</i>	2
0.501	0.002942169	<i>Aspartate aminotransferase, mitochondrial</i>	16
0.502	2.17E-07	<i>Vinculin</i>	10
0.507	0.003600485	<i>Peroxiredoxin-1</i>	1
0.529	0.002578807	<i>14-3-3 protein theta</i>	2
0.538	0.00341071	<i>Spectrin alpha chain, non-erythrocytic 1</i>	9
0.539	3.38E-05	<i>Alpha-enolase</i>	1
0.54	0.001274333	<i>Lumican</i>	12
0.557	0.000266962	<i>Moesin O</i>	X
0.559	0.0003635	<i>Tropomyosin beta chain</i>	9
0.564	0.000761283	<i>Isocitrate dehydrogenase [NADP] cytoplasmic</i>	2
0.566	8.73E-05	<i>Alpha-actinin-4</i>	19

0.567	0.002192938	<i>Chloride intracellular channel protein 4</i>	1
0.577	0.00233123	<i>Adenylyl cyclase-associated protein 1</i>	1
0.598	0.000563399	<i>Chloride intracellular channel protein 1</i>	6
0.599	0.00300038	<i>Fructose-bisphosphate aldolase A</i>	16
0.6	4.73E-06	<i>Plectin</i>	8
0.602	2.09E-05	<i>Dihydropyrimidinase-related protein 3</i>	5
0.606	0.000458531	<i>14-3-3 protein gamma</i>	7
0.614	1.14E-05	<i>Prelamin-A/C</i>	1
0.617	0.001213683	<i>Transgelin-2</i>	1
0.664	0.002896551	<i>Profilin-1</i>	17
0.677	2.68E-08	<i>Caldesmon</i>	22
0.677	0.002971031	<i>14-3-3 protein eta</i>	7
0.713	9.71E-06	<i>Calponin-1</i>	19
0.72	0.000115752	<i>Transketolase</i>	3
0.733	0.000230883	<i>Peptidyl-prolyl cis-trans isomerase A</i>	7
0.746	0.002432943	<i>Cysteine and glycine-rich protein 1</i>	1
0.785	0.00171249	<i>Macrophage-capping protein</i>	2
0.813	0.000979338	<i>Tropomyosin alpha-4 chain</i>	19
0.875	0.000898748	<i>Hepatoma-derived growth factor</i>	1
0.953	0.003525814	<i>Heterogeneous nuclear ribonucleoproteins A2/B1</i>	7