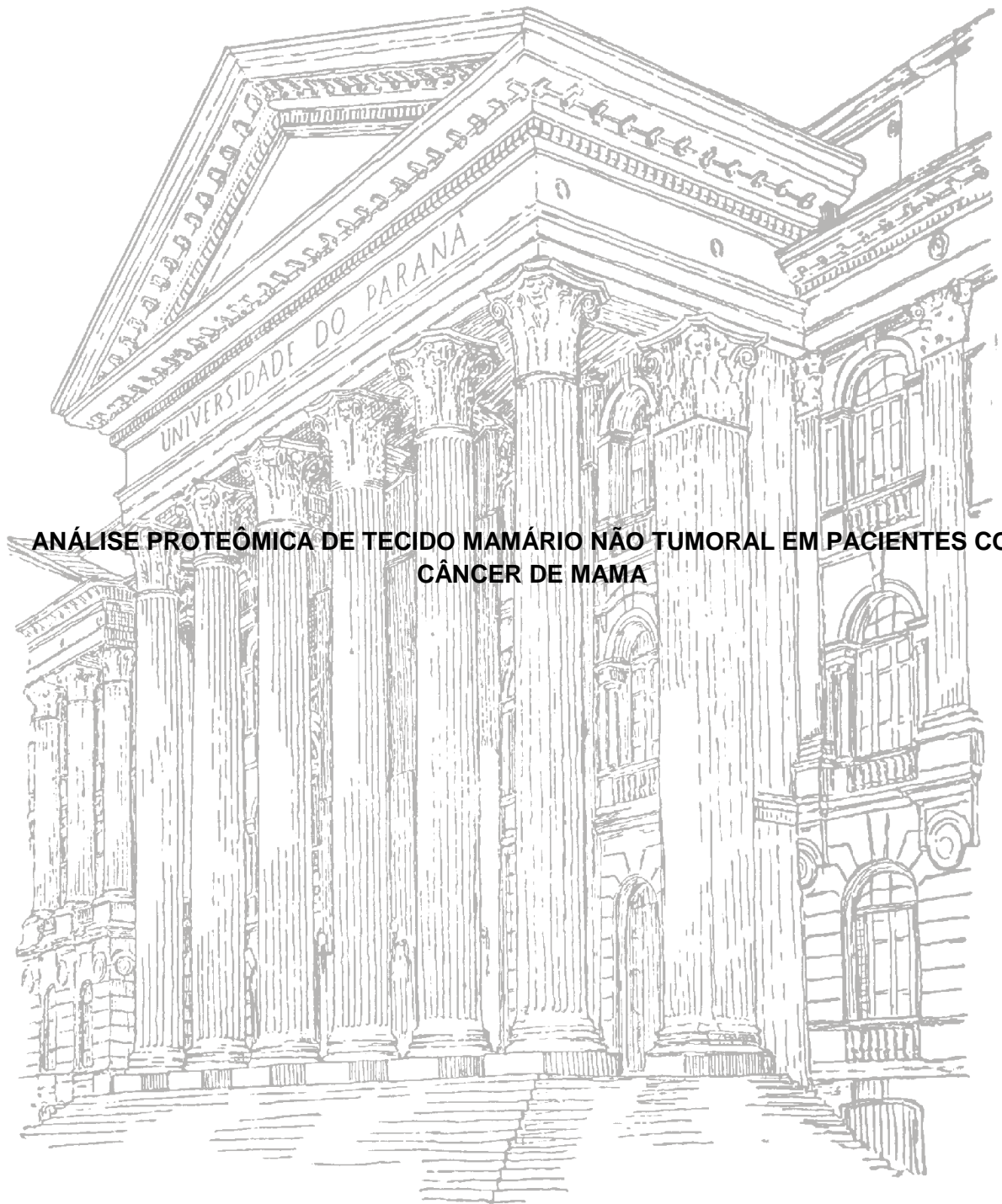


**UNIVERSIDADE FEDERAL DO PARANÁ**

**GUSTAVO GÓES DA COSTA**



**ANÁLISE PROTEÔMICA DE TECIDO MAMÁRIO NÃO TUMORAL EM PACIENTES COM  
CÂNCER DE MAMA**

**CURITIBA**

**2010**

GUSTAVO GÓES DA COSTA

**ANÁLISE PROTEÔMICA DE TECIDO MAMÁRIO NÃO TUMORAL EM PACIENTES COM  
CÂNCER DE MAMA**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Genética, Departamento de Genética, Setor de Ciências Biológicas, Universidade Federal do Paraná, para obtenção do título de Mestre em Genética.

Orientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Enilze Maria de Souza  
Fonseca Ribeiro

Co-orientador: Prof. Dr. Iglénir João Cavalli

CURITIBA

2010

Universidade Federal do Paraná  
Sistema de Bibliotecas

Costa, Gustavo Góes da  
Análise proteômica de tecido mamário não tumoral em pacientes com  
câncer de mama. / Gustavo Góes da Costa. – Curitiba, 2010.  
102 f.: il. ; 30cm.

Orientadora: Enilze Maria de Souza Fonseca Ribeiro  
Co-orientador: Iglénir João Cavalli

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

1. Mamas - Cancer I. Título II. Ribeiro, Enilze Maria de Souza Fonseca  
III. Cavalli, Iglénir João IV. Universidade Federal do Paraná. Setor de  
Ciências Biológicas. Programa de Pós-Graduação em Genética.

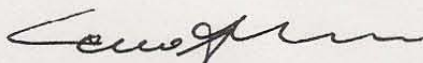
CDD (20. ed.) 616.99449

## PARECER

Os abaixo-assinados, membros da Banca Examinadora da Defesa de Dissertação de Mestrado, a qual se submeteu **GUSTAVO GÓES DA COSTA** para fins de obter o título de Mestre em Ciências Biológicas na Área de Genética pela Universidade Federal do Paraná, no Programa de Pós-Graduação em Genética, são de parecer que se confira ao candidato o conceito "A".

Secretaria da Coordenação do Programa de Pós-Graduação em Genética do Setor de Ciências Biológicas da Universidade Federal do Paraná.

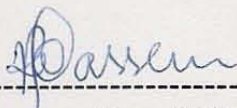
Curitiba, 30 de abril de 2010



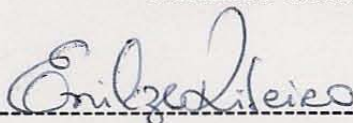
Professor Doutor Cícero Andrade Urban da Universidade Positivo  
Membro Titular



Professor Doutor Ricardo Lehtonen R. Souza - UFPR  
Membro Titular

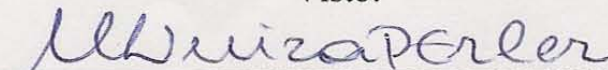


Professora Doutora Roseli Wassem - UFPR  
Membro Titular



Professora Doutora Enilze Ribeiro - UFPR  
Orientadora e Presidente da Banca Examinadora

Visto:



Professora Doutora Maria Luiza Petzl-Erler  
Coordenadora do Programa de Pós Graduação em Genética

**Dedico este trabalho aos meus pais, *Fernando* e *Silvana*. A força para chegar até aqui e a vontade de ir ainda mais longe são qualidades que herdei de vocês!”**

## **AGRADECIMENTOS**

Aos meus orientadores, Prof<sup>a</sup>. Dr<sup>a</sup>. Enilze Ribeiro e ao Prof. Dr. Iglénir João Cavalli, pela oportunidade de trabalhar em algo novo e pela amizade durante estes dois anos. Obrigado por todo o conhecimento compartilhado e por sempre acreditar que seríamos (eu e o Rodrigo) capazes de realizar este trabalho.

À minha banca de acompanhamento e de defesa, Prof<sup>a</sup>. Dr<sup>a</sup>. Roseli Wassem, Prof. Dr. Emanuel Maltempi de Souza, Prof. Dr. Silvio Marques Zanata, Dr. Cícero Urban, Prof. Dr. Ricardo Lehtonen R. de Souza. Obrigado pela disponibilidade, pelas sugestões e correções. A participação de vocês foi fundamental para a conclusão e qualidade deste trabalho.

Ao doutorando e amigo Luiz Eduardo Rizzo de Souza, pela dedicação e envolvimento com este trabalho e por todo conhecimento transmitido.

Aos amigos laboratoriais, pelo apoio e por tornarem o laboratório um ótimo ambiente para trabalhar. Em especial ao Márcio, Marcos e Rodrigo (em ordem alfabética para que vocês não briguem). A amizade de vocês foi muito importante durante estes anos.

A todos meus amigos e amigas pela força e amizade sincera.

Aos meus pais, Fernando e Silvana, pelo amor, educação, confiança, e apoio emocional e financeiro! Aos meus irmãos, Guilherme (que me deu uma sobrinha linda!) e Gregório, que mesmo não morando na mesma cidade (ou continente) que eu, estão sempre presentes em minha vida. A toda minha família (avós, tios, primos...)

À Luana, pelo amor, companheirismo, por aturar minhas maluquices e me acompanhar na torcida pelo Mengão!

A todos os indivíduos e instituições envolvidos nesta pesquisa, sobretudo aos Drs. Rubens Lima e Cícero Urban, aos funcionários da Oncoclínica de Curitiba, do Hospital de Clínicas, Hospital Nossa Senhora das Graças, e à CAPES e CNPq pelo apoio financeiro.

“O difícil a gente tira de letra, o impossível a gente demora um pouquinho!”

*Autor desconhecido*



## RESUMO

O tecido mamário é controlado por um sistema complexo, envolvendo hormônios, que atuam através de seus receptores, e diversos fatores locais. As estruturas anatômicas do tecido mamário podem ser modificadas devido a fatores como a multiparidade, idade, estado pré-menstrual e terapia de reposição hormonal. Portanto, compreender os fatores e mecanismos que regulam as mudanças dependentes de hormônios no tecido mamário é crucial, pois alterações na estrutura da mama e suas funções durante o ciclo menstrual podem predispor este tecido ao desenvolvimento de doenças mamárias, incluindo o câncer. O câncer de mama é o tipo mais comum de câncer em mulheres. É o quinto tipo de câncer em mortalidade no mundo, devido principalmente pela detecção no estágio avançado da doença. A abordagem proteômica é um campo promissor que possui diversas aplicações potenciais, incluindo a identificação de biomarcadores para o diagnóstico, o prognóstico e a terapêutica. O objetivo deste estudo foi a análise proteômica de tecido mamário não tumoral de seis pacientes acometidas pelo câncer de mama, obtendo informações para a caracterização deste tecido e fornecendo dados para a análise comparativa com tecido de carcinomas mamários. A análise dos géis bidimensionais foi realizada no programa *ImageMaster™ 2D Platinum v6.0*. Setenta bandas foram removidas manualmente, digeridas *in gel* e submetidas à identificação por PMF na plataforma MASCOT. Esta identificação resultou em 44 proteínas diferentes distribuídas em nove categorias: Citoesqueleto e proteínas associadas (29%); Proteínas com funções de ligação (27%); Enzimas metabólicas (11%); Chaperonas moleculares / proteínas heat shock (7%); Detoxificação e proteínas redox (5%); Proteínas membrana-associadas com múltiplas atividades (5%); Crescimento celular e reguladores de proliferação (5%); Degradação protéica (2%); e Proteínas com outras funções (9%). Algumas destas proteínas têm função no controle da proliferação celular e adesão, podendo atuar na tumorigênese quando sua expressão é alterada. Portanto, a análise comparativa não deve se limitar ao estudo das proteínas expressas isoladamente em um dos tecidos, mas também considerar a variação da sua expressão. Estas informações são importantes para a caracterização do proteoma do tecido mamário não tumoral e posterior análise comparativa com tecido de carcinoma mamário.

**Palavras-Chave:** Tecido mamário não tumoral. Proteoma. 2D-PAGE.

## ABSTRACT

The mammary tissue is controlled by a complex system ruled by hormones acting through its receptors and different local elements. The anatomic structures of the mammary tissue can be modified by factors as multiparity, age, premenstrual state and hormonal replacement therapy. Therefore, it's important to understand the factors and mechanisms that regulate crucial hormone alterations in the mammary tissue. The breast structure and functions can change during the menstrual cycle therefore predisposing the tissue to mammary disease, including cancer. Breast cancer is the most common type of cancer in women. This type is the number five leading cause of death per cancer in the world, mainly because of its late detection in an advanced disease stage. The proteomic approach is a promising field for the diagnostic, prognostic and therapeutics of cancer. The objective of this study was to carry out a proteomic analysis of non-tumoral mammary tissue of six patients with breast cancer, aiming to provide information for the tissue characterization and data for the comparative analysis with mammary carcinoma tissue. The two-dimensional gel analysis was performed using the software *ImageMaster™ 2D Platinum v6.0*. Seventy bands were manually removed, digested *in gel* and submitted to PMF identification using the MASCOT platform. This identification resulted in 44 different proteins distributed in nine categories. Cytoskeleton-associated protein (29%); binding proteins (27%), metabolic enzymes (11%); molecular chaperones/ heat shock proteins (7%); detoxifying and redox proteins (5%); membrane associated proteins with multiple activities (5%); degradation proteins (2%) and proteins with other functions (9%). Some of these proteins are involved in the control of cellular proliferation and adhesion, promoting tumorigenesis when its expression is altered. For that reason the comparative analysis can't be limited to the study of isolated expressed proteins in one tissue, it's also necessary to consider the variation of its expression. This information is important for the proteomics characterization of the non-tumoral tissue and posterior comparative analysis with the mammary carcinoma tissue.

KEY WORDS: Non-tumoral mammary tissue, proteomics, 2D-PAGE.

## LISTA DE ILUSTRAÇÕES

FIGURA 1 – Anatomia da mama feminina.....	13
FIGURA 2 – Fatores de controle do desenvolvimento mamário.....	15
FIGURA 3 – Esquema de funcionamento do MALDI-Tof-MS.....	25
QUADRO 1 – Informações sobre a amostra.....	28
QUADRO 2 – Programa para o Ettan IPGphor II (tira 13 cm; pH 4-7).....	30
QUADRO 3 – Dosagem de proteínas.....	35
QUADRO 4 – Quantidade de bandas detectadas nos géis bidimensionais.....	36
FIGURA 4 – Triplicata de géis bidimensionais da amostra CP632 N.....	37
FIGURA 5 – Géis de referência.....	39
QUADRO 5 – Relação das bandas (n=118) detectadas em pelo menos três amostras.....	42
FIGURA 6 – Espectros de massas.....	44
FIGURA 7 – Bandas identificadas identificadas no gel <i>Master</i> .....	47
FIGURA 8 – Proteínas identificadas e sua localização no gel <i>Master</i> .....	52
GRÁFICO 1 – Classificação das proteínas identificadas.....	53
FIGURA 9 – Detalhe das bandas da proteína vimentina.....	62
FIGURA 10 – Detalhe das bandas da proteína alfa-1 antitripsina.....	66
FIGURA 11 – Detalhe das bandas da proteína sorotransferrina.....	67
FIGURA 12 – Detalhe das bandas da proteína albumina sérica.....	72
FIGURA 13 – Detalhe das bandas da cadeia alfa-1 do colágeno VI.....	81

## LISTA DE TABELAS

TABELA 1 –	Bradford do tampão de lise.....	36
TABELA 2 –	Identificações obtidas por MS e dados referentes aos critérios considerados.....	48

## LISTA DE ABREVIATURAS E SIGLAS

2D - Bidimensional  
2D-PAGE - eletroforese bidimensional em gel de poliacrilamida  
BSA - Albumina de soro bovino  
CHAPS - Ciclohexilamino dimetilamônio propano sulfonato  
Da - Dalton  
DNA - Ácido Desoxirribonucléico  
DTT - Ditioneitol  
ESI - Ionização por eletronspray  
g - Grama  
GTP - guanosina trifosfato  
GTPase - enzimas que hidrolisam GTP  
h - Hora  
HCCA - Ácido  $\alpha$ -ciano-4-hixiamínico  
IEF - Focalização isoeétrica  
INCA - Instituto Nacional do Câncer  
IPG - Gradiente de pH imobilizado  
kDa - Quilodalton  
kV - Quilovolts  
kV/h – Quilovolt/ hora  
LMW – Marcador de massa molecular  
m/z - Razão massa carga  
mA - Miliampére  
MALDI - Dessorpção por *laser* ionizante assistida por uma matriz  
mg - Miligrama  
min - Minuto  
ml - Mililitro  
mM - Milimolar  
MM - Massa molecular  
MPT - Modificação pós-traducional  
MS - Espectrometria de massa  
MS/MS - Espectrometria de massa em *tandem*  
ng - Nanograma  
ns - Nano segundos  
pI - Ponto isoeétrico  
PMF - Impressão digital de peptídeos  
PMSF - Fluoreto de fenil-metil-sulfonilo  
ppm – Partes por milhão  
mRNA - Ácido Ribonucléico mensageiro  
SDS - Dodecil-sulfato de sódio  
SDS-PAGE - Eletroforese em gel de poliacrilamida na presença de dodecil-sulfato de sódio  
ToF - Tempo de vôo  
 $\mu$ l - Microlitro

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	12
<b>2. REVISÃO DA LITERATURA</b> .....	13
2.1. A ESTRUTURA DO TECIDO MAMÁRIO.....	13
2.1.1. Influência Hormonal no Desenvolvimento Mamário.....	14
2.2. CÂNCER E O CÂNCER DE MAMA.....	17
2.2.1. Fatores de Risco no Câncer de Mama.....	19
2.3. DA ERA GENÔMICA PARA A PROTEÔMICA.....	20
2.4. PROTEÔMICA NA BUSCA POR BIOMARCADORES.....	22
2.5. ELETROFORESE BIDIMENSIONAL E ESPECTROMETRIA DE MASSAS.....	23
<b>3. JUSTIFICATIVA</b> .....	26
<b>4. OBJETIVOS</b> .....	27
4.1. GERAL.....	27
4.2. ESPECÍFICOS.....	27
<b>5. MATERIAL E MÉTODOS</b> .....	28
5.1. CARACTERIZAÇÃO DA AMOSTRA .....	28
5.2. PROCESSAMENTO DAS AMOSTRAS.....	28
5.2.1. Coleta do Material Biológico e Armazenamento.....	28
5.2.2. Extração de Proteínas.....	29
5.3. DOSAGEM DE PROTEÍNAS.....	29
5.4. ELETROFORESE BIDIMENSIONAL (2D-PAGE).....	29
5.4.1. Isoeletofocalização.....	29
5.4.2. Equilíbrio das Tiras de IPG.....	30
5.4.3. Eletroforese em Gel de Poliacrilamida na Presença de SDS (SDS-PAGE)...	30
5.5. COLORAÇÃO, DIGITALIZAÇÃO E ANÁLISE COMPUTACIONAL.....	31
5.6. ESPECTROMETRIA DE MASSAS.....	32
5.6.1. Preparo das Amostras.....	32
5.6.2. Purificação Peptídica.....	32

5.6.3. Obtenção e Análise dos Espectros de Massas.....	33
5.6.4. Identificação das Proteínas.....	33
<b>6. RESULTADOS.....</b>	<b>35</b>
6.1. EXTRAÇÃO E QUANTIFICAÇÃO DAS PROTEÍNAS.....	35
6.2. ANÁLISE DOS GÉIS BIDIMENSIONAIS.....	36
6.3. PURIFICAÇÃO PEPTÍDICA E ESPCTROMETRIA DE MASSAS.....	43
6.4. IDENTIFICAÇÃO PROTÉICA.....	43
<b>7. DISCUSSÃO .....</b>	<b>54</b>
7.1 IMPLANTAÇÃO DA METODOLOGIA PROTEÔMICA.....	55
7.2. QUANTIFICAÇÃO PROTÉICA.....	56
7.3. ESPECTRO DE MASSAS.....	57
7.4. PROTEÍNAS IDENTIFICADAS.....	57
7.4.1. Citoesqueleto e Proteínas Associadas.....	59
7.4.2. Proteínas com Funções de Ligação.....	65
7.4.3. Enzimas Metabólicas.....	72
7.4.4. Chaperonas Moleculares / Proteínas <i>Heat Shock</i> .....	73
7.4.5. Crescimento Celular e Reguladores de Proliferação.....	76
7.4.6. Detoxificação e Proteínas redox.....	77
7.4.7. Proteínas Membrana-associadas com Múltiplas Atividades.....	78
7.4.8. Degradação Protéica.....	79
7.4.9. Proteínas com Outras Funções.....	80
<b>8. CONCLUSÃO.....</b>	<b>83</b>
<b>REFERÊNCIAS.....</b>	<b>84</b>
<b>APÊNDICE.....</b>	<b>98</b>
<b>ANEXOS.....</b>	<b>100</b>

## 1. INTRODUÇÃO

O tecido mamário é uma estrutura complexa, que apresenta heterogeneidade celular, controlada pela ação hormonal e por vários fatores locais (COURTILLOT, 2005). A compreensão dos fatores e mecanismos que regulam as mudanças dependentes de hormônios no tecido mamário é fundamental, pois alterações na estrutura da mama e suas funções durante o ciclo menstrual podem predispor este tecido ao desenvolvimento de doenças mamárias (NAVARRETE *et al.*, 2005).

Segundo a Organização Mundial da Saúde (OMS, 2010), o câncer é, em todo o mundo, uma das principais preocupações na área de saúde pública.

O câncer de mama é o quinto tipo de câncer com maior mortalidade no mundo e o segundo tipo mais incidente entre as mulheres. Muitos cânceres podem ser prevenidos evitando-se a exposição a fatores de risco e uma proporção significativa pode ser curada por procedimentos terapêuticos como a cirurgia, radioterapia ou quimioterapia, especialmente se detectados precocemente (OMS, 2010).

A abordagem proteômica surge como um campo promissor na caracterização de tecidos, mas devido às alterações hormonais durante o ciclo menstrual, o perfil protéico do tecido mamário também está sendo alterado e, portanto, torna-se difícil definir um padrão (HONDERMARCK *et al.*, 2008). A análise do tecido mamário não tumoral é também essencial na busca por biomarcadores do câncer de mama. Gradualmente torna-se aceitável acreditar que a análise direta das proteínas possa ser mais produtiva para um melhor conhecimento do câncer (HONDERMARCK *et al.*, 2002).

Baseado nas informações apresentadas, este trabalho busca, através da análise proteômica de tecido mamário não tumoral, auxiliar na caracterização do proteoma deste tecido e fornecer dados para futura análise comparativa.



## 2. REVISÃO DA LITERATURA

### 2.1. ESTRUTURA DO TECIDO MAMÁRIO

A glândula mamária humana (FIGURA 1) é composta por um epitélio ducto-lobular ramificado (TOILLON *et al.*, 2007). Este epitélio secretor (parênquima) é envolvido por um estroma de tecido conectivo e adiposo e uma rede de capilares (NANDI, 1995). Este sistema é delimitado por uma camada interna de células epiteliais luminais e uma camada externa de células mioepiteliais separadas do estroma intersticial por uma membrana basal íntegra (TOILLON *et al.*, 2007). As células luminais são glandulares e polarizadas (TOILLON *et al.*, 2007), responsáveis por formar a camada interior dos ductos e alvéolos e pela síntese e passagem do leite (NANDI, 1995). Apresentam domínios de membrana apical e basolateral que expressam respectivamente sialomucina e moléculas de adesão célula-célula (TOILLON *et al.*, 2007). As células mioepiteliais contribuem significativamente para a formação da membrana basal e a sua diferenciação miogênica é responsável pela função contrátil que permite a ejeção do leite (TOILLON *et al.*, 2007). Células luminais e mioepiteliais, juntamente com a porção estromal, produzem os componentes da matriz extracelular, que assim como hormônios e fatores de crescimento são essenciais para o funcionamento e morfogênese da glândula mamária (NANDI, 1995).

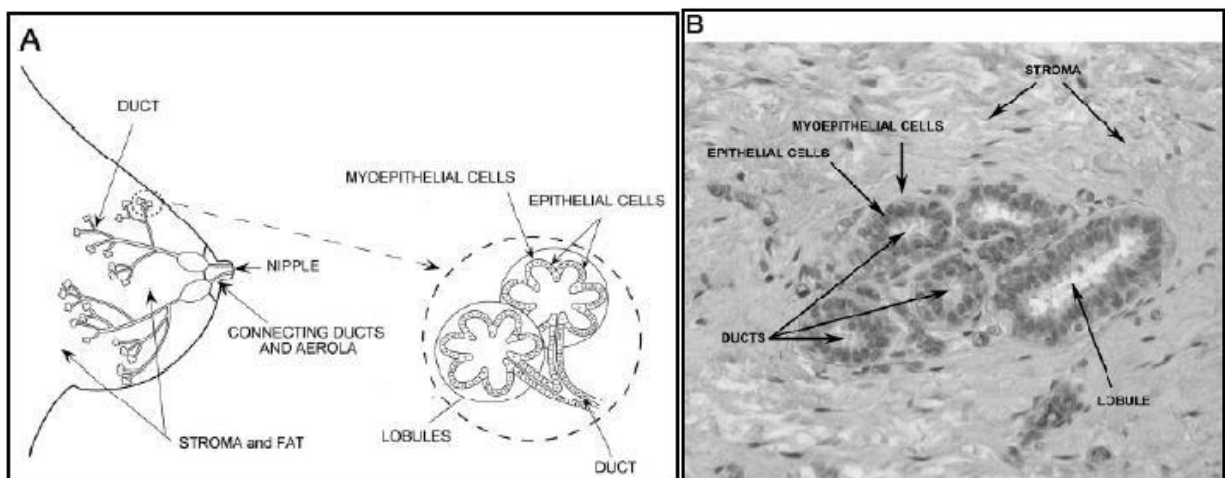


FIGURA 1 – ANATOMIA DA MAMA FEMININA

FONTE: Hondermack (2003)

LEGENDA: Esquema anatômico (A) e corte histológico (B) da glândula mamária humana feminina.

Esse modelo de dois tipos celulares (epiteliais luminais e mioepiteliais) foi muito importante para a compreensão dos processos patológicos do tecido mamário humano nas últimas décadas. Entretanto, existe um acúmulo de evidências que indicam que o mesmo é muito simplista e que a mama apresenta uma estrutura celular mais complexa, com um sistema integrado de renovação celular do epitélio mamário, semelhante ao da próstata e da epiderme. Recentemente, estudos de ultraestrutura, imunohistoquímica e expressão gênica identificaram um terceiro tipo celular na mama, localizado tanto nas camadas basais quanto nas luminais (BÁNKFALVI *et al.*, 2004). Baseada nessas informações é proposta a existência de células tronco epiteliais mamárias responsáveis pela formação de novas estruturas durante o desenvolvimento da mama (CLARKE e FULLER, 2006). Evidências indicam que essas células têm capacidade de originar células luminais e mioepiteliais (DONTU *et al.*, 2003; ALLRED *et al.*, 2004).

Em humanos, o estroma intersticial corresponde a mais de 80% do volume da mama. Existem quatro tipos principais de estroma na mama: tecido adiposo; tecido conectivo intersticial/interlobular denso; tecido conectivo intralobular frouxo; e vasos sanguíneos. O tecido adiposo é composto quase totalmente de células contendo grande quantidade de lipídeos, envolvidas por membrana basal e uma rede de capilares. A quantidade total de estroma e a proporção de estroma intersticial fibroso em relação ao tecido adiposo variam bastante de indivíduo para indivíduo (RONNOV-JESSEN *et al.*, 1996).

### 2.1.1. Influência Hormonal no Desenvolvimento Mamário

A mama humana é um órgão complexo, o qual tem sua proliferação e diferenciação reguladas pela ação combinada de fatores de crescimento, hormônios esteróides e interação célula-célula (PAGE, *et al.*, 1999). As principais moléculas que realizam esse controle são: fatores juxtácrinos, situados na superfície de uma célula e que influenciam as células adjacentes por contato direto; hormônios parácrinos, que são secretados por uma célula para influenciar outros tipos celulares adjacentes; e hormônios autócrinos, que atuam na própria célula secretora através de receptores intra ou extracelulares (FIGURA 2). A interação desses hormônios e outros fatores locais derivados (como proteínas de adesão celular) produzem as vias

de sinalização que são responsáveis pela regulação e estímulo celular dos tecidos mamários (COURTILLOT, 2005).

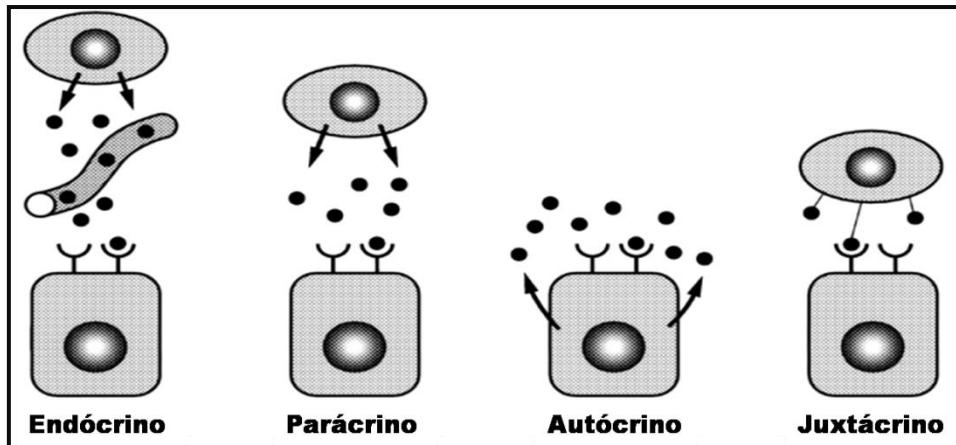


FIGURA 2 – FATORES DE CONTROLE DO DESENVOLVIMENTO MAMÁRIO  
 FONTE: Modificado de ABE, 2000

Os ovários produzem três classes de esteróides sexuais: estrógenos, progesterona e andrógenos. O mecanismo pelo qual os hormônios esteróides estimulam a proliferação do epitélio mamário ainda é controverso e está sobre intensa investigação. Estudos sugerem que a regulação da proliferação celular ocorra por três mecanismos principais: mediado por receptor; *loop* autócrino/parácrino; e *feedback* negativo (MUTI, 2005; NAVARRETE *et al.*, 2005).

Além dos hormônios, diversas famílias de fatores de crescimento já foram descritas como importantes para a regulação do desenvolvimento mamário, como a do fator de crescimento epidermal/ fator de crescimento transformador- $\alpha$  (EGF/TGF- $\alpha$ ), a do fator de crescimento semelhante à insulina (IGF) e a do fator de crescimento de fibroblastos (FGF) (NANDI *et al.*, 1995).

As células epiteliais respondem às flutuações hormonais do ciclo menstrual através de proliferação, diferenciação ou morte celular programada (apoptose). Esta resposta epitelial é limitada a uma pequena fração de células, sugerindo também a existência de regulação local neste tecido. Fatores como a idade, o uso de contraceptivos orais e a gravidez, além da influência da fase do ciclo menstrual, modificam a proliferação e a apoptose no tecido mamário (NAVARRETE *et al.*, 2005).

O desenvolvimento da glândula mamária ocorre em estágios definidos (embrionário, pré-puberdade, puberdade, gestação, lactação e involução) que estão ligados ao desenvolvimento sexual e reprodutivo (HENNIGHAUSEN e ROBINSON, 2001).

A principal característica do desenvolvimento mamário entre os 15 e 25 anos de idade é a adição de estruturas lobulares ao sistema de ductos já formado. Até aproximadamente os 35 anos de idade, a fase lútea do ciclo menstrual está associada com uma proliferação de ácinos e lóbulos, estimulada pela produção de estrogênios pelos ovários. O ciclo de mudanças (proliferação e involução das estruturas mamárias devido à menstruação e gravidez), que ocorre entre os 15 e 50 anos de idade, geram muitas oportunidades para que ocorram pequenas disfunções (HUGHES, 2000; COURTILLOT, 2005). A puberdade é a fase onde ocorre o rápido crescimento das glândulas mamárias, o qual é desencadeado pelo aumento nos níveis de estrogênio e progesterona. A glândula mamária sofre uma expansão substancial da população celular do epitélio durante este período (KIM *et al.*, 2008). Kim e colaboradores (2008) evidenciaram a expressão proteica diferencial entre ratas no estágio pré-puberdade e pós-puberdade.

O risco de que ocorram disfunções no tecido mamário é aumentado também pela involução concomitante que ocorre por 20 anos, entre os 35 e 55 anos de idade. Essa involução afeta os lóbulos e é altamente dependente da relação entre o epitélio e o estroma especializado do lóbulo (HUGHES, 2000). Durante o processo de involução lobular, os tecidos conectivos intralobulares com resposta a hormônios são perdidos e substituídos por um tecido intralobular fibroso. Finalmente, ao término da menopausa, a involução ocorreu extensivamente, restando apenas alguns ductos e poucos (ou nenhum) lóbulos (COURTILLOT, 2005). Após a menopausa, também devido à carência hormonal, ocorre atrofia glandular e a substituição do tecido parenquimatoso por gordura (INCA, 2000).

Embora a involução mamária ocorra por uma consistente sequência de mudanças histológicas, a taxa e a extensão da involução podem variar consideravelmente entre as mulheres. Este processo já foi documentado em mulheres com menos de 40 anos de idade e, portanto, os fatores envolvidos não são limitados ao início da menopausa. Dentre os possíveis fatores que podem influenciar o início e a taxa da involução estão o histórico de amamentação e de gravidez da mulher (MILANESE *et al.*, 2006).

As estruturas anatômicas do tecido mamário podem ser modificadas devido a fatores como a multiparidade, idade, estado pré-menstrual e terapia de reposição hormonal (BUDEL, 2007). Portanto, compreender os fatores e mecanismos que regulam as mudanças dependentes de hormônios no tecido mamário é crucial, pois alterações na estrutura da mama e suas funções durante o ciclo menstrual podem predispor este tecido ao desenvolvimento de doenças mamárias, incluindo o câncer (NAVARRETE *et al.*, 2005).

## 2.2. CÂNCER E O CÂNCER DE MAMA

O câncer, em geral, não é uma doença única, mas designa as formas mais agressivas de neoplasia, um processo de doença caracterizado por uma proliferação celular descontrolada, que resulta em uma massa ou tumor (neoplasma), mesmo após ter cessado o estímulo que o gerou (NUSSBAUM *et al.*, 2002).

Apesar de ser responsável por grande número de mortes no mundo, o câncer não é uma doença da vida moderna. A evidência mais antiga data de 8.000 a.C e as primeiras descrições de tumores foram encontradas em papiros do Egito, e datam de 1.600 a.C (AAPC, 2010).

O envelhecimento é um fator fundamental no desenvolvimento do câncer que, apesar de ocorrer em todas as idades, alcança seu pico na faixa etária de 65 a 70 anos (INCA, 2010). Por isso, devido à melhoria das condições de saúde em geral e o avanço da medicina, fatores importantes do aumento da expectativa de vida populacional, o câncer passou a ser considerado com maior relevância nas últimas três décadas.

Estudos epidemiológicos indicam que os fatores ambientais são responsáveis pela maioria dos casos de câncer, sendo que lesões no DNA estão consistentemente associadas ao risco do desenvolvimento de neoplasias (SMITH *et al.*, 2003; CRESPI e SUMMERS, 2005; CEBRIANT *et al.*, 2006). As alterações pelas quais uma célula normal torna-se tumoral são o resultado de interações entre os fatores genéticos e carcinógenos, físicos, químicos ou biológicos (OMS, 2010).

Um tumor é considerado benigno quando as células neoplásicas permanecem agregadas formando uma massa única. Um tumor é considerado um câncer apenas se for maligno, isto é, somente se suas células tiverem adquirido a

capacidade de invadir tecidos adjacentes (ALBERTS *et al.*, 2004). Neste último caso, além da autonomia proliferativa, as células apresentam também perda da coesão. Isto permite a mobilidade celular tornando-as invasivas (MAREEL e LEROY, 2003).

Capacidade de invasão geralmente implica na habilidade de entrar na corrente sangüínea ou vasos linfáticos e formar tumores secundários, ou metástases, em outros locais do organismo. Quanto mais metástases um câncer for capaz de induzir, mais difícil será a sua erradicação (ALBERTS *et al.*, 2004).

As principais anormalidades genéticas observadas no câncer incluem uma expressão aumentada de oncogenes, inativação de genes supressores de tumor, instabilidades cromossômicas, alterações nos genes de reparo de DNA, alterações epigenéticas (GARNIS *et al.*, 2004), e disfunção da telomerase (GILLEY *et al.*, 2005).

O câncer de mama é o segundo tipo de câncer mais freqüente no mundo e o mais comum entre as mulheres. A cada ano, cerca de 22% dos casos novos de câncer em mulheres são de mama. O número de casos novos de câncer de mama esperados para o Brasil, no ano de 2010, é de 49.240, com uma ocorrência estimada de 49 casos a cada 100 mil mulheres (INCA, 2010).

Admite-se que, o câncer de mama é o resultado da interação de fatores genéticos com o estilo de vida, hábitos reprodutivos e meio ambiente (JOHNSON-THOMPSON e GUTHRIE, 2000). Acredita-se que 90-95% dos cânceres sejam esporádicos (não-familiares) ou decorram de mutações somáticas que se acumulam durante a vida (BALMAIN *et al.*, 2003), e que 5-10% sejam hereditários, explicados por mutações germinais raras em genes de alta penetrância como o *BRCA1* e o *BRCA2* (NUSSBAUM *et al.*, 2002). Cerca de 30% de mulheres jovens que desenvolvem câncer de mama devem apresentar um padrão genético de predisposição à doença (DUFLOTH *et al.*, 2005).

Apesar de ser considerado um câncer de relativamente bom prognóstico, se diagnosticado e tratado precocemente, as taxas de mortalidade por câncer de mama continuam elevadas no Brasil (INCA, 2010). Isto se deve ao fato de que aproximadamente 60% dos casos são identificados em estágios avançados, de acordo com o Instituto Brasileiro de Controle do Câncer (IBCC, 2010). Na população mundial a sobrevida média após cinco anos do diagnóstico é de 61% (INCA, 2010),

sendo que um terço dos casos de câncer no mundo poderia ser evitado através de ações preventivas, como mudanças no hábito de vida (IARC, 2010).

Os métodos mais usados para a detecção de tumores de mama, benignos ou malignos são baseados na mamografia. Contudo, há sugestões de que raios-X podem potencialmente induzir a carcinogênese e é certo de que para ser detectado na mamografia o tumor precisa de um tamanho mínimo de alguns milímetros, sendo que neste estágio já seria constituído por aproximadamente cem milhões de células (HONDERMARCK, 2003).

### 2.2.1. Fatores de Risco no Câncer de Mama

Os fatores de risco reprodutivos e hormonais aumentam o risco de câncer de mama em mulheres que apresentam menarca precoce (antes dos 11 anos de idade) e menopausa tardia (após os 50 anos de idade, em média), nuliparidade e idade da primeira gestação acima de 30 anos. O risco é aumentado devido à exposição a altas taxas de estrogênio endógeno (PARKIN *et al.*, 2001). Em uma revisão realizada por Collins e colaboradores (2005) foi relatado que o risco é maior em mulheres em tratamento de reposição hormonal na menopausa do que em mulheres não tratadas ou que interromperam o tratamento.

Morch e colaboradores (2005) evidenciaram que o consumo regular de álcool, mesmo que moderadamente, é identificado como fator de risco. Seus resultados mostraram que o consumo entre 22-27 doses por semana apresentou um aumento do risco em duas vezes para o desenvolvimento do câncer de mama, quando comparado às mulheres que consomem 1-3 doses semanais. Lash e Aschengrau (2000) acreditam que a explicação mais plausível seja que a ingestão de álcool modifique o metabolismo e remoção dos componentes do estrogênio, desta forma aumentando a exposição do tecido mamário a este hormônio.

A exposição ao tabaco é um dos fatores de risco mais estudados, mas apesar destas consideráveis pesquisas a relação entre câncer de mama e tabagismo permanece controversa. A fumaça do tabaco contém carcinógenos e metabólitos desta fumaça foram encontrados em fluidos da mama de fumantes. Contudo, o fumo também tem efeitos antiestrogênicos que, paradoxalmente, atuam na redução do risco de câncer de mama. A inconsistência na literatura pode ser

devido à heterogeneidade do risco em relação ao tempo de exposição, idade do diagnóstico ou susceptibilidade genética (REYNOLDS *et al.*, 2004). Mulheres regularmente expostas ao cigarro, mesmo fumantes passivas, apresentaram um aumento no risco de desenvolver câncer de mama, em oposição àquelas que nunca se expõem ao tabaco, num risco particularmente associado ao período premenopausa (JOHNSON, 2005). Essa relação foi observada principalmente em mulheres com histórico de hábitos tabagistas por longo período de tempo (TERRY *et al.*, 2002).

Como é característico em qualquer tipo de câncer, a idade continua sendo um dos mais importantes fatores de risco. O câncer de mama é relativamente raro antes dos 35 anos de idade, mas acima desta faixa etária sua incidência aumenta rápida e progressivamente (INCA, 2010). Isto se deve provavelmente ao maior tempo de exposição aos fatores de risco. Os mecanismos de reparo celular tendem também a ser menos efetivos à medida que a pessoa envelhece (OMS, 2010).

### 2.3. DA ERA GENÔMICA PARA A PROTEÔMICA

A aplicação da genômica na biologia do câncer possui um importante potencial para identificação de mecanismos que conduzem a malignidade e para o desenvolvimento de estratégias terapêuticas (MARTIN e NELSON, 2001).

A genômica tinha como objetivo primordial a completa e precisa elucidação da seqüência de DNA de um genoma haplóide representativo de uma determinada espécie (GRIFFITHS *et al.*, 2002). O seqüenciamento completo do genoma humano pelo Projeto Genoma Humano (HGP) e Celera só foi possível devido à melhoria das tecnologias para seqüenciamento (MARTIN e NELSON, 2001). Tornou-se evidente que a disponibilidade de genomas totalmente seqüenciados apenas estimulou a motivação científica por informações globais adicionais. A genômica foi, então, dividida em estrutural, caracterizando a natureza física dos genomas, e funcional, caracterizando o transcriptoma (toda a gama de transcritos produzidos por um determinado organismo) e o proteoma (toda a gama de proteínas codificadas) (GRIFFITHS *et al.*, 2002). O termo proteoma (derivado de proteína+genoma) foi proposto por Wilkins em 1995, como sendo todo o conteúdo de proteínas expressas por um genoma (ROCHA *et al.*, 2005).



O genoma foi o alvo primário das explorações pelas bases moleculares de doenças (AZAD *et al.*, 2006) já que os genes são o material genético essencial para o controle das características biológicas da célula, porém, são as proteínas que realizam funções celulares diversas e vitais, sendo que todos os comportamentos celulares em vida podem em geral ser interpretados no nível protéico (DENG *et al.*, 2006).

O genoma de um organismo apresenta um caráter permanente, pois é totalmente transferido durante a replicação celular. Ao contrário disso, o proteoma é extremamente dinâmico, sendo entendido como o conjunto de proteínas produzidas por uma célula e dependendo de fatores como: o estágio de desenvolvimento e diferenciação; e condições específicas e temporais do ambiente (BISCH, 2004). Recentemente, várias abordagens proteômicas têm sido utilizadas para identificar interações de proteínas com outras proteínas, ácidos nucleicos, lipídeos, anticorpos e drogas (GARNIS, 2004).

A proteômica tem potencial para complementar e ampliar as informações geradas pela genômica no conhecimento da fisiologia de tecidos normais e de várias doenças, entre elas, o câncer de mama. Os níveis de RNA mensageiro (mRNA) não correspondem necessariamente com a abundância das proteínas, pois estas sofrem modificações pós-traducionais, como fosforilações, acetilações, glicosilações ou clivagens, que aumentam a sua complexidade. Estas modificações não são detectadas no nível de mRNA, mas desempenham funções significantes na estabilidade protéica, localização, interações e mecanismos de ação (BERTUCCI *et al.*, 2006).

A glândula mamária e os tumores mamários são estruturas celulares heterogêneas e dinâmicas e, portanto, uma análise diferencial dos seus proteomas se refere ao estudo comparativo de misturas de células em diferentes proporções, como células epiteliais, mioepiteliais, endoteliais, fibroblastos, suplementadas por macrófagos, adipócitos e outras células circulantes. (HONDERMARCK *et al.*, 2008).

Há ainda o questionamento de o que seria um tecido de mama normal, devido às mudanças nas concentrações hormonais (estrógeno e progesterona) durante o ciclo menstrual. O estímulo hormonal será diferente dependendo do período do ciclo no qual for retirada a amostra (HONDERMARCK, *et al.*, 2008). A diversidade e a complexidade da glândula mamária permanecem, portanto, como os principais fatores limitantes à investigação proteômica diferencial.

## 2.4. PROTEÔMICA NA BUSCA POR BIOMARCADORES

O proteoma reflete tanto o programa genético intrínseco da célula quanto o impacto do meio ambiente e é, portanto valioso na busca por biomarcadores (SRINIVAS *et al.*, 2001).

O biomarcador pode ser definido como método, estrutura ou processo que se encontra na via causal, ou intimamente ligado a esta, entre o momento da exposição e a doença, em qualquer etapa do processo (BARTSCH, 2000). Mutações gênicas, alterações na transcrição e tradução e alterações nos produtos protéicos podem potencialmente ser utilizadas como biomarcadores específicos de doenças (WULFKUHLE *et al.*, 2003).

Estes biomarcadores para doenças (classificados como marcadores de efeito) podem ser agrupados em três classes principais: **diagnóstico**, que auxiliam na correta definição da doença, **prognóstico**, que fornecem informações sobre o curso clínico e **preditivo**, que podem antever o curso da doença e a melhor forma de tratamento (AZAD *et al.*, 2006).

O perfil proteômico é útil para a descoberta de biomarcadores, contudo, tem limitações intrínsecas: (1) devido à complexidade da amostra e a dificuldade de encontrar marcadores entre centenas a milhares de outras proteínas da amostra biológica; (2) não fornece (ou fornece apenas limitadamente) informação sobre proteínas específicas, como o papel biológico da proteína, exceto em um subproteoma em particular; (3) a proteômica em geral é uma metodologia cara que requer bastante tempo e não é compatível, para a análise, com uma grande quantidade de amostras clínicas (GUO *et al.*, 2007).

Além da proteômica de tecidos, outras abordagens têm sido propostas. O plasma, por exemplo, é uma excelente fonte de material clinicamente relevante para a detecção do câncer de mama em estágio inicial, pois é facilmente obtido, usado rotineiramente para diagnóstico de várias doenças e possui um proteoma muito rico. A única dificuldade é que a proteômica baseada no plasma é composta por proteínas originárias de vários tecidos (DAVIS e HANASH, 2006).

Existem centenas de candidatos a biomarcadores identificados, mas ainda necessitam de validação para provar sua especificidade e relevância clínica. É necessário que eles tenham uma eficácia comprovada e que não sejam apenas resultados de uma resposta adaptativa do organismo (SRINIVAS *et al.*, 2001). Para

o câncer de mama ainda nenhum biomarcador foi validado (DAVIS e HANASH, 2006).

O fato de até o momento nenhum marcador descrito pela proteômica ter se tornado um marcador de rotina na prática médica, não deve ser visto com pessimismo, mas um estímulo para novos estudos que descrevam os aspectos básicos do proteoma humano.

## 2.5. ELETROFORESE BIDIMENSIONAL E ESPECTROMETRIA DE MASSAS

A Eletroforese bidimensional em gel de poliacrilamida (2D-PAGE) é o método de referência para separação de mistura de proteínas complexas em larga escala e é considerado o mais prático e mais utilizado em proteômica (SOMIARI *et al.*, 2005). Conhecida desde os anos 70, a eletroforese bidimensional ganhou grande destaque apenas na década de 1990, após uma revolução silenciosa da química de proteínas e com o surgimento da proteômica (WILKINS *et al.*, 1997).

Tecidos e fluídos corporais são misturas complexas e precisam ser inicialmente separadas para uma análise posterior (AZAD *et al.*, 2006). Na 2D-PAGE, as proteínas intactas são separadas por dois parâmetros físico-químicos independentes (SOMIARI *et al.*, 2005). A primeira dimensão é baseada no ponto isoelétrico (pI) intrínseco através da focalização isoelétrica (IEF) e a segunda dimensão é baseada na massa molecular pela eletroforese em gel de poliacrilamida desnaturante (SDS-PAGE) (GUO *et al.*, 2007). O resultado final consiste em um gel com diversas bandas dispersas, cada uma correspondendo a uma proteína particular (ROCHA *et al.*, 2005), possibilitando a separação e identificação de centenas ou mesmo milhares de proteínas nestes géis (BISCH, 2004). Uma análise de imagem sofisticada é necessária para comparação de múltiplos géis e quantificação precisa de proteínas, que podem ser alcançados através de modernos algoritmos de comparação em programas de bioinformática (GUO, *et al.*, 2007).

Há alguns anos, a 2D-PAGE seguida pela identificação por espectrometria de massa (MS) tem sido a principal técnica para descoberta de biomarcadores em análises proteômicas convencionais. Esta técnica é destinada à identificação de proteínas diferencialmente expressas entre tecidos normais e tumorais em vários cânceres, inclusive no de mama (WULFKUHLE *et al.*, 2003). A MS é uma técnica

analítica que fornece alta sensibilidade e seletividade, sendo que o espectro pode ser produzido rapidamente dando informação sobre a massa molecular ou a estrutura de um composto (CAÑAS *et al.*, 2007).

O atual progresso na proteômica é em grande parte devido ao recente desenvolvimento das tecnologias baseadas em espectrometria de massa, particularmente as novas técnicas para a ionização de proteínas e peptídeos, tal como a MALDI (*matrix-assisted laser desorption-ionization*) combinada com TOF (*time-of-flight*). Estas técnicas estão agora se tornando ferramentas de escolha para caracterização de proteínas (ALAOUI-JAMALI e XU, 2006).

As bandas de interesse são recortadas do gel e submetidas a diversas etapas de lavagem, para depois sofrerem digestão por enzimas proteolíticas, como a tripsina, capazes de clivar a ligação peptídica em seqüências específicas (BISCH, 2004). Henzel *et al.*, em 1989, apresentaram o primeiro trabalho empregando o método por PMF (*peptide mass fingerprint*), um diferente conceito para identificar proteínas (citado por CAÑAS *et al.*, 2007). Os espectros de peptídeos resultantes da digestão por enzimas proteolíticas consistem em identificações únicas de proteínas (BISCH, 2004).

No espectrômetro MALDI-ToF-MS pequenos fragmentos ou peptídeos são cristalizados dentro de uma matriz sólida, a qual é bombardeada com um laser para ionizar os peptídeos. Os peptídeos ionizados são separados por ToF, onde partículas menores e com pouca carga passam através de um tubo de vôo mais rápido do que as partículas grandes e com maior carga, gerando um espectro de picos massa/carga (DAVIS e HANASH, 2006). As proteínas são analisadas e identificadas com base em suas assinaturas espectrométricas exclusivas (AZAD *et al.*, 2006). A FIGURA 3 ilustra as etapas da espectrometria de massas MALDI-ToF.

Todas as abordagens técnicas têm vantagens e limitações. Embora existam várias novas alternativas, sobretudo na tentativa de maior automação, apesar da dificuldade em separar proteínas com massa ou pontos iso-elétricos extremos e com baixa solubilidade (ALAOUI-JAMALI e XU, 2006), a flexibilidade e fácil manipulação garantem que a eletroforese 2D continuará sendo importante nos estudos proteômicos futuros (BISCH, 2004).

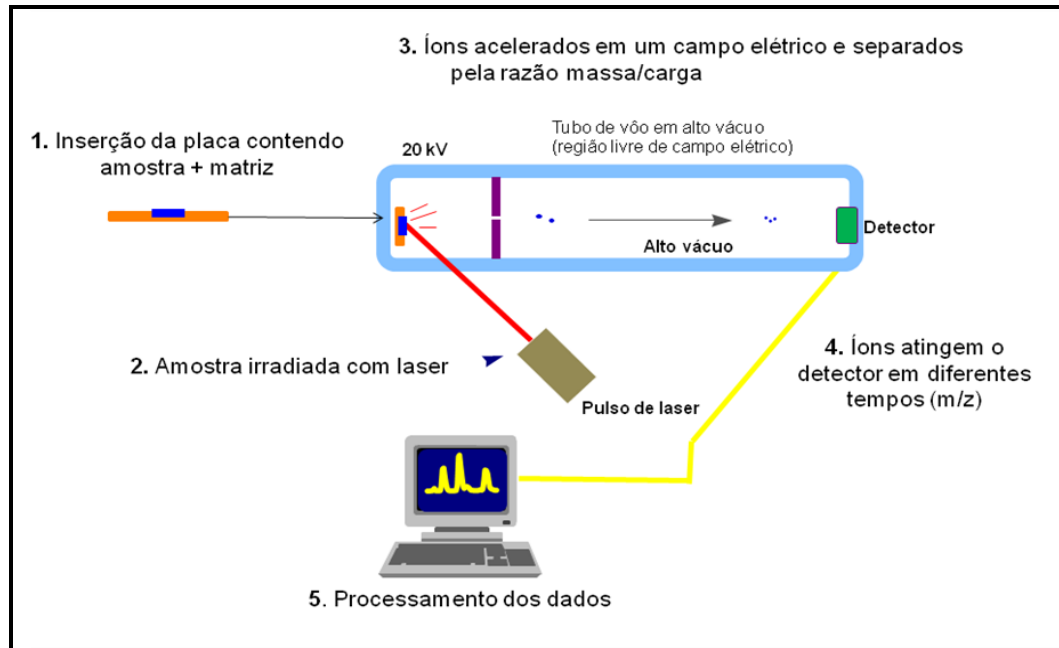


FIGURA 3 - ESQUEMA DE FUNCIONAMENTO DO MALDI-ToF-MS  
FONTE: O autor (2010)

### 3. JUSTIFICATIVA

A caracterização do proteoma do tecido mamário é uma das etapas iniciais para a compreensão da biologia da glândula mamária assim como de suas alterações, culminando como uma poderosa ferramenta na descoberta de novos marcadores de diagnóstico e progressão no câncer mamário.

O câncer de mama é atualmente o quinto tipo de câncer com maior mortalidade no mundo e o segundo em incidência no sexo feminino. Contudo, poderia ter um terço das mortes evitado se a doença fosse detectada no estágio inicial. A mamografia, principal exame para a detecção, não permite a observação de tumores iniciais (muito pequenos) e o diagnóstico tardio piora o prognóstico e dificulta o tratamento.

A identificação de biomarcadores proteicos que possam auxiliar no diagnóstico precoce em exames de rotina, como por exemplo, através de amostras de sangue ou urina, tem motivado pesquisadores de todo o mundo pelo alto potencial oferecido pelas análises proteômicas.

Os dados obtidos desta análise serão essenciais para a descrição de potenciais biomarcadores, pois se constituirá como referência na comparação com a análise proteômica do tecido tumoral.

## 4. OBJETIVOS

### 4.1. GERAL

Analisar amostras de tecido mamário não tumoral pelo método proteômico gerando informações que contribuam para a caracterização do proteoma deste tecido.

### 4.2. ESPECÍFICOS

- Implantar o estudo proteômico voltado para a genética humana no Laboratório de Citogenética Humana e Oncogenética do Departamento de Genética da UFPR;
- Detectar e identificar as proteínas mais freqüentes entre todas as amostras de tecido mamário não tumoral;
- Criar um banco de dados com as informações obtidas com as identificações para futura comparação com dados de tecido tumoral.

## 5. MATERIAIS E MÉTODOS

### 5.1. CARACTERIZAÇÃO DA AMOSTRA

Este estudo foi realizado com amostras de tecido mamário não tumoral, coletado no momento da cirurgia de seis pacientes portadoras de neoplasias mamárias (QUADRO 1), atendidas nos Hospitais Nossa Senhora das Graças e de Clínicas, Curitiba, nos anos de 2008 e 2009. A média de idade das pacientes foi de  $60,16 \pm 10,94$ . O projeto deste estudo foi aprovado pelo Comitê de Ética do HNSG, Curitiba - PR. Processo nº 25000.007020/2003-93; registro no CONEP: 7220 e parecer nº 251/2003, de 20/02/2003.

<b>CÓDIGO</b>	<b>IDADE</b>	<b>*TIPO HISTOLÓGICO</b>
CP622 N	71 anos	Carcinoma Ductal Invasor
CP632 N	61 anos	Carcinoma Ductal Invasor
CP633 N	71 anos	Carcinoma Ductal Invasor
CP638 N	48 anos	Carcinoma Lobular Invasor
CP645 N	64 anos	Carcinoma Ductal Invasor
CP655 N	46 anos	Carcinoma Ductal Invasor

QUADRO 1 - INFORMAÇÕES SOBRE A AMOSTRA

FONTE: O autor (2010)

NOTA: \*Tipo histológico de tumor que acometia a mama da qual foi retirada a amostra de tecido não tumoral.

### 5.2. PROCESSAMENTO DAS AMOSTRAS

#### 5.2.1. Coleta do Material Biológico e Armazenamento

Após a retirada do tumor de mama por excisão cirúrgica, uma amostra de tecido mamário não tumoral foi coletada de cada paciente. Este tecido foi retirado da mesma mama, contra lateralmente, respeitando a margem de segurança. Após a retirada, a amostra foi depositada em tubo rotulado e encaminhada ao Laboratório de Citogenética Humana e Oncogenética (UFPR). O tecido foi fragmentado com tesoura e pinça e alocado em tubos de 2ml, sendo armazenado a 80° C negativos.



Todo o procedimento foi realizado com material estéril, no fluxo laminar e com luvas sem talco.

### 5.2.2. Extração de Proteínas

Para extração e solubilização das proteínas, foi adicionado 1ml tampão de Lise (7M Uréia, 2M Tiuréia, 4% CHAPS, 40mM Tris e 0,2% inibidor de protease PMSF) ao tubo contendo a amostra. A desagregação do tecido e a homogeneização foram realizadas no próprio tubo com um homogeneizador elétrico. Durante todo o processo as amostras ficaram em banho de gelo para preservar a integridade do material.

## 5.3. DOSAGEM DE PROTEÍNAS

A quantificação da proteína total do extrato foi feita pelo método de Bradford (BRADFORD, 1976) que se baseia na ligação do azul brilhante de Coomassie com a proteína e comparação dessa ligação com diferentes concentrações de uma proteína padrão. A albumina de soro bovino (*Bovine Serum Albumine*, BSA) foi utilizada como padrão. Para o uso, o reagente de Bradford (Sigma) foi diluído 1:5 em água ultra pura.

## 5.4. ELETROFORESE BIDIMENSIONAL (2D-PAGE)

### 5.4.1. Isoeletrofocalização

A isoeletrofocalização (IEF), primeira dimensão, foi realizada no *Ettan IPGphor II* (GE Healthcare). Foram utilizadas tiras de IPG (*immobilized pH gel*) de 13cm de comprimento (130x3x0,5mm; GE Healthcare) com gradiente de pH de 4-7. A reidratação foi realizada no *Immobilize DryStrip Reswelling Tray* (GE Healthcare) com tampão de Reidratação (7M Uréia, 2M Tiuréia, 2% CHAPS, sendo adicionado

50mM DTT e 0,5% IPG *buffer* apenas no momento do uso). Ao tampão foi misturado 1000µg de proteína total da amostra atingindo um volume de 250µl. Posteriormente a tira foi coberta com óleo mineral (*DryStrip Cover fluid* / GE Healthcare).

A reidratação foi feita à temperatura ambiente por 16 horas. As tiras foram transferidas para a placa de cerâmica (*Manifold* / GE Healthcare) e novamente cobertas com óleo mineral para o início da focalização. A IEF foi realizada com o programa do *Ettan IPGphor II* (QUADRO 2). As tiras foram imediatamente equilibradas ou congeladas a -80° C até a realização da segunda dimensão.

FASE	VOLTAGEM (V)	TEMPO (h:min)	kV/h
1	500	1:00	0,5
2	1000	1:00	0,8
3	8000	2:30	11,3
4	8000	0:30	4,1
Total	-	5:00	16,7

QUADRO 2 - PROGRAMA PARA O *ETTAN IPGphor II* (tira de 13cm; pH 4-7).

FONTE: GE HEALTHCARE HANDBOOK, 2-D Electrophoresis *Principles and Methods* (2005)

#### 5.4.2. Equilíbrio das Tiras de IPG

Previamente à SDS-PAGE, as tiras foram equilibradas em 5ml de tampão de equilíbrio (50mM Tris-HCl – pH 8,8; 6M Uréia; 30% (p/v) Glicerol e 2% (p/v) SDS e um traço de azul de bromofenol) e 50mM DTT (agente redutor) por 15 minutos sob leve agitação à temperatura ambiente. Em seguida as tiras foram equilibradas por mais 15 minutos, sob as mesmas condições, em tampão de equilíbrio acrescido de 4,5% de iodoacetamida (responsável por alquilar grupos tiol prevenindo reoxidações durante a eletroforese).

#### 5.4.3. Eletroforese em Gel de Poliacrilamida na Presença de SDS (SDS-PAGE)

Após o equilíbrio, as tiras foram submersas por alguns segundos em tampão de corrida (25mM Tris-base, 192mM glicina e 0,1% SDS) e em seguida posicionadas horizontalmente sobre o gel SDS-PAGE 10% com dimensões de 18cm x 16cm. Marcadores de massa molecular (LMW) conhecida entre 14 e 97 kDa (GE

Healthcare) foram aplicados em papel filtro que foi posicionado sobre o gel à esquerda da tira. As tiras e o LMW foram seladas com agarose 1% em tampão de corrida. Para que a frente de migração ficasse visível, foi acrescentado traços de azul de bromofenol à solução de agarose.

A segunda dimensão foi realizada em tampão de corrida sob temperatura de 11°C, mantida por refrigeração com circulador termostático *Multitemp II*, utilizando o sistema eletroforético *Hoefler SE 600 Ruby* (GE Healthcare). Para que a passagem das proteínas da tira para o gel SDS-PAGE ocorresse de maneira suave e homogênea, a eletroforese foi desenvolvida a 15 mA/gel nos 30 minutos iniciais. O restante da corrida foi realizada a 30 mA/gel, sendo interrompida quando a frente de migração estava a aproximadamente 1cm do limite inferior do gel. A duração total foi de aproximadamente 4 horas e 30 minutos. Todas as etapas e soluções utilizadas no 2D-PAGE foram adaptadas e desenvolvidas a partir do protocolo fornecido pelo fabricante do aparelho (*2-D Electrophoresis Principles and Methods*, 2005).

## 5.5. COLORAÇÃO, DIGITALIZAÇÃO E ANÁLISE COMPUTACIONAL

Para a fixação das proteínas, os géis permaneceram durante uma hora, sob leve agitação, em solução contendo 1,3% (v/v) de ácido orto-fosfórico 85% e 20% (v/v) de metanol. Foi utilizada a coloração com Coomassie coloidal adaptada de WESTERMEIER, 2006. Os géis ficaram submersos em 300ml de solução corante (1,5% de ácido orto-fosfórico 85%, 7,7% de sulfato de amônio e 0,01% de Coomassie G-250) sob leve agitação por um período de 16 a 20 horas. Para permanecerem apenas as marcações das bandas, os géis foram descorados em água destilada. O armazenamento foi realizado a 4 °C em solução contendo 1% de ácido acético.

Os géis de eletroforese bidimensionais foram digitalizados no *ImageScanner™ II* (GE Healthcare) e analisados com o programa *ImageMaster™ 2D Platinum v6.0* (GE Healthcare). Os parâmetros definidos para a detecção das bandas no programa foram os seguintes: *min area* = 5; *smooth* = 2; e *saliency* = 100.

## 5.6. ESPECTROMETRIA DE MASSAS

### 5.6.1. Preparo das Amostras

As bandas de interesse foram retiradas manualmente do gel utilizando ponteiras novas estéreis tipo P1000 e colocadas em tubos tipo eppendorf (0,5 ml) contendo 300µl de solução para descorar (50% de acetonitrila, 25mM de bicarbonato de amônio) por uma hora. Este procedimento foi repetido quando necessário até a completa descoloração. Em seguida, foram submetidas a duas etapas de desidratação com 100µl de acetonitrila 100% por 4 minutos cada. A acetonitrila foi descartada e os fragmentos de gel ficaram por uma hora secando à temperatura ambiente, sendo posteriormente armazenadas a 20° C negativos.

Para a digestão proteolítica *in gel*, as bandas selecionadas foram reidratadas em 15µl de solução de tripsina (*Sequencing Grade Modified Trypsin - Promega*) (15ng/µL em tampão contendo 40mM de bicarbonato de amônio e 10% de acetonitrila). A reidratação ocorreu durante 30 minutos em banho de gelo. Após esse período, a digestão foi realizada a 37° C por 16 a 20 horas. A solução contendo os peptídeos foi misturada na proporção 1:1 com solução de 50% de acetonitrila, 0,1% de ácido trifluoracético saturada com matriz HCCA (ácido  $\alpha$ -ciano-4-hidroxicinâmico). O volume final de 2µl foi aplicado em placa para MALDI (*scout MTP MALDI ion source 384 target - Bruker Daltonics*).

### 5.6.2. Purificação Peptídica

Algumas das bandas que apresentaram *score* limítrofe ou percentual de cobertura de seqüência abaixo do estabelecido para identificação válida foram submetidas à purificação de peptídeos. Para tal, foi utilizado uma ponteira com uma coluna de fase reversa "*PerfectPure C-18 Tip*" (Eppendorf). O procedimento foi realizado de acordo com instruções do fabricante. A ponteira foi pré-umidificada duas vezes com 10µl de solução de 50% de acetonitrila, 0,1% de ácido trifluoracético e equilibrada duas vezes com 10µl de solução de 0,1% de ácido trifluoracético. Em seguida, a amostra foi aspirada e expelida cerca de dez vezes,

sempre evitando a passagem de ar pela coluna de fase reversa. A amostra foi lavada cinco vezes novamente com solução de 0,1% de ácido trifluoracético para retirar contaminantes. Os peptídeos foram eluídos diretamente em tubo eppendorf (0,2ml) contendo 2µl de solução de 50% de acetonitrila, 0,1% de ácido trifluoracético saturada com matriz HCCA e aplicados na placa para o MALDI.

### 5.6.3. Obtenção e Análise dos Espectros de Massa

As massas dos peptídeos tripticos foram determinadas utilizando um espectrômetro de massa MALDI-ToF/ToF/MS/MS modelo *Autoflex II* (Bruker Daltonics). O programa de computador utilizado foi o *FlexControl 2.0* (Bruker Daltonics). Os espectros foram obtidos no modo refletor positivo, com uma voltagem de aceleração de 20 kV, intervalo de 150 ns entre o pulso de laser e aplicação da voltagem, e faixa de aquisição entre 800 – 3200 Da. A calibração externa foi feita utilizando padrões de peptídeos de massa conhecida (angiotensina, ACTH e somatostatina) e como calibrante interno foram utilizados picos de autólise da tripsina (842,50 Da e 2211,10 Da) quando presentes. A análise e tratamento dos espectros foi feita utilizando o programa *FlexAnalysis 2.0* (Bruker Daltonics).

### 5.6.4. Identificação das Proteínas

As proteínas foram identificadas por PMF utilizando o programa MASCOT (Matrix Science, UK - [http://www.matrixscience.com/search\\_form\\_select.html](http://www.matrixscience.com/search_form_select.html)) associado ao uso do banco de dados *Swiss-Prot*. Os parâmetros de busca utilizados foram: enzima tripsina, um sítio de clivagem perdido, carbamidometilação de cisteínas como modificação fixa e oxidação de metionina como modificação variável (ANEXO 1). A tolerância de erro utilizada foi de no máximo 200 ppm (partes por milhão).

Os critérios considerados para a identificação incluem:

- MASCOT score de probabilidade. O MASCOT score é o valor atribuído à identificação pelo programa de busca (ANEXO 2). O valor mínimo a partir do qual a identificação deixa de ser considerada aleatória é estipulado pelo

programa e varia de acordo com a taxonomia da amostra. Em buscas por proteínas humanas, utilizando o banco de dados *Swiss-Prot* e o programa MASCOT, esse valor é 56. Quanto maior o *score* maior a confiabilidade na identificação.

- Percentual da cobertura de seqüência. Este valor indica o quanto da seqüência de aminoácidos da proteína sugerida foi coberto pelos peptídeos resultantes da digestão tríptica da amostra (ANEXO 3). Valores acima de 20% foram considerados satisfatórios.
- Quantidade de massas peptídicas pareadas (mínimo 4). Este valor corresponde ao número de picos de massa obtidos por MS que foi pareado com as informações das massas teóricas (provenientes de clivagem *in silico*) para cada proteína (ANEXO 3). Os valores teóricos são fornecidos pelo programa de busca MASCOT
- Dados de pI e MM experimentais. Condizentes com os dados teóricos.

Os nomes e siglas utilizados para as proteínas foram os sugeridos no banco de dados *UniProtKB / Swiss-Prot* (<http://www.uniprot.org/>). Os respectivos genes são representados pelas siglas oficiais de acordo com o *National Center for Biotechnology Information* (NCBI). A tabela contendo o nome oficial em inglês das proteínas e o erro (em ppm) tolerado nas identificações está no APÊNDICE.

## 6. RESULTADOS

### 6.1. EXTRAÇÃO E QUANTIFICAÇÃO DE PROTEÍNAS

A técnica estabelecida para extração de proteínas de amostras de tecido apresentou resultados satisfatórios. Os resultados da quantificação protéica estão dispostos no QUADRO 3.

<b>CÓDIGO.</b>	<b>†MASSA (g)</b>	<b>CONCENTRAÇÃO (µg/µl)</b>
CP622 N	0,4	8,8
CP632 N	0,4	9,4
CP633 N	0,4	3,8
CP638 N	*0,35	13,4
CP645 N	0,4	15
CP655 N	0,4	8

QUADRO 3 - DOSAGEM DE PROTEÍNAS

FONTE: O autor (2010)

NOTA: †Massa de tecido utilizada para a extração de proteínas. \*Todo o material da amostra CP638 N foi utilizado.

O método de Bradford foi previamente testado para assegurar que o tampão de lise utilizado não interferiria na quantificação protéica (TABELA 1). Para isto, foram realizadas leituras da absorbância da solução de Bradford (200 µl de reagente de Bradford diluído 1:5 em água ultra pura), que é considerado como “branco” na quantificação, e da solução de Bradford acrescido de tampão de lise em três volumes: 1µl (0,5%), 5µl (2,5%) e 10µl (5%). A análise de variância entre as médias das leituras não apresentou diferença estatisticamente significativa ( $F=3,60$ ;  $p>0,05$ ) entre as leituras da absorbância do “branco” e as três concentrações de tampão de lise.

TABELA 1 - BRADFORD DO TAMPÃO DE LISE

AMOSTRA	ABSORBÂNCIA (595nm)	MÉDIA
Sol. de Bradford	0,398	0,387
	0,376	
Tampão de lise (1µl)	0,436	0,411
	0,387	
Tampão de lise (5µl)	0,455	0,429
	0,404	
Tampão de lise (10µl)	0,588	0,532
	0,477	

FONTE: O autor (2010)

NOTA: Volumes de 1µl, 5µl e 10µl de tampão foram adicionados à solução de Bradford para avaliar sua influência na quantificação. As diferenças entre as médias, avaliadas pela análise de variância, não foram estatisticamente significativas ( $F=3,60$ ;  $p>0,05$ ).

## 6.2. ANÁLISE DOS GÉIS BIDIMENSIONAIS

A confecção dos géis bidimensionais também foi implantada com êxito. Para cada amostra foram confeccionados géis em triplicata (FIGURA 4), totalizando 18 géis bidimensionais.

A QUADRO 4 apresenta o número de bandas detectadas através do programa *ImageMaster™ 2D Platinum v6.0* (GE Healthcare) em todos os géis bidimensionais confeccionados.

CP622 N		CP632 N		CP633 N	
Gel	nº de bandas	Gel	nº de bandas	Gel	nº de bandas
N1 (Ref.)	96	N1	182	N1 (Ref.)	139
N2	67	N2	135	N2	125
N3	86	N3 (Ref. MASTER)	193	N3	113
CP638 N		CP645 N		CP655 N	
Gel	nº de bandas	Gel	nº de bandas	Gel	nº de bandas
N1	97	N1	131	N1 (Ref.)	137
N2 (Ref.)	116	N2	143	N2	185
N3	95	N3 (Ref.)	146	N3	130

QUADRO 4 - QUANTIDADE DE BANDAS DETECTADAS NOS GÉIS BIDIMENSIONAIS

FONTE: O autor (2010)

NOTA: Os géis escolhidos como referência estão indicados entre parênteses.



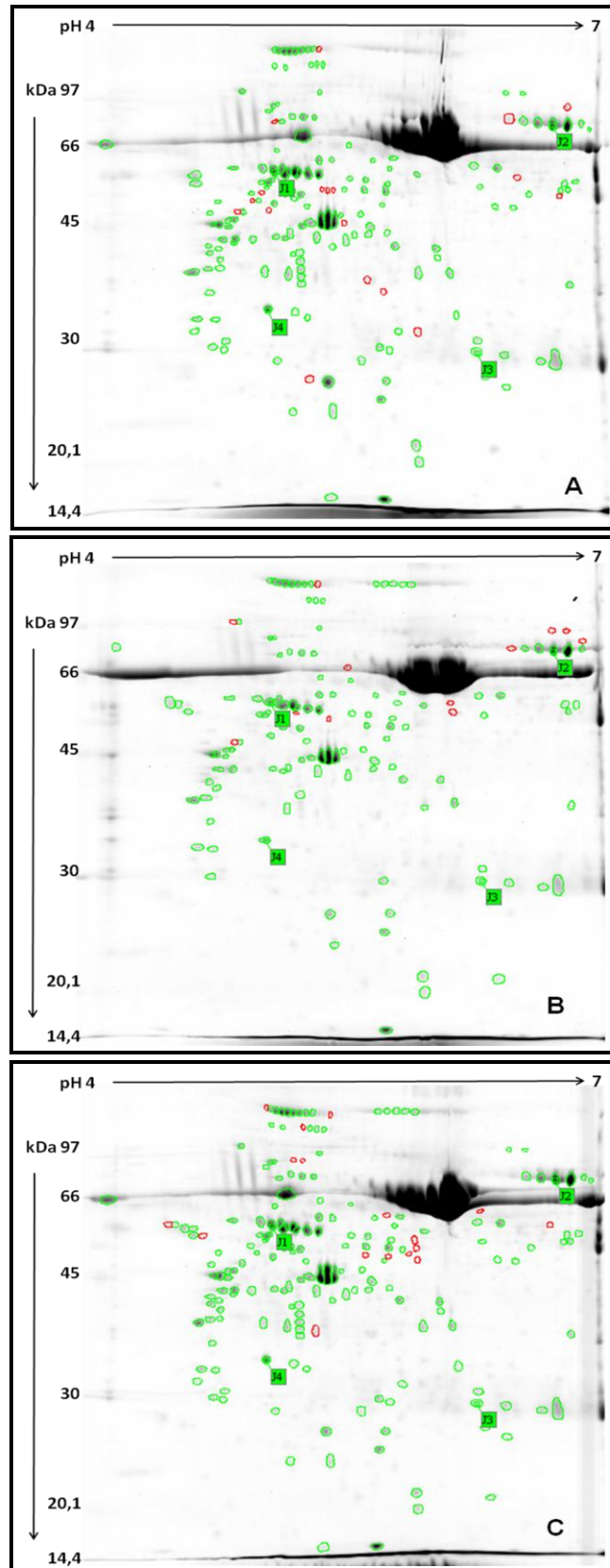


FIGURA 4 – TRIPLICATA DE GÉIS BIDIMENSIONAIS DA AMOSTRA CP632 N  
 FONTE: O autor (2010)

LEGENDA: Gel CP632 N1 (A), CP632 N2 (B) e, o escolhido como referência, CP632 N3 (C). Bandas destacadas em vermelho são exclusivas de cada gel. As bandas destacadas em verde foram detectadas em pelo menos dois géis da triplicata. J1, J2, J3 e J4 evidenciados pelos quadrados verdes são as bandas utilizadas no pareamento entre o gel de referência e os demais.

O gel de referência para cada amostra foi escolhido baseado na representatividade e resolução das bandas (FIGURA 5).

Os géis de referência de cada amostra foram submetidos à análise comparativa entre eles. O gel CP632 N3 foi escolhido como *Master*, servindo como base para a comparação e determinação de pI e MM das bandas no programa *ImageMaster™ 2D Platinum v6.0* (GE Healthcare). Desta análise, observou-se que 118 bandas detectadas estavam presentes no gel Master e em pelo menos dois dos outros cinco géis de referência, sendo que 43 (36,4%) estavam presentes em todas as amostras (QUADRO 5).

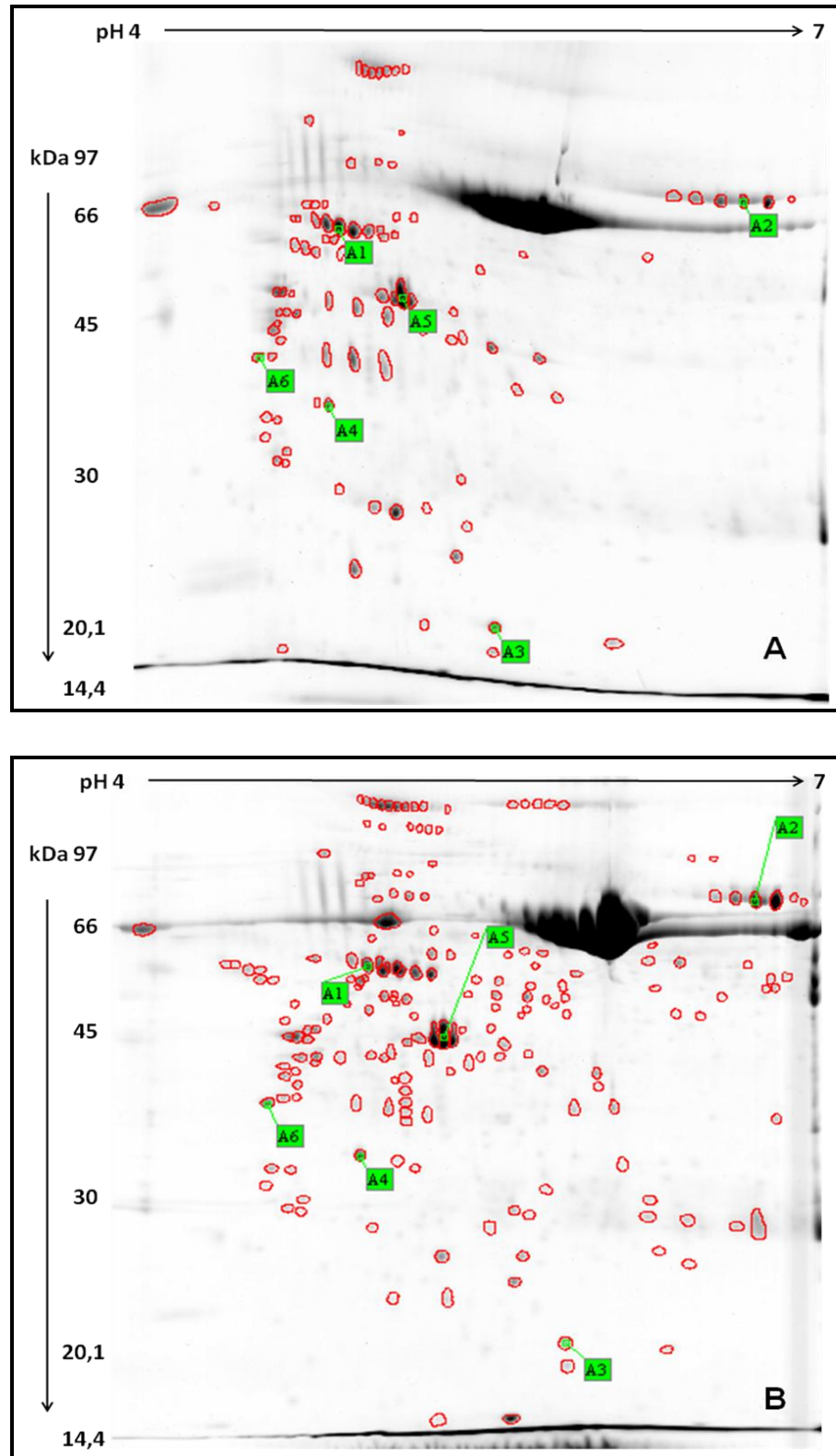


FIGURA 5 - GÉIS DE REFERÊNCIA (continua)

FONTE: O autor (2010)

LEGENDA: Gel CP622 N1 (A) e CP622 N3 (MASTER) (B). Bandas detectadas pelo programa *ImageMaster* destacadas em vermelho. A1, A2, A3, A4, A5 e A6, destacadas pelo quadrado verde, representam as bandas utilizadas para pareamento entre o gel *Master* (B) e os demais.

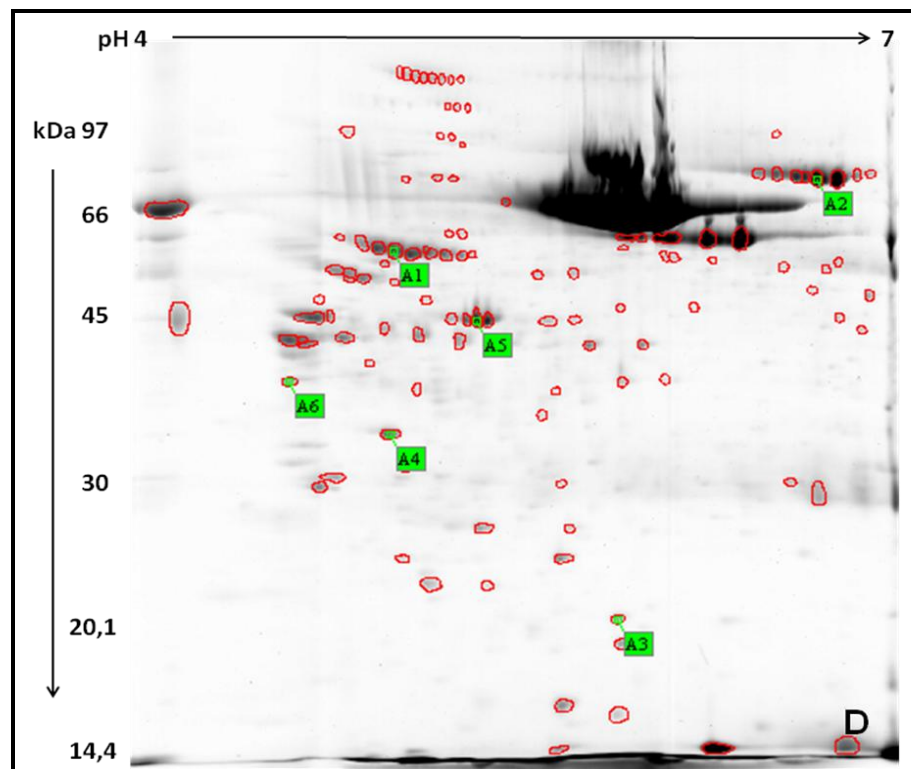
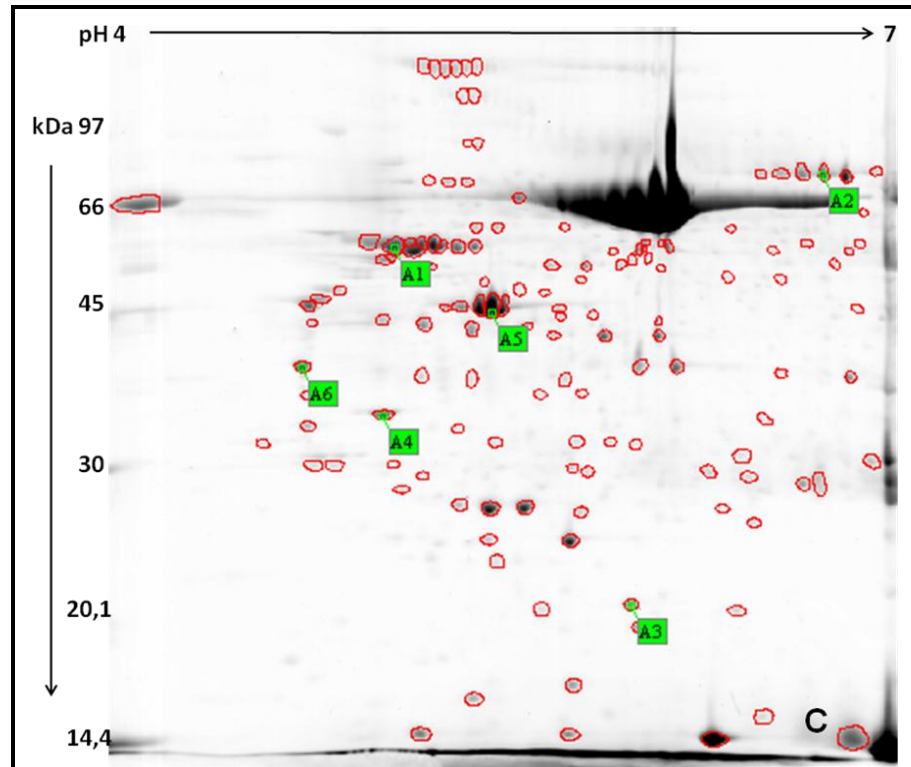


FIGURA 5 - GÉIS DE REFERÊNCIA (continuação)

FONTE: O autor (2010)

LEGENDA: Gel CP633 N1 (C) e CP638 N2 (D). Bandas detectadas pelo programa *ImageMaster* destacadas em vermelho. A1, A2, A3, A4, A5 e A6, destacadas pelo quadrado verde, representam as bandas utilizadas para pareamento entre o gel *Master* (B) e os demais.

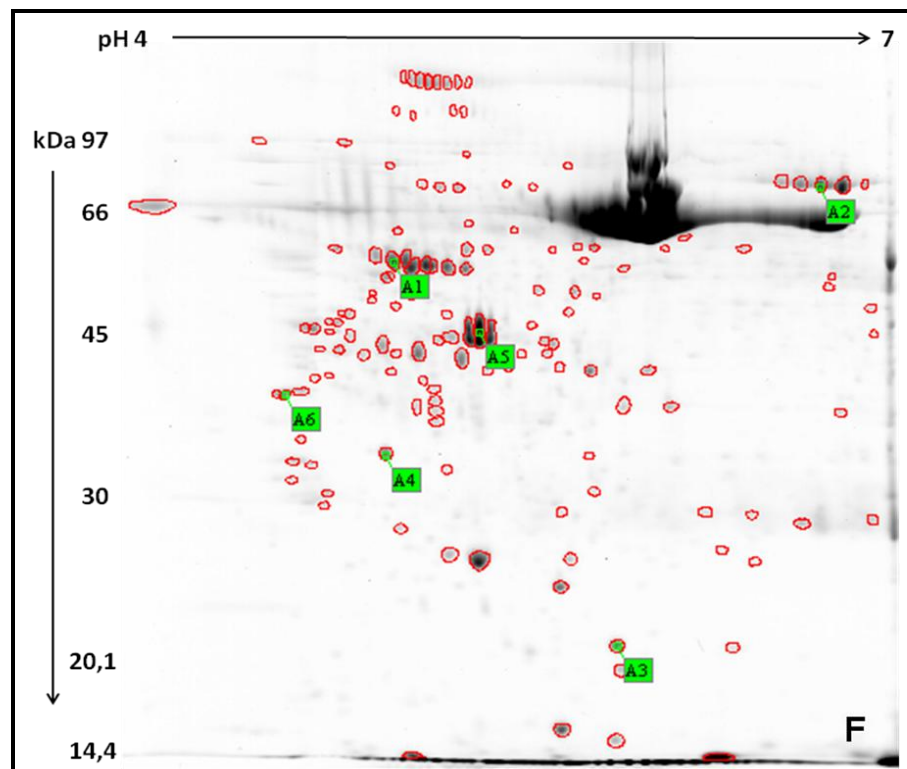
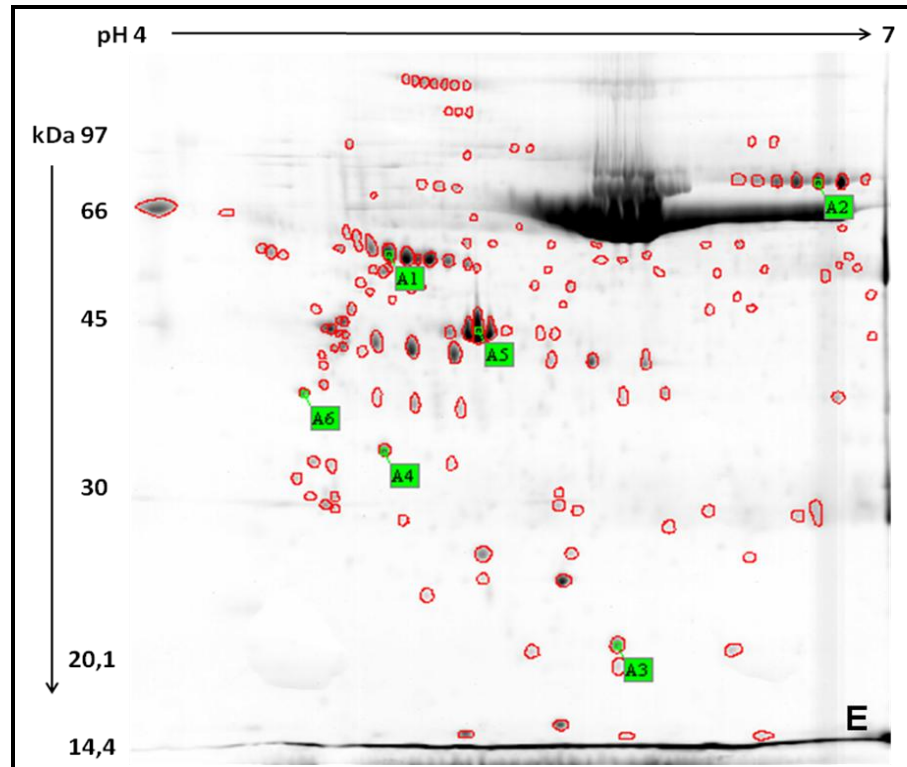


FIGURA 5 - GÉIS DE REFERÊNCIA (conclusão)

FONTE: O autor (2010)

LEGENDA: Gel CP645 N1 (E) e CP655 N2 (F). Bandas detectadas pelo programa *ImageMaster* destacadas em vermelho. A1, A2, A3, A4, A5 e A6, destacadas pelo quadrado verde, representam as bandas utilizadas para pareamento entre o gel *Master* (B) e os demais.

DETECTADAS EM TODAS AS AMOSTRAS	DETECTADAS EM 5 AMOSTRAS	DETECTADAS EM 4 AMOSTRAS	DETECTADAS EM 3 AMOSTRAS
140	139	134	167
142	156	185	203
144	157	196	239
145	164	264	250
146	262	276	261
170	268	292	269
189	317	314	275
192	355	333	287
195	364	339	293
197	375	346	306
198	378	351	315
201	441	363	330
230	453	405	337
260	460	418	349
267	470	422	350
271	492	429	365
272	514	440	371
283	-	442	374
285	-	447	376
286	-	463	377
298	-	468	385
319	-	469	396
320	-	480	400
359	-	486	408
361	-	619	414
362	-	-	450
367	-	-	457
369	-	-	462
379	-	-	465
386	-	-	476
392	-	-	485
394	-	-	513
411	-	-	617
415	-	-	-
417	-	-	-
433	-	-	-
459	-	-	-
478	-	-	-
479	-	-	-
481	-	-	-
491	-	-	-
497	-	-	-
618	-	-	-
<b>Total: 43</b>	<b>Total: 17</b>	<b>Total: 25</b>	<b>Total: 33</b>

QUADRO 5 - RELAÇÃO DAS BANDAS (n=118) DETECTADAS EM PELO MENOS 3 AMOSTRAS

FONTE: O autor (2010)

NOTA: O número das bandas é referente ao ID atribuído pelo programa *ImageMaster™ 2D Platinum v6.0* para o gel CP632 N3. As bandas submetidas à identificação estão destacadas em vermelho.

### 6.3. PURIFICAÇÃO PEPTÍDICA E ESPECTROMETRIA DE MASSAS

A purificação peptídica foi realizada em 21 bandas e gerou uma melhora considerável na obtenção dos espectros de massas e na identificação em dez ocasiões (47,6%) (TABELA 2). As bandas que não obtiveram melhora foram: 140, 142, 156, 157, 170, 262, 375, 411, 491, 497 e 618. O espectro de massas de algumas bandas e comparação entre o espectro obtido antes e depois da purificação peptídica está demonstrado na FIGURA 6.

### 6.4. IDENTIFICAÇÃO PROTÉICA

Das 118 bandas detectadas, 70 foram submetidas aos procedimentos para identificação, sendo que para todas obteve-se êxito. (QUADRO 5). A FIGURA 7 apresenta o gel de *Master* com as bandas identificadas. As bandas identificadas corresponderam a 44 proteínas diferentes, conforme demonstrado na TABELA 2. A FIGURA 8 apresenta o gel *Master* com as proteínas identificadas associadas às suas referentes bandas.

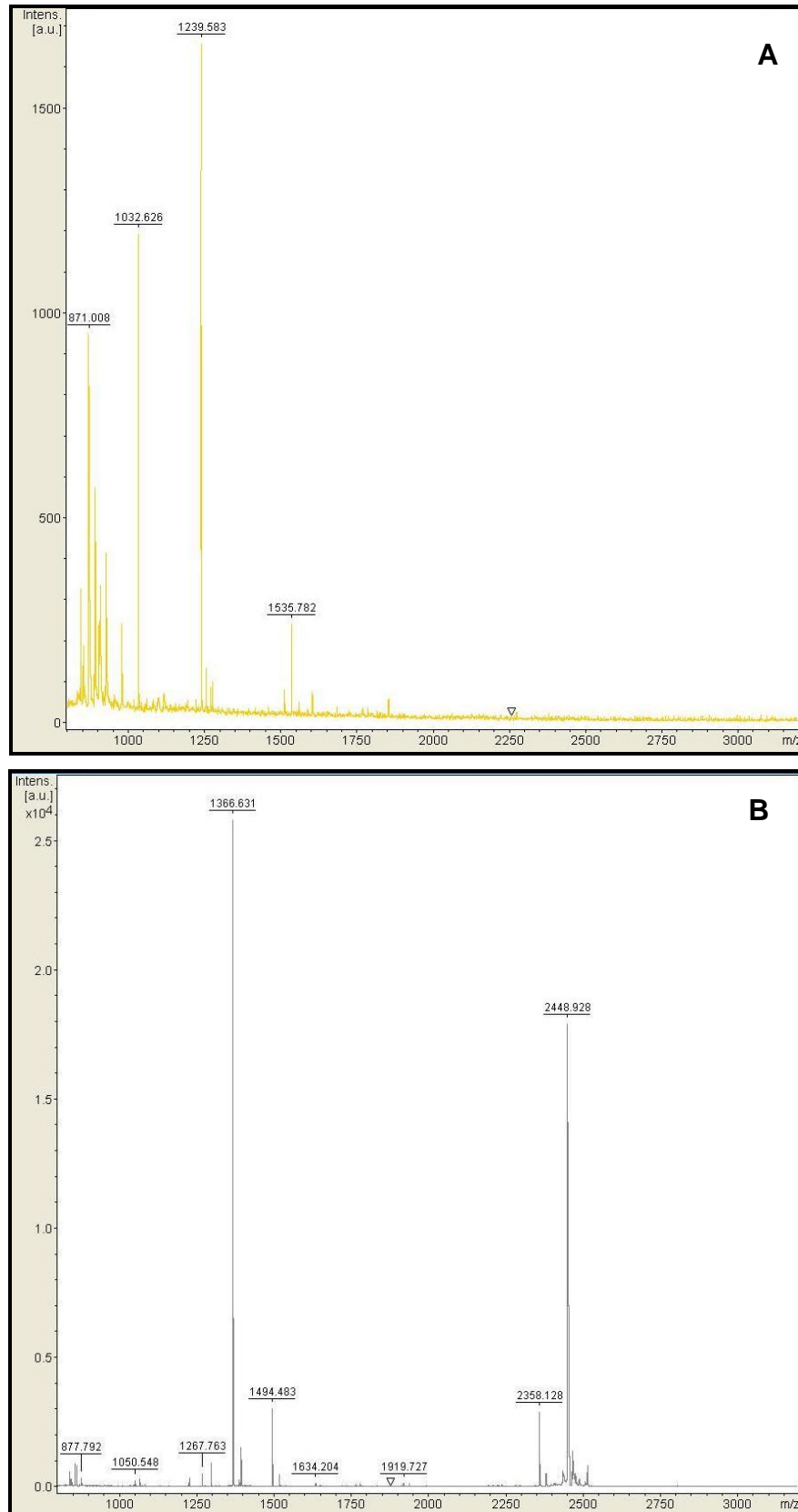


FIGURA 6 – ESPECTROS DE MASSAS (continua)

FONTE: O autor (2010)

LEGENDA: Espectro de massas da banda 319 que não resultou em identificação (A), com ruído alto muito sinal de matriz e poucos picos de massa; e o espectro de massas utilizado na obtenção da identificação para a mesma banda (B).



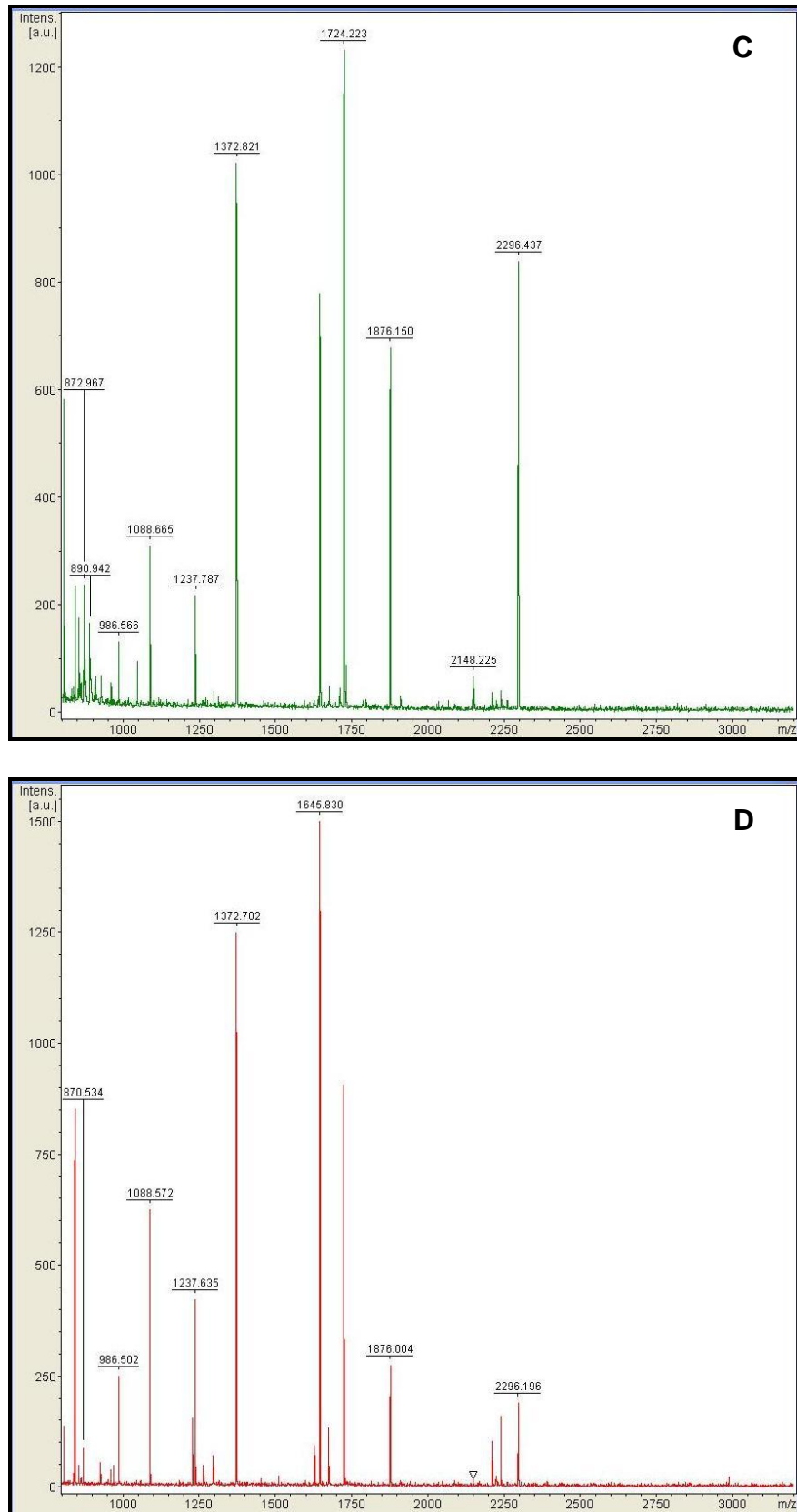


FIGURA 6 – ESPECTROS DE MASSAS (continuação)

FONTE: O autor (2010)

LEGENDA: Banda 189 sem (C) e com purificação peptídica (D). A identificação só foi possível com a purificação.

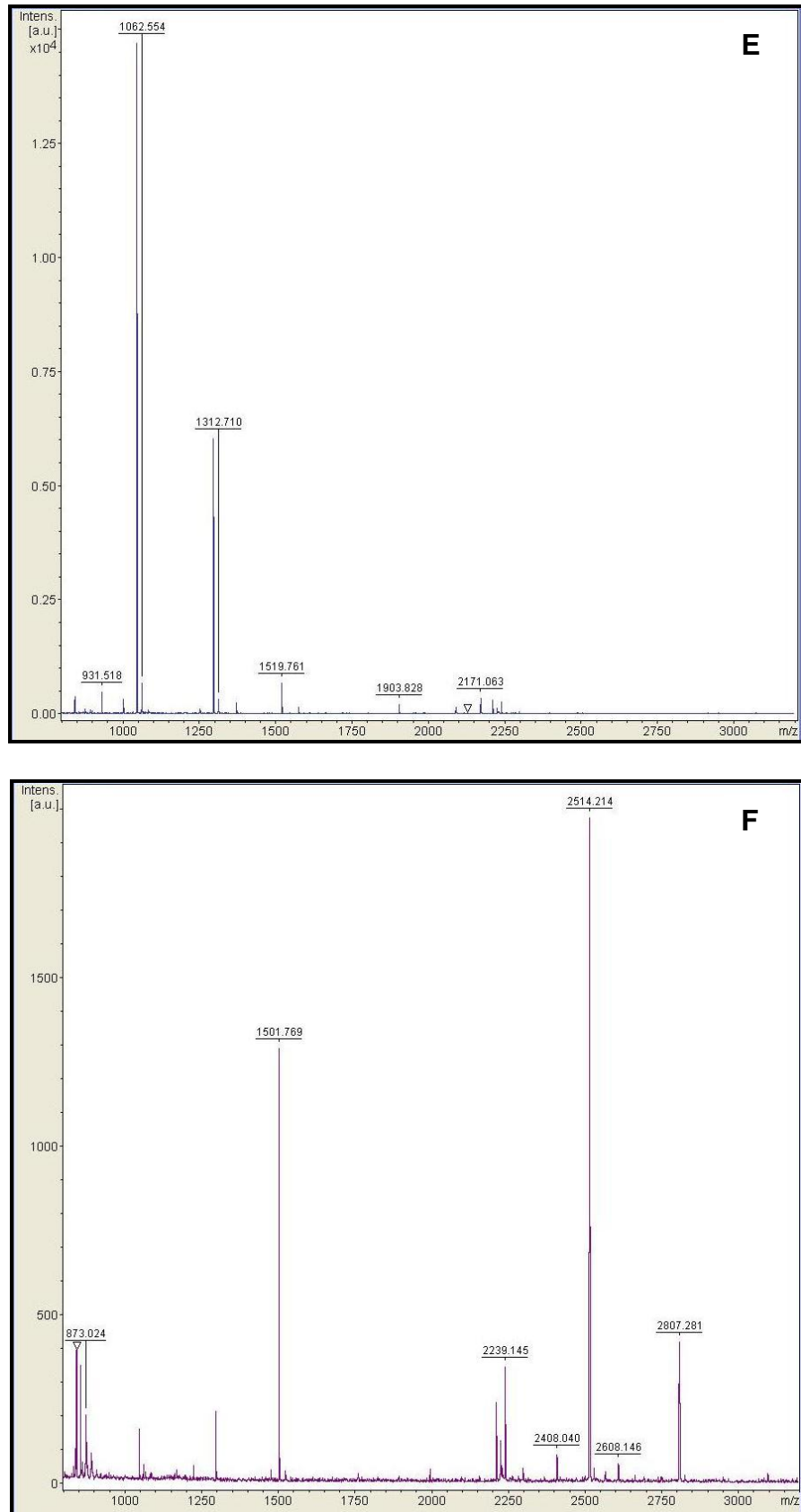


FIGURA 6 – ESPECTROS DE MASSAS (conclusão)

FONTE: O autor (2010)

LEGENDA: Bandas 157(E) e 497(F). Ambas foram submetidas à purificação peptídica, mas não houve melhora de *score* ou de cobertura de seqüência.

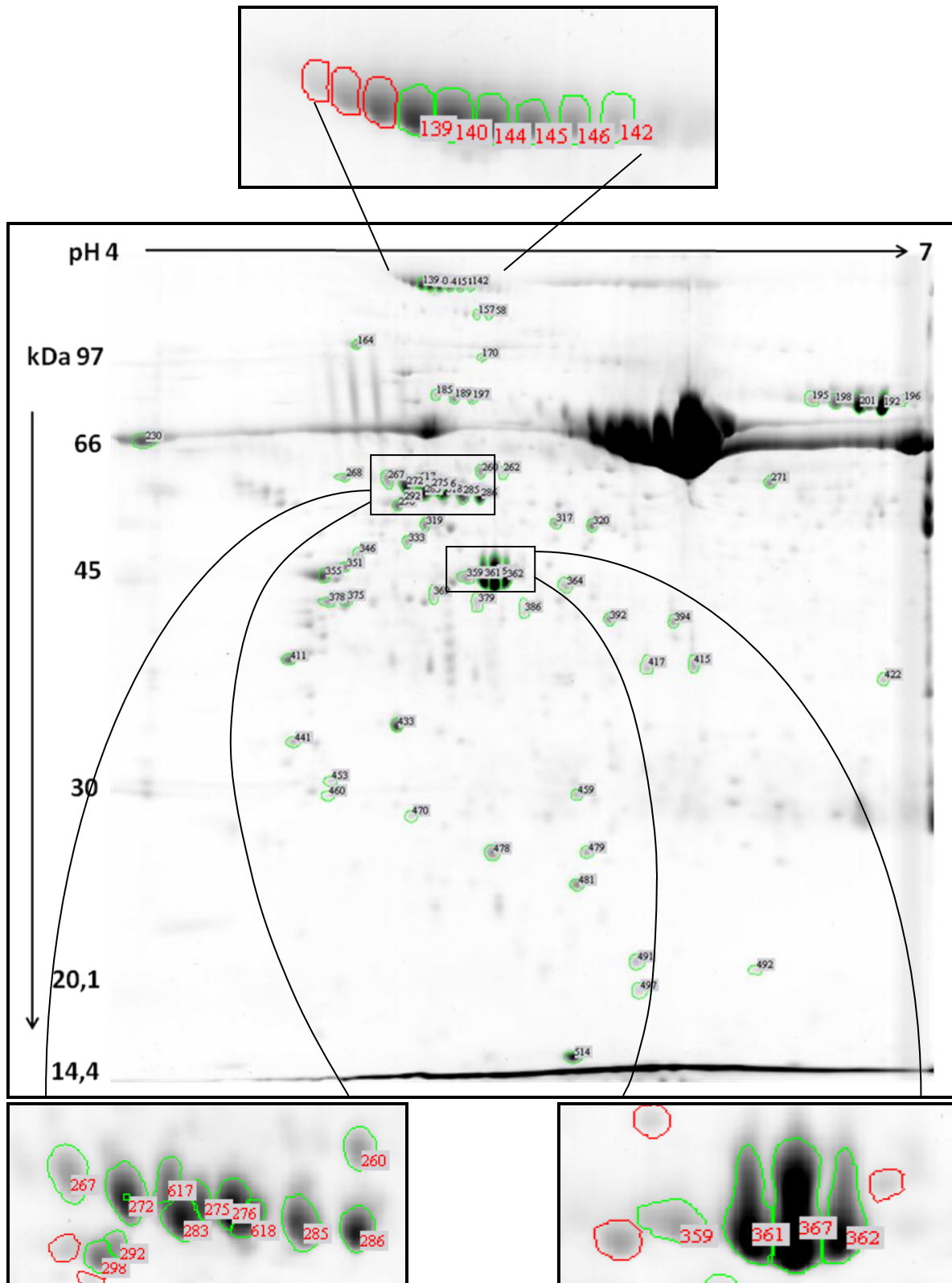


FIGURA 7 – BANDAS IDENTIFICADAS INDICADAS NO GEL MASTER

FONTE: O autor (2010)

LEGENDA: As bandas estão demonstradas no gel *Master* (CP632 N3) destacadas em verde e com a respectiva numeração atribuída pelo programa *ImageMaster*.

TABELA 2: IDENTIFICAÇÕES OBTIDAS POR MS E DADOS REFERENTES AOS CRITÉRIOS CONSIDERADOS (continua)

Nome da proteína (SIGLA)	ID	Código UniprotKB	Mascot score / 2º score	MM (kDa) / pl teóricos	MM (kDa) / pl observados	ΔMM / Δpl	Cobertura de seqüência (%)	Massas pareadas
Cadeia alfa-1 do colágeno VI (CO6A1)	139	P12109	71 / 25	109,6 / 5,26	125,0 / 5,12	16,3% / 0,16	12	10 / 22 (45%)
	140 *	P12109	106 / 27	109,6 / 5,26	125,0 / 5,16	15,7% / 0,12	17	14 / 27 (52%)
	142 *	P12109	99 / 30	109,6 / 5,26	125,0 / 5,27	15% / 0,03	16	10 / 20 (50%)
	144 *	P12109	67 / 26	109,6 / 5,26	125,0 / 5,19	15% / 0,09	11	10 / 22 (45%)
	145 *	P12109	150 / 42	109,6 / 5,26	125,0 / 5,21	14,4% / 0,05	20	14 / 18 (78%)
	146 *	P12109	82 / 34	109,6 / 5,26	125,0 / 5,24	15% / 0,01	12	7 / 13 (54%)
Glicoproteína alfa-1B (A1BG)	185 *	P04217	61 / 26	54,8 / 5,58	82,0 / 5,17	33,1% / 0,41	17	5 / 16 (31%)
	189 *	P04217	143 / 33	54,8 / 5,58	81,0 / 5,23	32,3% / 0,35	27	11 / 17 (65%)
	197 *	P04217	106 / 27	54,8 / 5,58	81,0 / 5,30	32,3% / 0,28	24	8 / 13 (61%)
Sorotransferrina (TRFE)	192 *	P02787	70 / 23	79,3 / 6,81	80,0 / 6,80	1,7% / 0,01	13	8 / 20 (40%)
	195 *	P02787	133 / 29	79,3 / 6,81	81,0 / 6,50	0,9% / 0,26	18	12 / 19 (63%)
	198 *	P02787	225 / 28	79,3 / 6,81	81,0 / 6,64	2,1% / 0,17	29	21 / 35 (60%)
	201 *	P02787	142 / 28	79,3 / 6,81	80,0 / 6,73	0,9% / 0,08	22	15 / 31 (48%)
	196	P02787	94 / 22	79,3 / 6,81	80,0 / 6,89	0,9% / 0,08	21	12 / 31 (39%)
Albumina sérica (ALBU)	230 *	P02768	165 / 40	71,3 / 5,92	70,0 / 4,08	1,85% / 1,84	25	13 / 15 (87%)
	271 *	P02768	180 / 27	71,3 / 5,92	61,0 / 6,39	14,4% / 0,47	27	15 / 20 (75%)
	A <sup>■</sup>	P02768	220 / 25	71,3 / 5,92	67,0 / 6,10	6% / 0,18	37	24 / 51 (47%)
Antitrombina-III (ANT3)	260 *	P01008	114 / 25	53,0 / 6,32	62,0 / 5,33	13,1% / 1	29	13 / 22 (59%)
Alfa-1-Antitripsina (A1AT)	267 *	P01009	106 / 31	46,9 / 5,37	61,0 / 4,98	23,1% / 0,39	22	10 / 25 (40%)
	272 *	P01009	123 / 19	46,9 / 5,37	59,0 / 5,05	20,5% / 0,32	24	13 / 35 (37%)
	283 *	P01009	162 / 31	46,9 / 5,37	58,0 / 5,12	19,1% / 0,25	37	15 / 33 (45%)
	617 *	P01009	62 / 24	46,9 / 5,37	61,0 / 5,09	23,1% / 0,28	21	8 / 31 (26%)
	618 *	P01009	67 / -	46,9 / 5,37	58,0 / 5,19	19,1% / 0,18	27	9 / 36 (25%)
Proteína ligante de vitamina D (VTDB)	286 *	P02774	84 / 18	54,5 / 5,40	57,0 / 5,32	4,4% / 0,08	25	8 / 23 (35%)
Proteína ligante de vitamina D (VTDB)	285 * (mistura) score 162	P02774	97 / 19 - quando não em mistura	54,5 / 5,40	58,0 / 5,26	6% / 0,14	30	10 / 29 (34%)
Alfa-1-Antitripsina (A1AT)		P01009	85 / 23 - quando não em mistura	46,9 / 5,37	58,0 / 5,26	19,1% / 0,11	17	9 / 29 (31%)
Cadeia beta-2C da tubulina (TBB2C)	292	P68371	99 / 70	50,2 / 4,79	57,0 / 5,04	11,9% / 0,25	31	12 / 29 (41%)
Cadeia gama do fibrinogênio (FIBG)	320 *	P02679	85 / 27	52,1 / 5,37	53,0 / 5,74	1,7% / 0,37	25	8 / 22 (36%)
	317	P02679	82 / 25	52,1 / 5,37	53,0 / 5,60	1,7% / 0,23	28	8 / 25 (32%)

TABELA 2: IDENTIFICAÇÕES OBTIDAS POR MS E DADOS REFERENTES AOS CRITÉRIOS CONSIDERADOS (continuação)

Nome da proteína (SIGLA)	ID	Código UniprotKB	Mascot score / 2º score	MM (kDa) / pl teóricos	MM (kDa) / pl observados	ΔMM / Δpl	Cobertura de seqüência (%)	Massas pareadas
Cadeia beta da tubulina (TBB5)	298 *	P07437	216 / 162	50,1 / 4,78	56,0 / 5,01	11,9% / 0,25	52	19 / 29 (66%)
Subunidade beta daATP sintetase mitocondrial (ATPB)	319 *	P06576	194 / 40	56,5 / 5,26	53,0 / 5,13	10,5% / 0,23	37	16 / 21 (76%)
Actina citoplasmática 1 (ACTB); Actina citoplasmática 2 (ACTG)	359 *	P60709	96 / 41	42,0 / 5,29	45,0 / 5,27	6,7% / 0,02	23	6 / 7 (86%)
		P63261	96 / 41	42,1 / 5,31	45,0 / 5,27	6,4% / 0,04	23	6 / 7 (86%)
	362 *	P60709	79 / -	42,0 / 5,29	45,0 / 5,42	6,7% / 0,13	32	10 / 37 (27%)
		P63261	79 / -	42,1 / 5,31	45,0 / 5,42	6,4% / 0,11	32	10 / 37 (27%)
	367 *	P60709	88 / 33	42,0 / 5,29	44,0 / 5,38	4,5% / 0,09	28	9 / 25 (36%)
		P63261	88 / 33	42,1 / 5,31	44,0 / 5,38	4,3% / 0,07	28	9 / 25 (36%)
Actina do músculo liso da aorta (ACTA); Actina do músculo cardíaco alfa (ACTC); Actina do músculo esquelético alfa (ACTS)	361*	P62736	75 / 59	42,4 / 5,23	45,0 / 5,34	5,8% / 0,11	20	7 / 29 (24%)
		P68032	75 / 59	42,3 / 5,23	45,0 / 5,34	6% / 0,11	20	7 / 29 (24%)
		P68133	75 / 59	42,4 / 5,23	45,0 / 5,34	5,8% / 0,11	20	7 / 29 (24%)
Haptoglobina (HPT)	369 *	P00738	130 / 57	45,9 / 6,13	43,0 / 5,15	6,3% / 1,02	25	10 / 16 (62%)
	379 *	P00738	138 / 31	45,9 / 6,13	43,0 / 5,25	6,3% / 0,88	24	9 / 14 (64%)
	386 *	P00738	87 / 27	45,9 / 6,13	42,0 / 5,46	8,5% / 0,64	19	8 / 13 (61%)
Glicerol-3-fosfato desidrogenase [NAD+] (GPDA)	415 *	P21695	142 / 31	38,2 / 5,81	38,0 / 6,1	0,5% / 0,29	43	12 / 24 (50%)
Cadeia beta da L-lactato desidrogenase (LDHB)	417 * (mistura) score 184	P07195	111 / 40 - quando não em mistura	36,9 / 5,71	38,0 / 5,94	2,9% / 0,23	25	8 / 14 (57%)
Glicerol-3-fosfato desidrogenase [NAD+] (GPDA)		P21695	73 / 31 - quando não em mistura	38,2 / 5,81	38,0 / 5,94	0,5% / 0,13	18	6 / 14 (43%)
Anexina A5 (ANXA5)	433 *	P08758	193 / 52	36,0 / 4,94	34,0 / 5,01	5,5% / 0,07	52	16 / 30 (53%)
Anexina A1 (ANXA1)	422	P04083	104 / 27	38,9 / 6,57	37,0 / 6,81	4,9% / 0,24	32	9 / 22 (41%)
Componente amilóide-P sérico (SAMP)	459 *	P02743	68 / 32	25,5 / 6,1	30,0 / 5,69	15% / 0,41	25	5 / 23 (22%)
Apolipoproteína A-I (APOA1)	478 *	P02647	229 / 34	30,7 / 5,56	26,0 / 5,37	15,3% / 0,19	62	18 / 30 (60%)
Glutathione S-transferase P (GSTP1)	479 *	P09211	68 / 27	23,6 / 5,43	26,0 / 5,71	9,2% / 0,28	41	6 / 33 (18%)
Peroxiredoxina-2 (PRDX2)	481 *	P32119	76 / 27	22,0 / 5,66	25,0 / 5,68	12% / 0,02	22	5 / 20 (25%)

TABELA 2: IDENTIFICAÇÕES OBTIDAS POR MS E DADOS REFERENTES AOS CRITÉRIOS CONSIDERADOS (continuação)

Nome da proteína (SIGLA)	ID	Código UniprotKB	Mascot score / 2º score	MM (KDa) / pl teóricos	MM (KDa) / pl observados	$\Delta$ MM / $\Delta$ pl	Cobertura de seqüência (%)	Massas pareadas
Vimentina (VIME)	355	P08670	90 / 24	53,7 / 5,06	45,0 / 4,75	16,2% / 0,31	29	11 / 29 (38%)
	378	P08670	261 / 30	53,7 / 5,06	43,0 / 4,76	19,9% / 0,3	40	21 / 24 (87%)
	276	P08670	66 / 44	53,7 / 5,06	60,0 / 5,17	10,5% / 0,11	25	12 / 52 (23%)
	346	P08670	126 / 32	53,7 / 5,06	48,0 / 4,87	10,6% / 0,19	37	14 / 27 (52%)
	351	P08670	132 / 33	53,7 / 5,06	46,0 / 4,82	14,3% / 0,24	33	14 / 24 (58%)
	275	P08670	115 / 28	53,7 / 5,06	60,0 / 5,13	10,5% / 0,07	35	14 / 31 (45%)
Creatina quinase tipo B (KCRB)	364 (mistura) score 210	P12277	111 / 24 - quando não em mistura	42,9 / 5,34	44,0 / 5,64	2,5% / 0,3	32	9 / 23 (39%)
Albumina sérica (ALBU)		P02768	110 / 34 - quando não em mistura	71,3 / 5,92	44,0 / 5,64	38,2% / 0,28	22	11 / 23 (48%)
Cadeia alfa-4 da tropomiosina (TPM4)	441	P67936	109 / 49	28,6 / 4,67	33,0 / 4,63	13,3% / 0,04	20	8 / 12 (67%)
Proteína 14-3-3 gama (1433G)	453	P61981	60 / 26	28,4 / 4,8	31,0 / 4,77	8,4% / 0,03	25	6 / 22 (27%)
Proteína 14-3-3 beta/alfa (1433B)	460	P31946	82 / 29	28,2 / 4,76	30,0 / 4,75	6% / 0,01	34	7 / 16 (44%)
Inibidor de dissociação Rho GDP (GDIR1)	470	P52565	57 / 25	23,2 / 5,02	28,0 / 5,07	17,1% / 0,05	29	5 / 22 (23%)
Proteína <i>heat shock</i> beta-6 (HSPB6)	492	O14558	62 / 31	17,2 / 5,95	21,0 / 6,34	18,1% / 0,39	29	4 / 10 (40%)
Transtiretina (TTHY)	514	P02766	69 / 26	16,0 / 5,52	15,0 / 5,66	6,25% / 0,14	48	4 / 19 (21%)
Endoplasmina (ENLP)	164	P14625	72 / 24	92,7 / 4,76	97,0 / 4,87	9,1% / 0,1	15	10 / 22 (45%)
Tripartite motif-containing protein 13 (TRI13)	262	O60858	58 / 32	47,8 / 5,75	63,0 / 5,4	21,6% / 0,34	15	6 / 23 (26%)
Proteína dissulfeto-isomerase (PDIA1)	268	P07237	103 / 38	57,5 / 62,0	4,76 / 4,85	4,2% / 0,06	18	8 / 19 (42%)
ATPase transicional do retículo endoplasmático (TERA)	170 *	P55072	92 / 33	89,9 / 5,14	95,0 / 5,33	6,3% / 0,19	13	10 / 13 (77%)
Cadeia beta do fibrinogênio (FIBB)	392 *	P02675	204 / 39	56,6 / 8,54	41,0 / 5,8	27,6% / 2,74	38	21 / 24 (87%)
	394 *	P02675	195 / 26	56,6 / 8,54	41,0 / 6,13	27,6% / 2,5	39	20 / 26 (77%)
Cadeia beta da tropomiosina (TPM2)	411 *	P07951	86 / 43	32,9 / 4,66	39,0 / 4,61	13,4% / 0,04	19	7 / 13 (54%)
Fosfoenolpiruvato carboxiquinase [GTP] citosólica (PCKGC)	491 *	P35558	56 / 28	69,9 / 5,8	22,0 / 5,86	69,9% / 0,11	13	6 / 30 (20%)

TABELA 2: IDENTIFICAÇÕES OBTIDAS POR MS E DADOS REFERENTES AOS CRITÉRIOS CONSIDERADOS (conclusão)

Nome da proteína (SIGLA)	ID	Código UniprotKB	Mascot score / 2º score	MM (KDa) / pI teóricos	MM (KDa) / pI observados	$\Delta$ MM / $\Delta$ pI	Cobertura de seqüência (%)	Massas pareadas
Ceruloplasmina (CERU)	156	P00450	58 / 22	123,0 / 5,44	119,0 / 5,27	6,5% / 0,17	13	9 / 27 (33%)
	157	P00450	53 / 25	123,0 / 5,44	119,0 / 5,27	6,5% / 0,13	10	7 / 19 (37%)
Ubiquitina E3 proteína-ligase RAD18 (RAD18)	375	Q9NS91	57 / 25	57,0 / 7,17	43,0 / 4,86	24,5% / 2,35	19	6 / 23 (26%)
Proteína tipo-actina 3B (ARP3B)	497 *	Q9P1U1	62 / 25	48,1 / 5,61	19,0 / 5,88	58,4% / 0,3	14	5 / 21 (24%)
Queratina do citoesqueleto tipo I – 17 (K1C17)	333	Q04695	60 / 20	48,4 / 4,97	48,0 / 5,08	3,2% / 0,09	20	7 / 28 (25%)

FONTE: O autor (2010)

NOTA: As bandas cujas identificações foram obtidas apenas após purificação protéica estão destacadas com fundo cinza. A coluna das Massas pareadas corresponde ao número de picos de massa obtidos com a espectrometria de massas e que foram pareados com as informações das massas teóricas (provenientes de clivagem *in silico*) para cada proteína. Estes valores são fornecidos pelo programa de busca MASCOT (*Matrix Science, UK*)

\* Bandas detectadas em todas as amostras. ■ Marcação característica em todas as amostras, retirada para identificação.

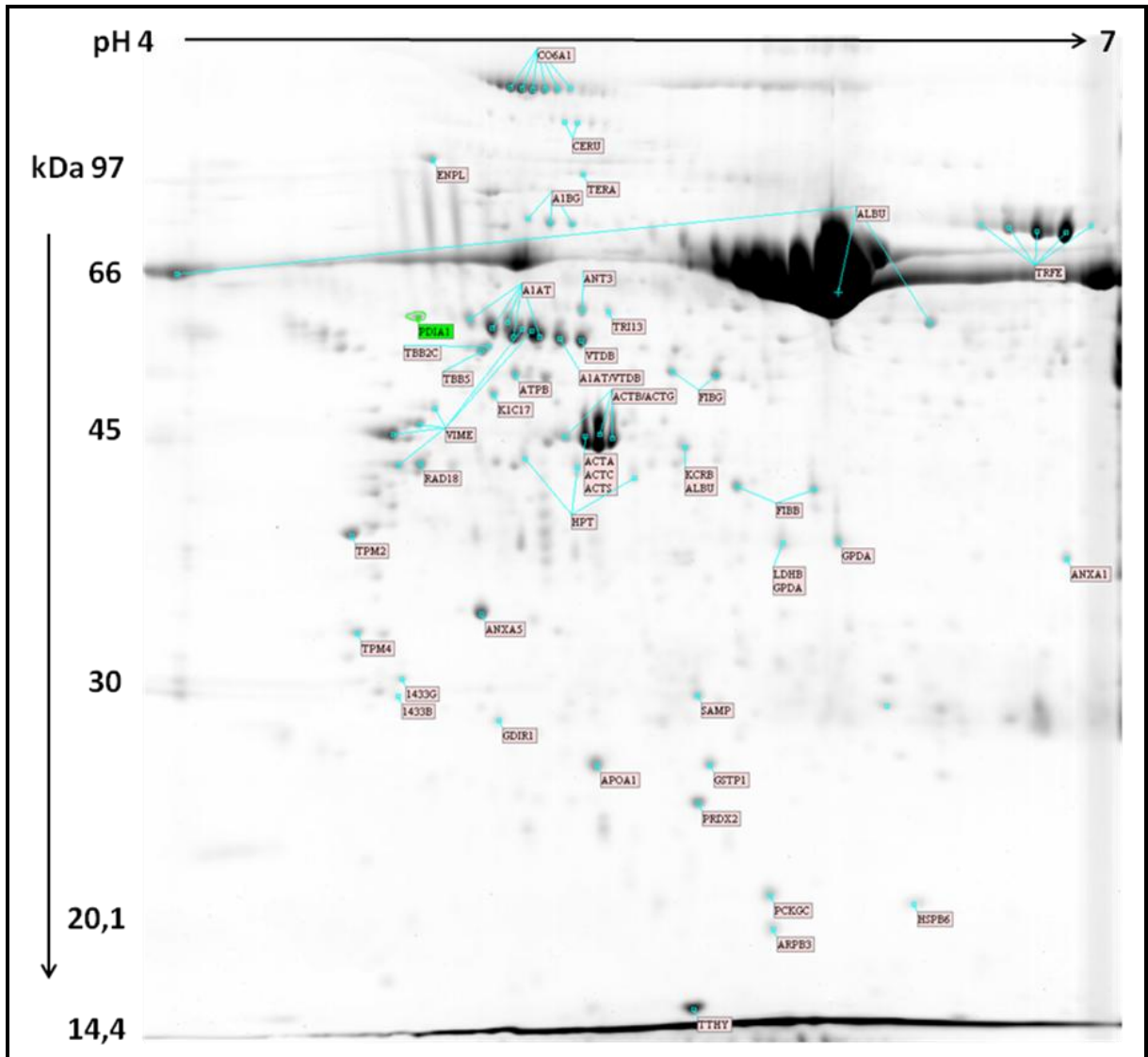


FIGURA 8 – PROTEÍNAS IDENTIFICADAS E SUA LOCALIZAÇÃO NO GEL *MASTER*

FONTE: O autor (2010)

LEGENDA: Bandas detectadas pelo programa *ImageMaster* estão destacadas em vermelho. Todas as identificações estão demonstradas através das siglas das proteínas ligadas às suas bandas correspondentes.



Através das informações obtidas na literatura e nos bancos de dados do NCBI (<http://www.ncbi.nlm.nih.gov/>) e *UniProtKB / Swiss-Prot* (<http://www.expasy.ch/>), as 44 proteínas identificadas foram agrupadas de acordo com sua função biológica em nove classes: Citoesqueleto e proteínas associadas; Proteínas com funções de ligação; Enzimas metabólicas; Chaperonas moleculares / proteínas *heat shock*; Detoxificação e proteínas redox; Proteínas membrana-associadas com múltiplas atividades; Crescimento celular e reguladores de proliferação; Degradação protéica; e Proteínas com outras funções (GRÁFICO 1).

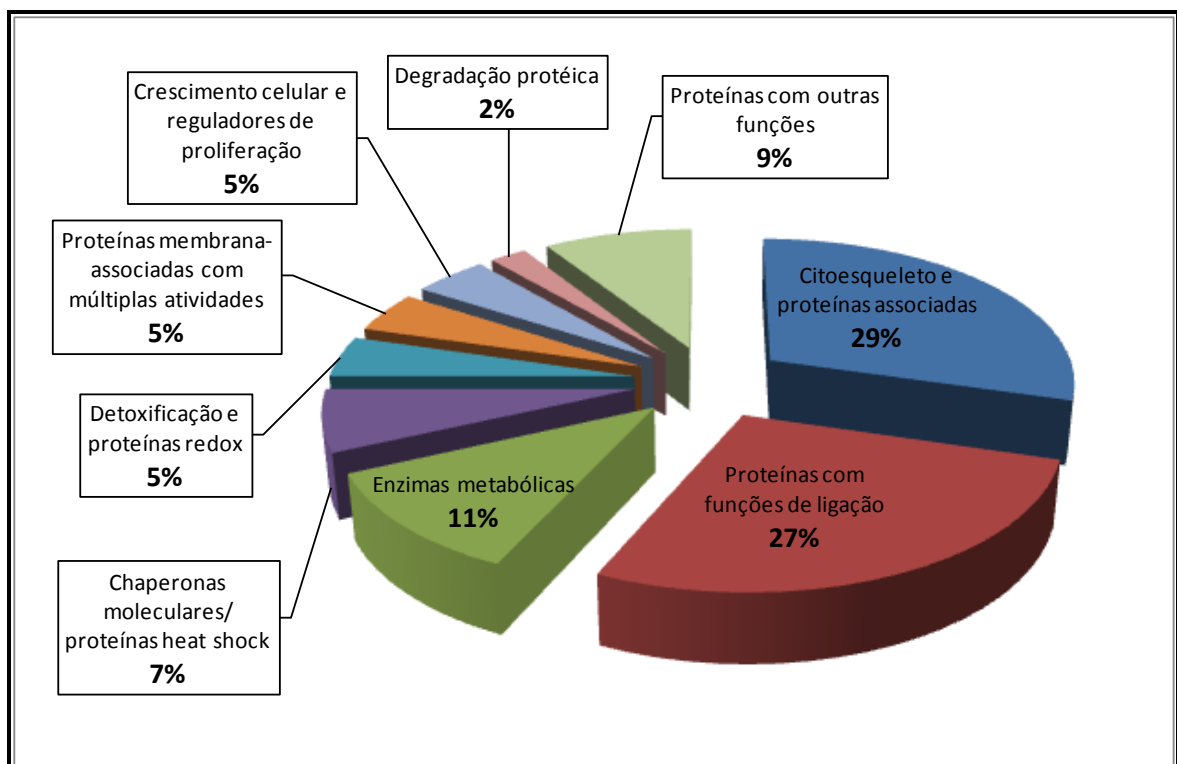


GRÁFICO 1 - CLASSIFICAÇÃO DAS PROTEÍNAS IDENTIFICADAS

FONTE: O autor (2010)

NOTA: Gráfico representando a distribuição das 41 proteínas identificadas de acordo com a função biológica. A classificação foi feita com base nas informações disponíveis nos bancos de dados do NCBI e do *UniProtKB / Swiss-Prot*.

## 7. DISCUSSÃO

A análise proteômica tem sido utilizada em diversas áreas e com variadas finalidades. O potencial demonstrado por esta metodologia, para complementar as informações obtidas com o seqüenciamento de genomas e para fornecer um novo panorama sobre as interações intracelulares, abriu muitas portas tanto para a bioquímica e a biologia celular como para a genética, tendo se tornado uma ferramenta importante na área da saúde humana. A compreensão dos mecanismos moleculares envolvidos no desenvolvimento de diversas doenças parece mais próxima. A busca por marcadores protéicos que possibilitem a detecção precoce e a intervenção terapêutica em estágios mais responsivos de várias doenças está em destaque atualmente.

O câncer de mama, por ser o segundo mais incidente entre as mulheres e o quinto tipo com maior mortalidade no mundo, é foco de diversos trabalhos. Além dos fatores genéticos, já são conhecidos muitos fatores de risco envolvidos na sua origem e progressão. A idade da menarca, idade da menopausa e idade da primeira gravidez, por exemplo, influenciam o tempo de exposição da mulher ao estrogênio endógeno (PARK *et al.*, 2003). Estes fatores, somados à influência de hábitos e exposição à carcinógenos ao longo da vida, acarretam uma série de variáveis no estudo deste tipo de câncer.

Quando comparamos o tecido mamário tumoral com o tecido não tumoral, uma questão fisiológica essencial surge: o que é um tecido mamário normal? Devido às alterações hormonais durante o ciclo menstrual, o perfil do tecido mamário não tumoral também está sendo alterado e, portanto, torna-se difícil definir um estado referencial (HONDERMARCK *et al.*, 2008). Para compreender o que está alterado no funcionamento do tecido tumoral é necessário o conhecimento sobre o respectivo tecido saudável em diferentes condições fisiológicas. A análise proteômica de tecido mamário não tumoral humano é, em grande parte, realizada apenas para estudos comparativos, sendo difícil encontrar na literatura informações sobre proteínas expressas neste tecido que não tenham relação aparente com o câncer de mama.

Para este estudo foram utilizadas amostras de tecido mamário não tumoral obtido de mamas acometidas pelo câncer. A amostra foi retirada do quadrante

oposto ao tumor, excedendo consideravelmente a margem de segurança. Na literatura científica não existe ainda uma definição bem estabelecida de margem cirúrgica. Contudo, a mais aceita é que a partir de uma determinada distância do tumor, a margem passa a ser classificada de negativa. São necessários novos estudos para adequar o tratamento cirúrgico à segurança oncológica, enquanto isso a margem de segurança indicada é de no mínimo 1,0 cm em pacientes tratadas conservadoramente (ROSOSCHANSKY e SOARES, 2007).

Mesmo tomando precauções para que o material estudado fosse não tumoral, a progressão patológica é lenta, sendo estimado que o desenvolvimento do tumor a um estado clinicamente detectável possa durar entre seis e oito anos (TOILLON *et al.*, 2007). Portanto, o tecido que no momento é considerado não tumoral pode estar no início da carcinogênese, e já surtir efeito na expressão protéica, retornando à questão sobre a dificuldade de estabelecer um perfil de normalidade.

### 7.1. IMPLANTAÇÃO DA METODOLOGIA PROTEÔMICA

Por ser o método mais freqüentemente encontrado na literatura científica, a idéia inicial para a coleta do material era de congelamento em nitrogênio líquido logo após a remoção por cirurgia. Isso não ocorreu por dois motivos: (1) as amostras eram provenientes de dois hospitais (Hospital de Clínicas e Hospital Nossa Senhora das Graças), sendo necessária a manutenção de recipientes com nitrogênio líquido em ambos. A aquisição destes recipientes levaria algum tempo e algumas amostras teriam que ser processadas por outro método, não havendo posteriormente uma homogeneidade quanto ao método de coleta; (2) O congelamento deveria ser igualmente realizado para todas as amostras de forma gradual, pois do contrário poderia prejudicar a integridade da amostra. A equipe cirúrgica nem sempre é a mesma e isso dificultaria a padronização para o congelamento adequado, podendo diferir entre as amostras. Desta forma, as amostras eram coletadas, acondicionadas em tubos de polipropileno estéreis e imediatamente transportadas ao Laboratório de Citogenética Humana e Oncogenética.

No laboratório, a amostra era processada antes do armazenamento a 80 °C negativos. Esse processamento consistiu na remoção do excesso de tecido adiposo que circundava o tecido glandular. Por não ser possível a completa remoção do tecido adiposo e havendo contaminação por pequenos derrames vasculares, assumiu-se a heterogeneidade celular da amostra. As técnicas de microdissecção como, por exemplo, a microdissecção por captura a laser (LCM), permite a obtenção de populações celulares puras de seções de tecidos heterogêneos (POSADAS *et al.*, 2005). Mas o entusiasmo inicial vindo com esta técnica por enquanto foi reduzido devido a alguns fatores. Estudos proteômicos requerem uma quantidade inicial de material significativa, já que não é possível utilizar técnicas de amplificação em proteínas, e seriam necessárias no mínimo 100 mil células para a análise. O segundo problema é que a integridade bioquímica das proteínas também poderia ser alterada devido ao uso de fixadores e pelo tempo necessário para realizar a microdissecção (HONDERMARCK *et al.*, 2008). Outro aspecto que deve ser levado em consideração é que isolando tipos celulares perdem-se informações sobre as interações entre os componentes do microambiente no qual as células estão.

A aplicação das técnicas e metodologias demonstrada nos resultados deste trabalho, desde a extração protéica da amostra (passando pela confecção dos géis bidimensionais e análise pelo *ImageMaster™ 2D Platinum v6.0*, excisão das bandas selecionadas e digestão trípica, e pela espectrometria de massas) até a identificação de proteínas no MASCOT, demonstra que a abordagem proteômica foi implantada no Laboratório de Citogenética Humana e Oncogenética.

## 7.2. QUANTIFICAÇÃO PROTÉICA

Nesta etapa chamou atenção a baixa concentração de proteínas obtida a partir da amostra CP633 N (3,8µg/µl) quando comparada às demais. Contudo, este fato não se mostrou relevante visto que esta foi uma das amostras com maior número de bandas detectadas.

### 7.3. ESPECTRO DE MASSAS

Após a obtenção dos espectros de massas houve o processamento destes no programa *FlexAnalysis 2.0 (Bruker Daltonics)* (FIGURA 4). Neste programa, os picos de massa de contaminantes conhecidos foram removidos. Ao observar as condições do espectro de massas, antes mesmo de submeter as informações no programa de busca, já é possível ter idéia da qualidade da identificação devido à quantidade de picos de massa, intensidade, sinais de matriz e ruído. Influenciam nestes aspectos fatores como: a intensidade do laser para ionização dos peptídeos; a correta calibração (externa e interna) do equipamento e a quantidade e condição da amostra. Todos estes fatores são importantes para determinar uma identificação confiável.

### 7.4. PROTEÍNAS IDENTIFICADAS

Baseado na comparação entre os valores teóricos de massa (MM) e ponto isoelétrico (pI) para cada proteína e os valores obtidos nos géis, observou-se que a maioria das proteínas identificadas e suas respectivas bandas (76%) foram observadas na posição esperada nos géis ( $\Delta MM < 20\%$  e  $\Delta pI < 0,6$ ) (DUPONT *et al.*, 2005) (TABELA 2). As proteínas que ficaram fora da variação esperada serão discutidas posteriormente. Em alguns casos, diferentes bandas foram identificadas como sendo a mesma proteína. As diferenças entre MM e/ou pI teórica e observada entre estas bandas são devido a presença de diferentes isoformas da mesma proteína. Aproximadamente 30% das 44 proteínas identificadas estavam representadas por mais de uma banda, sugerindo a ocorrência de modificações pós-traducionais (MPTs) que alteram a massa molecular e o pI das proteínas (MAYR *et al.*, 2005). As MPTs, às quais muitas proteínas estão sujeitas, são freqüentemente necessárias para que estas exerçam sua função biológica ou sejam inativadas. Além da localização, estas modificações também podem alterar a meia-vida da proteína (ONG e PANDEY, 2001).

Das modificações pós traducionais já relatadas no banco de dados *UniProtKB / Swiss-Prot* para as proteínas identificadas, as mais freqüentes foram a

fosforilação, glicosilação e acetilação. Entre as diferentes modificações pós traducionais, a fosforilação é a mais comum e melhor estudada, sendo importante em diversos processos celulares (KUMAR *et al.*, 2004). Esta modificação altera o pI, de acordo com a quantidade de resíduos fosforilados, deslocando-o no sentido ácido. Sabe-se que a fosforilação causa maiores variações em proteínas com pI acima de 7,0 (ZHU *et al.*, 2005). A glicosilação é a MPT mais variada, tanto em relação aos tipos de aminoácidos que são modificados quanto às estruturas anexadas a eles. Estas modificações, responsáveis pelo endereçamento das proteínas, são muito comuns nas proteínas secretadas, ocorrendo na maioria das proteínas do soro (LIS e SHARON, 1993). Por anexar cadeias de açúcares às proteínas, as glicosilações são as MPTs que mais influenciam na variação de massa. Löster e Kannicht (2008), através da deglicosilação de proteínas, verificaram que a glicosilação pode variar o padrão de banda das proteínas no gel em aproximadamente 0,4 unidades de pH e 5-8 kDa. A acetilação acarreta na mudança de pI no sentido ácido. Esta alteração é devida ao aumento das cargas negativas como resultado da substituição de grupamentos amino, positivamente carregados, de lisinas por grupos hidrofóbicos acetil com carga neutra (RAMOS e BORA, 2004).

Até chegar à identificação, algumas bandas precisaram ser retiradas de mais de um gel para digestão tríptica. Geralmente isso foi necessário para bandas muito fracas e/ou pequenas (ex.: 157, 158, 164, 170, 264, 453, 460, 492, entre outras) contendo pouca quantidade de proteína para a análise, mas algumas bandas bastante evidentes também geraram dificuldades na identificação (375, 491 e 497).

As bandas que permaneceram sem identificação ou que apresentaram *score* ou percentual de cobertura muito baixos foram submetidas à purificação peptídica. Houve melhora na identificação apenas para as bandas 145, 189, 320, 362, 369, 379, 392, 415, 433 e 514. Teoricamente, a purificação deveria sempre melhorar a identificação (ainda mais se for o produto de digestão tríptica de uma proteína que já tenha sido identificada sem a purificação), mas podem ocorrer falhas. Segundo o fabricante, devido à natureza extremamente hidrofóbica da resina, é importante que não haja circulação de ar dentro da ponteira *Eppendorf® PerfectPure C-18*, pois isso cria uma contra-pressão (*backpressure*) e prejudica a qualidade dos resultados. É possível que isto tenha ocorrido durante o procedimento, mas a condição do material

(pouca quantidade, degradação, homogeneização inadequada, etc.) também pode interferir.

Algumas bandas continham mais de uma proteína, talvez representando proteínas co-migrantes com estruturas primárias ou peso molecular similares. Por exemplo, as bandas 359, 362 e 367 foram identificadas como actina citoplasmática 1 e/ou actina citoplasmática 2, ambas apresentam homologia. A mesma situação ocorre com a banda 361, identificada como três tipos de actinas. As bandas 285 e 417 continham duas proteínas diferentes, mas como massa molecular e pI próximos que a resolução da tira de pH utilizada não permitiu a separação. A banda 364 também apresentou mistura de duas proteínas, mas neste caso, os dados teóricos de MM da albumina sérica diferiam bastante da creatina quinase tipo-B (71,3 kDa e 42,9 kDa). Isso pode ser explicado pela detecção da isoforma 2 da albumina sérica, explicado posteriormente.

As proteínas foram separadas em classes, baseado em um estudo de Minafra e colaboradores (2006), de acordo com as suas principais funções biológicas. A organização das proteínas em categorias funcionais é fundamental para o desenvolvimento de mecanismos válidos para o estudo das funções bioquímicas das células e seu envolvimento no câncer (MINAFRA *et al.*, 2006). No entanto, estas classificações funcionais devem ser tratadas com cuidado, pois elas estão sujeitas a mudanças, e geralmente as proteínas estão relacionadas a diversas funções (OU *et al.*, 2008).

#### 7.4.1. Citoesqueleto e Proteínas Associadas

A maior classe de proteínas identificadas é a que engloba proteínas do **citoesqueleto e proteínas associadas (30%)**. O citoesqueleto é um sistema complexo e dinâmico que compreende proteínas estruturais formando polímeros (actina, tubulina, filamentos intermediários e diversas proteínas associadas com funções regulatórias). Está direta ou indiretamente envolvido em diversos processos celulares importantes como, por exemplo, a polaridade celular, adesão, transporte intracelular, transdução de sinal e movimento celular (MINAFRA *et al.*, 2006).

Treze proteínas identificadas a partir de 17 bandas compuseram esta classe. Actina e vimentina apresentaram várias isoformas, quatro e seis respectivamente.

As actinas são proteínas extremamente conservadas evolutivamente e são fundamentais para as células eucarióticas. Por serem muito semelhantes entre si a distinção por impressão digital peptídica (PMF) acaba sendo difícil. É comum encontrar em trabalhos científicos, que utilizam PMF para identificação, a detecção dessas proteínas simultaneamente em uma mesma banda (DUPONT *et al.*, 2005; MINAFRA *et al.*, 2006). As actinas citoplasmáticas 1 e 2 são codificadas pelos genes *ACTB* (cromossomo 7 em 7p15-p12) e *ACTG* (cromossomo 17 em 17q25) respectivamente. Possuem homologia, diferindo apenas em quatro aminoácidos na região amino-terminal, e foram identificadas simultaneamente nas bandas 359, 362 e 367 com os mesmos valores de *score*, cobertura de seqüência e número de massas pareadas. Apesar das bandas poderem corresponder a apenas um dos tipos, pode ser também que os dois tipos estejam presentes na amostra. Portanto, tanto a actina citoplasmática 1 quanto a citoplasmática 2 foram consideradas na determinação do número total de proteínas identificadas. Com a banda 361 ocorreu algo semelhante, mas foram indicados três tipos de actina simultaneamente na identificação: (a) do músculo liso aórtico, codificado pelo gene *ACTA2* localizado no cromossomo 10 (10q23.3); (b) do músculo cardíaco alfa, codificado pelo gene *ACTC1* localizado no cromossomo 15 (15q11-q14); e (c) do músculo esquelético alfa, codificado pelo gene *ACTA1* localizado no cromossomo 1 (1q42.13). Os três tipos foram considerados na determinação do número total de proteínas identificadas. Para distinção entre os tipos de actina seria necessária a obtenção da seqüência de aminoácidos através de espectrometria de massas em *tandem* (MS/MS).

A vimentina (bandas 275, 276, 346, 351, 355 e 378) e a queratina do citoesqueleto tipo I - 17 (banda 333), também chamada de citoqueratina 17, são membros da família dos filamentos intermediários, um dos três maiores componentes do citoesqueleto nas células eucarióticas, juntamente com os microtúbulos e microfilamentos. São cruciais nas funções biomecânicas da célula e têm sido referidos como o “cinto de segurança da célula” por prevenir o estiramento excessivo da mesma (QIN *et al.*, 2009). Proteínas da família dos filamentos



intermediários são variavelmente expressas em diferentes tipos celulares e seus tumores correspondentes.

Citoqueratinas são um tipo de polipeptídeo fibroso com diâmetro de 7-11nm e usualmente expressas em células mioepiteliais e epiteliais luminais (CHEN *et al.*, 2008). Em 1992, um estudo de Troyanovsky e colaboradores mostrou que a citoqueratina 17 parece ser um marcador de diferenciação de células basais em epitélios complexos e, portanto indicativo de um certo tipo de “células tronco” epiteliais. Esta proteína é codificada pelo gene *KRT17*, localizado no cromossomo 17 em 17q12-q21. TRASK e colaboradores (1990) observaram que o padrão geral de expressão da proteína queratina é muito diferente em linhagens celulares normais e tumorais. As queratinas produzidas nas células normais foram K5, K6, K7, K14 e K17, enquanto que nas células tumorais foram K8, K18, K19.

A vimentina, codificada pelo gene *VIM* localizado no cromossomo 10 em 10p13, é expressa em tecido estromal conectivo, parede dos vasos sanguíneos, nervos e também nas células mioepiteliais (CHEN *et al.*, 2008). Esta proteína também está envolvida na resposta imune, no transporte de colesterol LDL, assim como na organização de diversas proteínas associadas à adesão, migração e sinalização celular (NCBI, 2010). De acordo com informações disponíveis no banco de dados *UniProtKB / Swiss-Prot*, a vimentina (53,7 kDa/ pI 5,06) é uma proteína que pode sofrer modificações pós-traducionais por acetilação e fosforilação. (FIGURA 9). Minafra e colaboradores (2006), em trabalho com linhagem celular neoplásica de tecido mamário, encontraram 7 isoformas de vimentina com uma variação de MM e pI entre elas similar ao encontrado no presente trabalho. O autor sugere que possivelmente a variação de bandas seja devido à modificações pós traducionais.

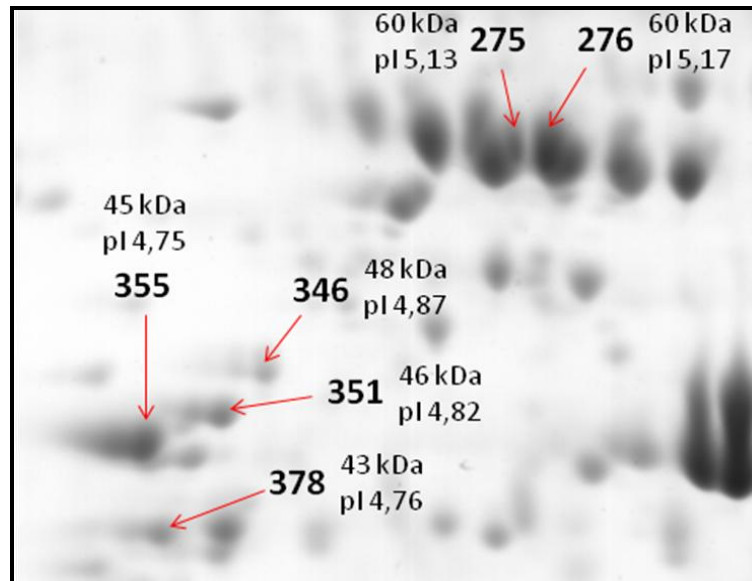


FIGURA 9 – DETALHE DAS BANDAS DA PROTEÍNA VIMENTINA

FONTE: O autor (2010)

LEGENDA: Os dados de pI e massa experimentais estão demonstrados para as bandas 275, 276 e 346, 351, 355 e 378.

Os microtúbulos também estão envolvidos em diversos processos essenciais, incluindo a divisão celular, motilidade ciliar e flagelar e transporte intracelular (McKEAN *et al.*, 2001). A tubulina é a principal constituinte dos microtúbulos. A banda 298 foi identificada como cadeia beta da tubulina com o *score* bastante alto (216). Para esta mesma banda, as cadeias beta-2A, 2B, 2C, 3 e 4 obtiveram *scores* 162, 162, 142, 110 e 89 respectivamente, todos acima do *score* mínimo (56). Com a banda 292 ocorreu a mesma situação. Apesar da tubulina cadeia beta-2C ter obtido *score* 99, as tubulinas cadeia-4, 2A e 2B obtiveram *score* de 70, 56 e 56 respectivamente. A seqüência de aminoácidos entre as diferentes cadeias beta da tubulina é bastante similar, gerando vários peptídeos coincidentes após a digestão trípica. Mesmo assim, o *score* e a cobertura de seqüência da cadeia beta da tubulina (banda 298), da cadeia beta-2C (banda 292), codificadas respectivamente pelos genes *TUBB* (cromossomo 6 em 6p21.33) e *TUBB2C* (cromossomo 9 em 9q34) (NCBI, 2010) foram consideravelmente maiores do que das demais proteínas sugeridas.

A actina, tanto como monômero quanto polímero, se liga a uma ampla variedade de proteínas-acessórias nas células de eucariotos. Neste estudo, foram encontradas duas proteínas ligantes de actina: cadeia beta da tropomiosina (gene *TPM2*, localizado no cromossomo 9 em 9p13.2-p13), identificada na banda 411, e a

cadeia alfa-4 da tropomiosina (gene *TPM4* localizado no cromossomo 19 em 19p13.1) identificada na banda 441. Ambas pertencem à família das tropomiosinas e se ligam a filamentos de actina tanto em células musculares como em não-musculares. Desempenham papel central, juntamente com o complexo troponina, na regulação cálcio-dependente da contração dos músculos estriados em vertebrados. Em células não-musculares estão associadas à estabilização dos filamentos de actina no citoesqueleto. Em 2005, Dupont e colaboradores identificaram estas duas proteínas em estudo de células de tecido muscular liso obtido de artérias mamárias. A partir desta informação, é possível inferir que, além de provenientes das células mioepiteliais que constituem a glândula mamária, estas proteínas podem ser provenientes da contaminação da amostra por microvasos sanguíneos que entremeiam o tecido mamário. De acordo com o banco de dados *UniProtKB / Swiss-Prot*, isoformas, originadas por *splicing* alternativo, destas duas cadeias da tropomiosina já foram relatadas. Neste trabalho, a MM observada para as cadeias beta e alfa-4 da tropomiosina foi de 38 kDa e 33 kDa respectivamente, enquanto o esperado teórico era de 32,9 kDa e 28,6 kDa (variação de 13,4% e 13,3%). Para a cadeia alfa-4, observou-se que a variação de massa quando comparada com a isoforma-2 (MM = 32,72), relatada no banco de dados, caiu para 0,8%, evidência de que a banda 411 corresponde a isoforma-2 da cadeia alfa-4 da tropomiosina. Apesar de isoformas para a cadeia beta também terem sido descritas, nenhuma possui MM mais próxima da observada para a banda 411, mas isso não exclui a possibilidade de que a forma encontrada seja decorrente de *splicing* alternativo.

A proteína inibidora de dissociação Rho GDP 1 (GDIR1) é codificada pelo gene *ARHGDI1*, localizado no cromossomo 17 (17q25.3), e pertence à família das GDIs (inibidores da dissociação do nucleotídeo guanina) que inclui também a D4-GDI e a RhoGDI-3. É uma proteína regulatória celular que atua primariamente controlando a distribuição e atividade das Rho-GTPases (ZHANG *et al.*, 2005). As Rho-GTPases de mamíferos compreendem uma família de 20 moléculas sinalizadoras intracelulares, documentadas pelo seu importante papel na regulação da actina no citoesqueleto. A maioria das Rho-GTPases alterna entre a forma ativa (GTP-ligado) e inativa (GDP-ligado) e este ciclo é regulado pelos fatores de troca de nucleotídeo guanina (GEFs) e pelas proteínas ativadoras de GTPase (GAPs), além das GDIs (HEASMAN e RIDLEY, 2008). A GDIR1 é expressa em variadas células e

tecidos e se liga à maioria das Rho-GTPases. Aceita-se que a regulação negativa realizada pela GDIR1 seja exercida de três maneiras: (a) restringindo a GTPase ao citosol, localização na qual permanece inativa; (b) impedindo a ligação com GEFs, inibindo a ativação GTPase; e (c) bloqueando a ligação a alvos moleculares *downstream* (ZHANG *et al.*, 2005). A super expressão de GDIR1 em várias linhagens celulares induz o desligamento das actinas do citoesqueleto e a perda da aderência ao substrato (ZHANG *et al.*, 2005). Apesar de baixo, o *score* para esta proteína foi suficiente (57) e todos os outros critérios avaliados foram atendidos.

A proteína tipo-actina 3B (banda 497) desempenha funções na organização da actina no citoesqueleto. Codificada pelo gene *ACTR3B*, localizado em 7q36.1, esta proteína também pode estar associada com a redução do potencial metastático de tumores (NCBI, 2010). Esta proteína se apresenta em duas isoformas relatadas até agora (*UniProtKB / Swiss-Prot*), mas nenhuma com a massa molecular próxima a encontrada neste trabalho, sendo que a variação de MM não fica abaixo de 52%. Pode ser que a forma identificada seja proveniente de um novo *splicing* alternativo, mas como a cobertura de seqüência e o *score* também foram baixos (14% e 62) a identificação foi tida como duvidosa. Outro fato que tira ainda mais a credibilidade da identificação é que a mesma banda, retirada de outro gel, acusou a superóxido desmutase [Cu – Zn] (*score* = 46; MM = 16,15 kDa; pI = 5,7; % de cob. = 32%; e 3 peptídeos pareados) como provável identificação. Mas apesar da MM próxima do teórico e cobertura de seqüência acima de 30%, o *score* foi abaixo de 56 e teve apenas 3 dos 21 peptídeos gerados por digestão trípica pareados com os peptídeos resultantes de clivagem *in silico*. A purificação protéica foi realizada para o produto da digestão trípica que originou as duas identificações prévias, mas não houve melhora em nenhuma das identificações.

A maioria destas proteínas, as quais desempenham funções regulatórias sobre o citoesqueleto, estão envolvidas em muitos processos celulares e a expressão diferencial entre células normais e tumorais da mama tem sido demonstrada (MINAFRA *et al.*, 2006).

#### 7.4.2. Proteínas com Funções de Ligação

Com 28 bandas correspondendo a 12 proteínas e suas isoformas, a classe das **proteínas com funções de ligação (27%)** é a segunda mais abundante. As proteínas identificadas neste grupo foram: antitrombina (banda 260), alfa-1 antitripsina (bandas 267, 272, 283, 285, 617 e 618), sorotransferrina (bandas 192, 195, 196, 198 e 201), haptoglobina (bandas 369, 379 e 386), albumina do soro (230, 271 e 364), cadeia beta do fibrinogênio (bandas 392 e 394), cadeia gama do fibrinogênio (bandas 320 e 317), proteína ligante à vitamina-D (bandas 285 e 286), apolipoproteína A-1 (banda 478), componente amilóide-P sérico (banda 459), transtiretina (banda 514) e ceruplasmina (banda 156 e 157). Muitas destas proteínas têm sua função de ligação associada ao transporte de moléculas.

A antitrombina III e a alfa-1 antitripsina pertencem à família serpina (*serine protease inhibitor*), mas, apesar da origem da palavra serpina, nem todas as proteínas inibidoras de protease serina pertencem a esta família (PATSTON, 2000). Os inibidores de protease serina estão envolvidos em processos biológicos como a coagulação sanguínea, ativação do sistema complemento, apoptose, angiogênese, inflamação e supressão de tumor (van GENT *et al.*, 2003).

A alfa-1 antitripsina, codificada pelo gene *SERPINA1* localizado em 14q32.1 (NCBI, 2010), tem como alvo primário a elastase, mas também tem afinidade moderada pela plasmina e pela trombina (DENG *et al.*, 2006). Esta inibidora de protease está em alta concentração no leite humano, particularmente no início da lactação (LÖNNERDAL, 2010). As seis bandas identificadas como alfa-1 antitripsina (FIGURA 10) tiveram variação de massa (acima do valor teórico) e de pI (abaixo do valor teórico), mas apenas as bandas 267, 272 e 617 superaram 20% de variação da MM teórica. Apesar disso, o *score* e o percentual de cobertura destas bandas foram bastante bons. Uma possível explicação para a variação de massa pode ser o *splicing* alternativo, gerando um mRNA que, no caso, codificaria para uma proteína maior. O banco de dados *UniProtKB / Swiss-Prot* apresenta três formas descritas para esta proteína originadas de *splicing* alternativo, mas nenhuma delas com a mesma massa observada neste trabalho. Glicosilações também já foram relatadas para esta proteína, mas a variação de MM foi muito grande (aproximadamente 12 kDa) para ser atribuída apenas a este fator. Quanto à variação do pI, possivelmente

seja devida às modificações pós-traducionais. A banda 285 foi identificada como mistura da proteína alfa-1 antitripsina e com a proteína ligante de vitamina D.

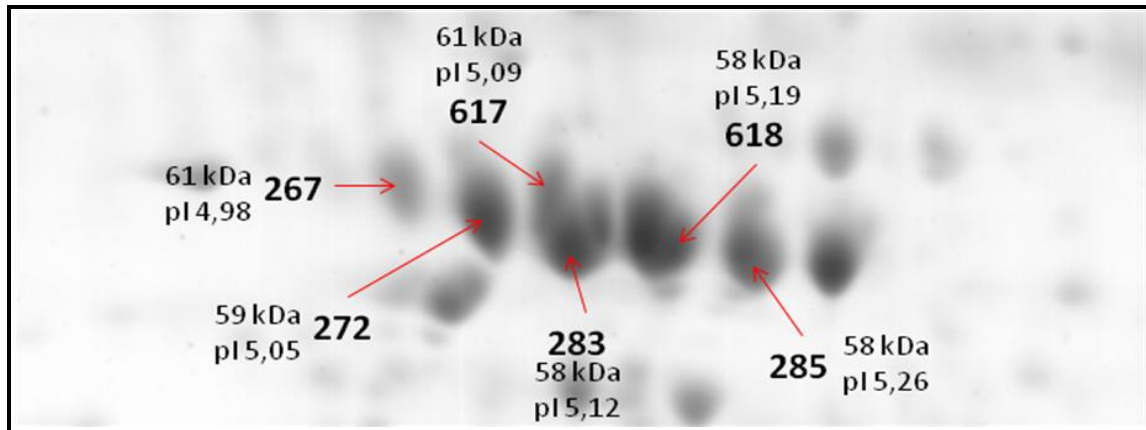


FIGURA 10 – DETALHE DAS BANDAS DA PROTEÍNA ALFA-1 ANTITRIPSINA

FONTE: O autor (2010)

LEGENDA: Os dados de pI e massa experimentais estão demonstrados para as bandas 267, 272, 283, 285, 617 e 618.

A antitrombina III, codificada pelo gene *SERPINC1* localizado em 1q23-q25.1 é o inibidor de protease serina mais importante no plasma. Inibe a trombina, e outras proteases serina ativadas, do sistema de coagulação, regulando a cascata de coagulação (NCBI, 2010). O pI teórico para esta proteína é de 6,32, mas o pI experimental foi de 5,32. O alto score obtido na identificação (114) e os demais critérios considerados levam a crer que a proteína foi corretamente identificada, sendo a variação de pI provavelmente devida às modificações pós-traducionais. De acordo com o banco de dados *UniProtKB / Swiss-Prot*, esta proteína pode sofrer fosforilações.

O ferro desempenha um papel central na replicação do DNA por ser necessário como co-fator de uma das enzimas chave deste processo, a ribonucleotídeo redutase (GOMME e McCANN, 2005). Contudo, o ferro livre no plasma pode ser tóxico por causar danos oxidativos aos tecidos, sendo essencial seu transporte em uma forma redox-inativa. Duas proteínas desta classe estão relacionadas com o transporte do ferro: a soro transferrina e a haptoglobina.

A soro transferrina, codificada pelo gene *TF* localizado no cromossomo 3 em 3q22.1, tem a função de se ligar ao ferro e transportá-lo do intestino, sistema retículo-endotelial e células parenquimais do fígado para todas as células em

proliferação no corpo. (NCBI, 2010). Em uma revisão de literatura, Gomme e McCann (2005) apresentam implicações desta proteína também em atividades de crescimento e diferenciação celular. Por esse motivo a sorotransferrina e seus receptores têm sido usados como alvo na intervenção farmacêutica, principalmente para o diagnóstico na inflamação e terapia em câncer (MACEDO e de SOUZA, 2008). A síntese da sorotransferrina é afetada pela redução da reserva de ferro e pelo estímulo de estrogênio endógeno e exógeno (RITCHIE *et al.*, 1999). As cinco bandas identificadas como sorotransferrina (FIGURA 11) apresentaram valores decrescentes de pI entre elas de 6,89 até 6,55. De acordo com o banco de dados *UniProtKB / Swiss-Prot*, várias modificações pós-traducionais já foram observadas para esta proteína, por exemplo a fosforilação, e estas modificações provavelmente são as responsáveis por esta variações de pI.

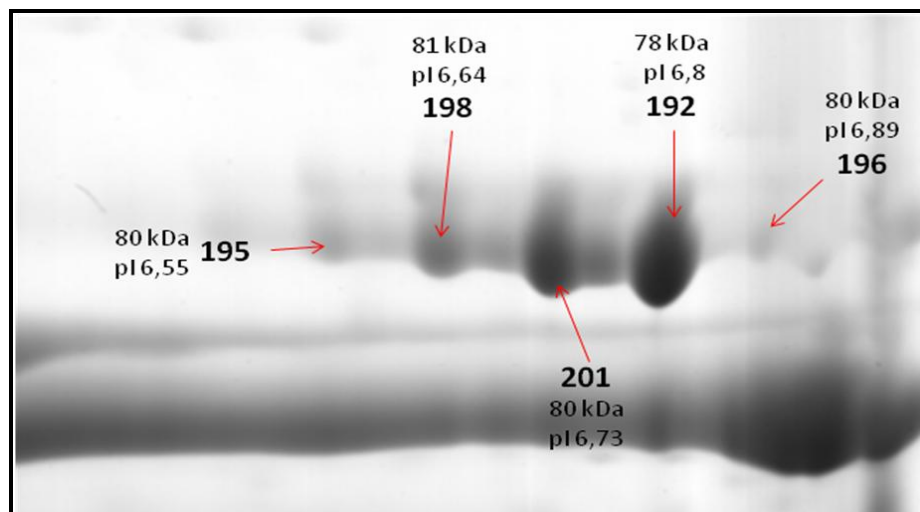


FIGURA 11 – DETALHE DAS BANDAS DA PROTEÍNA SOROTRANSFERRINA

FONTE: O autor (2010)

LEGENDA: Os dados de pI e massa experimentais estão demonstrados para as bandas 192, 195, 196, 198 e 201.

Localizado no cromossomo 16 em 16q22.1, o gene *HP* codifica uma pré-proteína, a qual é processada para gerar cadeias alfa e beta que subsequentemente são combinadas em tetrâmero para originar a haptoglobina (NCBI, 2010). A biossíntese da haptoglobina ocorre no fígado, mas também no tecido adiposo e no pulmão, provendo atividade antioxidante e antimicrobiana (DOBRYSZYCKA, 1997). Esta proteína é secretada no plasma, onde se combina com as hemoglobinas livres prevenindo a eliminação de ferro pelos rins e o dano renal enquanto as tornam

acessíveis às enzimas degradantes (WASSEL, 2000). Ainda desempenha um papel na resposta de fase aguda modulando vários aspectos (WASSEL, 2000). As três isoformas da haptoglobina observadas nos géis apresentam seus pontos isoelétricos deslocados no sentido ácido. Apesar da variação ser acima de 0,6 unidades de pH os demais parâmetros sustentaram a identificação.

Na região 18q12.1 do cromossomo 18, encontra-se o gene *TTR*, o qual codifica para a proteína transtiretina (NCBI, 2010). A transtiretina é uma proteína secretada no plasma, principalmente pelo fígado, que se liga aos hormônios da tireóide transportando-os pelo plasma (INGENBLEEK e YOUNG, 1994). A transtiretina, assim como a sorotransferrina, tem sua síntese influenciada pela redução da reserva de ferro e pelo estímulo de estrogênio endógeno e exógeno (RITCHIE *et al.*, 1999).

A vitamina D afeta o ciclo celular, a apoptose, os receptores hormonais e a angiogênese, todos fatores relacionados ao desenvolvimento tumoral, progressão e metástase (AUNG *et al.*, 2009). A proteína ligante de vitamina D, codificada pelo gene *GC*, localizado no cromossomo 4 em 4q12-q13 (NCBI, 2010), é uma proteína multifuncional e altamente expressa do plasma (WHITE e COOKE, 2000). Esta proteína possui um alto grau de polimorfismo (com mais de 120 espécies detectadas) e, apesar de sintetizada predominantemente no fígado, é detectada em outros fluidos além do plasma, por exemplo, no fluido cérebro-espinhal, no leite e também na superfície de muitas células (GOMME e BERTOLINI, 2004). A principal função desta proteína é se ligar à vitamina D e seus metabólitos no plasma e transportá-los aos tecidos alvo (VERBOVEN *et al.*, 2002).

O fibrinogênio é uma glicoproteína formada por três pares de cadeias polipeptídicas não-idênticas designadas alfa, beta e gama (MOESSON *et al.*, 2001). É uma macromolécula solúvel, mas forma um gel insolúvel (coágulo) quando clivada em fibrina pela trombina (WEISEL, 2005). Dentre as funções do fibrinogênio (e também da fibrina) destacam-se: a atuação na coagulação sanguínea; interações célula-matriz; resposta inflamatória; regeneração de tecidos; e neoplasia (MOESSON, 2005). O fibrinogênio pode se ligar, com alta afinidade, ao receptor de integrina nas plaquetas ativadas e também possui sítios de ligação aos íons cálcio, que são importantes para a estabilidade estrutural e funcional. Os sítios de alta afinidade pelo cálcio estão presentes na cadeia gama (WEISEL, 2005). Os genes



*FGB* e *FGG*, ambos localizados nos cromossomos 4 (4q28), codificam para a cadeia beta e para a cadeia gama do fibrinogênio respectivamente. O principal sítio de ligação para os receptores de plaquetas está na cadeia gama. O *splicing* alternativo do gene *FGG* gera dois transcritos variantes que codificam diferentes isoformas (NCBI, 2010). De acordo com o banco de dados *UniProtKB / Swiss-Prot*, o gene *FGB* é polimórfico. As bandas identificadas como cadeia beta do fibrinogênio (392 e 394) apresentaram uma variação de MM de 27,6% podendo ser devido *splicing* alternativo, embora nenhum caso tenha sido relatado até o momento. O *score* da para a banda 392 foi de 204 e o percentual de cobertura foi de 38%, sendo que estes valores foram obtidos após purificação peptídica (na identificação sem purificação, apesar do *score* ter sido satisfatório, o percentual de cobertura havia sido abaixo de 20%). Para a banda 394 foram obtidos valores altos para *score* e percentual de cobertura sem a necessidade de purificar o produto da digestão trípica. O fato que chamou a atenção foi a enorme variação de *pI* (acima de duas unidades de *pH*) na identificação destas duas bandas, sendo que o *pI* teórico da cadeia beta do fibrinogênio é de 8,54. Apesar de não relatadas até agora no banco de dado consultado, sabe-se que a fosforilação causa maiores variações em proteínas com *pI* acima de 7,0 (ZHU *et al.*, 2005).

A Ceruplasmina é uma metaloproteína que se liga a 95% do cobre (Cu) no plasma e é referida como uma proteína de fase aguda (SOKOLOV *et al.*, 2006). Suas principais funções são de reduzir o oxigênio molecular (formando água) e oxidar os íons ferrosos transformando-os na forma menos tóxica, sem liberar espécies reativas de oxigênio. Estas atividades enzimáticas conferem à ceruplasmina um papel importante na homeostase do ferro (Fe) e também um poder relevante como antioxidante (ALTAMURA *et al.*, 2009). A ceruplasmina humana é extremamente suscetível à proteólise (SOKOLOV *et al.*, 2006). Apesar de o fígado ser considerado a fonte primária de expressão, o gene *CP* (que codifica para a ceruplasmina), localizado no cromossomo 3 (3q23-q25) (NCBI, 2010), é expresso em outros locais. Em estudo com ratos, Jaeger e colaboradores (1991) observaram a expressão abundante do mRNA da ceruplasmina em células da glândula mamária. As duas identificações para esta proteína apresentaram *score* muito baixo (bandas 156 com *score* 58 e banda 157 com *score* 53) e foram submetidas à purificação peptídica, sem melhora na identificação. O único critério que foi atendido para esta

proteína foi a comparação entre MM e pl teórico e observado, a qual gerou pequena variação. Por ser uma molécula grande (115 kDa observado) pode justificar o baixo percentual de cobertura de seqüência. Há ainda a questão de não haver marcador de massa acima de 97 kDa e com isso, a estimativa para proteínas com massa superior a este valor pode ser comprometida. Contudo, as bandas da ceruplasmina (MM teórica de 123 kDa) deveriam estar acima das bandas da cadeia alfa-1 do colágeno VI (MM teórica de 109 kDa), salvo fosse uma isoforma de *splicing* alternativo, sendo que nenhuma foi relatada até agora de acordo com o banco de dados *UniProtKB / Swiss-Prot*. A identificação desta proteína, em suas duas bandas foi considerada duvidosa.

A apolipoproteína A-I, que é sintetizada no fígado e no intestino, desempenha um papel importante no metabolismo das lipoproteínas de alta densidade (HDL) (BALLANTYNE e NAMBI, 2004). Esta proteína é a mais abundante HDL no plasma, compreendendo cerca de 70% da massa total das HDLs plasmáticas. Cinco a dez por cento da apolipoproteína A-I plasmática está em estado livre, sendo provavelmente sintetizada no fígado ou no intestino. A forma livre tem despertado bastante interesse, pois a ausência de lipídeos parece ser necessária para que a apolipoproteína A-I se ligue ao transportador ABCA1 (*ATP-binding cassette transporters A1*), chave para a manutenção dos níveis de HDL no plasma (DAVIDSON e THOMPSON, 2004). Dentre as HDLs, é a principal responsável pela proteção das células contra a toxicidade dos lipopolissacarídeos, se ligando a estes e neutralizando-os (MA *et al.*, 2004). A apolipoproteína A-I promove o efluxo do colesterol dos tecidos ao fígado para excreção e atua como co-fator da colesterolaciltransferase (LCAT) na formação dos ésteres colesteril. O gene *APOA1*, que codifica para a apolipoproteína A-1, está localizado no cromossomo 11 em 11q23-q24 (NCBI, 2010).

O componente amilóide-P sérico (SAP) é uma glicoproteína comum do plasma e, juntamente com a proteína C reativa (CRP), é uma proteína da família das pentraxinas (HUTCHINSON *et al.*, 2000). As pentraxinas são formadas por moléculas compostas por 5 -10 subunidades protéicas idênticas dispostas em simetria pentamérica (GEWURZ *et al.*, 1995). Esta família de proteínas é altamente conservada e seus membros são conhecidos por se ligarem a vários ligantes de forma cálcio dependente (KIERNAN *et al.*, 2004). Dentre as atividades observadas

para a SAP estão: a forte ligação com o DNA e cromatina em condições fisiológicas (*in vitro*); descondensação da cromatina pela retirada de histonas H1 e solubilização da mesma (*in vitro*); e ligação às células apoptóticas e à cromatina exposta pela morte celular (*in vivo*). Estes dados indicam a importância da SAP em reduzir a imunogenicidade da cromatina, provavelmente se ligando a ela e regulando a sua degradação (BICKERSTAFF, 1999). A SAP é codificada pelo gene *APCS*, localizado no cromossomo 1 (1q21-q23) e sugere-se que também exerça função de chaperona (NCBI, 2010).

A albumina sérica, que tem como função principal a regulação da pressão osmótica coloidal do sangue, é a principal proteína do plasma (DENG, *et al.*, 2006). É a principal transportadora de zinco no plasma, se ligando à cerca de 80% de todo o zinco plasmático (LU *et al.*, 2008). Atua na estabilização do volume de fluido extracelular e funciona também como carreador protéico de esteróides, ácidos graxos, hormônios da tireóide, íons (sódio, potássio e cálcio) (NCBI, 2010), água, e drogas. A albumina sérica é sintetizada no fígado como pré-albumina a qual tem um peptídeo N-terminal removido antes de ser liberado do retículo endoplasmático rugoso. O produto (pró-albumina) é então clivado nas vesículas de Golgi para originar a albumina secretada (van der VUSSE, 2009). O gene *ALB* é o codificador da albumina sérica e está localizado no cromossomo 4 em 4q11-q13. (NCBI, 2010). Todas as bandas de albumina sérica (FIGURA 12) foram identificadas com *score* bastante alto. A marcação representada pela letra **A** foi retirada de uma região característica de géis de tecido mamário não tumoral, já relatada na literatura (DENG *et al.*, 2006; LUO *et al.*, 2005) como albumina sérica, sendo sua identificação realizada apenas para confirmação. Sabendo-se que a albumina sérica possui massa teórica de 71,3 kDa e *pI* teórico de 5,92, a banda 230 apresentou um deslocamento muito grande de *pI* na direção ácida, mas possivelmente foi devido ao excesso de proteína que dificultou a separação na focalização isoeletrica, gerando um rastro. A banda 271 apresentou variação dentro do aceitável. A albumina sérica presente em mistura com a creatina quinase tipo-B na banda 364 extrapolou a variação de 20% aceitável para a massa (38,2%), mas, de acordo com o banco de dados *UniProtKB / SwissProt*, a albumina sérica apresenta uma isoforma gerada por *splicing* alternativo com massa igual a 47,36 kDa. A variação de massa para a banda 364, quando comparada a esta isoforma, foi de 7,1%, ficando dentro do aceitável.

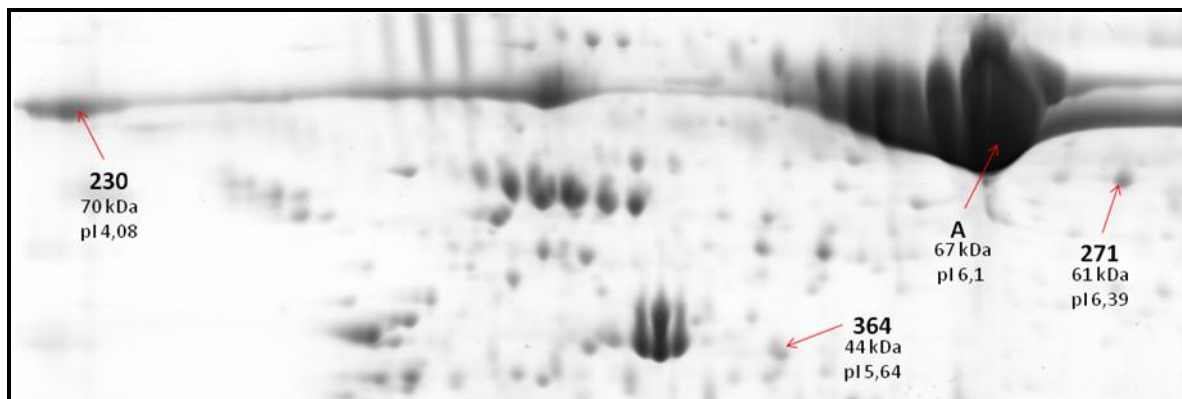


FIGURA 12 – DETALHE DAS BANDAS DA PROTEÍNA ALBUMINA SÉRICA

FONTE: O autor (2010)

LEGENDA: Os dados de pI e massa experimentais estão demonstrados para as bandas 230, 271 e 364, e também para a marcação representada pela letra **A** (retirada para confirmação).

#### 7.4.3. Enzimas Metabólicas

O grupo das **enzimas metabólicas (11%)** é composto por cinco bandas protéicas correspondendo a cinco enzimas: cadeia beta da L-lactato desidrogenase (banda 417), creatina quinase tipo-B (banda 364), glicerol-3-fosfato desidrogenase [NAD<sup>+</sup>] (bandas 415 e 417), a subunidade beta da ATP sintetase mitocondrial (banda 319), e a fosfoenolpiruvato carboxiquinase [GTP] citosólica (banda 491).

As isoenzimas lactato desidrogenase (LDH) são formadas por uma combinação aleatória entre duas subunidades codificadas por dois genes estruturalmente distintos, *LDHA* e *LDHB*. A expressão desses genes em mamíferos é regulada durante o desenvolvimento e é tecido-específica (MAEKAWA, *et al.*, 2003). Como muitas outras enzimas celulares, a LDH é liberada na corrente sanguínea por células danificadas (SHRINIVASAN *et al.*, 1999), portanto, alterações no padrão destas isoenzimas no soro serve como indicador patológico e de desenvolvimento do câncer (MAEKAWA, *et al.*, 2003). A cadeia beta da L-lactato desidrogenase é codificada pelo gene *LDBH*, localizado no cromossomo 12 em 12p12.2-p12.1, e é responsável pela catálise da conversão reversível de lactato e piruvato, e de NAD e NADH (NCBI, 2010).

A enzima creatina quinase tipo-B foi identificada em mistura com a albumina do soro na banda 364. O gene que codifica para esta proteína é o *CKB* e está localizado no cromossomo 14 em 14q32. A creatina quinase tipo B é uma enzima citoplasmática envolvida na homeostase energética através da catálise reversível da

transferência de fosfato entre o ATP e vários substratos, como a creatina fosfatase (NCBI, 2010).

O gene *GPD1*, localizado no cromossomo 12 em 12q12-q13, codifica a proteína glicerol-3-fosfato desidrogenase [NAD+] (NCBI, 2010). Esta enzima catalisa a conversão biológica de dihidroxiacetona em glicerol-3-fosfato e exibe expressão tecido-específica (OU *et al.*, 2005). A desidrogenase glicerol-3-fosfato [NAD+] foi identificada nas bandas 415 e 417, sendo que nesta última a identificação acusou uma mistura com a cadeia beta da L-lactato desidrogenase. As duas bandas apresentaram uma pequena diferença de pI entre elas .

Localizado no cromossomo 12 em 12q13.13, o gene *ATP5B* codifica para a subunidade beta da ATP sintase mitocondrial. A ATP sintase mitocondrial catalisa a síntese de ATP utilizando um gradiente eletro-químico de prótons através da membrana interna durante a fosforilação oxidativa (NCBI, 2010). Além da produção de ATP, estudos recentes sugerem que componentes da ATP sintase estão também presentes na superfície externa da membrana plasmática atuando como receptor de vários ligantes, e participando de processos biológicos como a regulação da proliferação e diferenciação em células endoteliais, e resposta imune em células tumorais (HUANG, *et al.*, 2008).

A fosfoenolpiruvato carboxiquinase [GTP] citosólica (banda 491) é codificada pelo gene *PCK1*, localizado em 20q13.31, e está envolvida na gliconeogênese. Catalisa a conversão de oxalato em fosfoenolpiruvato. É sintetizada principalmente pelo fígado, rins e adipócitos (NCBI, 2010). Esta proteína foi identificada como o *score* limite (56) e percentual de cobertura de seqüência de apenas 13% e mesmo com purificação peptídica não houve melhora. A variação entre a MM teórica e a experimental foi de 70%. Não são relatados casos de *splicing* alternativo no banco de dados *UniProtKB / SwissProt*. Essa identificação foi considerada duvidosa, sendo necessária nova análise para comprovação.

#### 7.4.4. Chaperonas Moleculares / Proteínas *Heat Shock*

O dobramento das proteínas na célula é auxiliada por enzimas catalíticas que aceleram o processo e por chaperonas que impedem o agregamento. O

desdobramento de uma proteína pode obviamente ocasionar a perda de função, mas o acúmulo de proteínas desdobradas em grandes e insolúveis agregados pode também interferir na função celular (WILKINSON e GILBERT, 2004). A classe das **chaperonas moleculares / proteínas *heat shock*** abrangeu **7%** das proteínas identificadas. Nesta categoria foram catalogadas três bandas, cada uma correspondendo a uma proteína diferente: endoplasmína (banda 164), proteína *heat shock* pequena beta-6 (banda 492), e proteína dissulfeto-isomerase (banda 268).

As proteínas *heat shock*, das quais as chaperonas moleculares representam um grupo importante, são classificadas de acordo com seu tamanho molecular em: Hsp100, Hsp90, Hsp70, Hsp60, Hsp40 e proteínas *heat shock* pequenas (*small heat shock proteins* – sHSPs); e sua expressão é induzida pelo estresse celular (MINAFRA *et al.*, 2006)

Localizado no cromossomo 19 (19q13.12), o gene *HSPB6* codifica para a proteína *heat shock* pequena beta-6, também conhecida como Hsp20 por ter massa molecular aparente de 20 kDa, pertence a uma família de 10 membros expressa em diferentes tecidos humanos (GUSEV *et al.*, 2004). As proteínas *heat shock* pequenas (sHSPs) funcionam como chaperonas moleculares prevenindo agregações, induzidas por estresse, de proteínas parcialmente desnaturadas e promovendo o retorno destas à sua conformação nativa quando em condições favoráveis. Similaridade de seqüência entre as sHSPs reside predominantemente no domínio  $\alpha$ -*crystallin*, uma região flanqueada por duas extensões. A extensão N-terminal é pouco conservada e influencia a atividade de chaperona. A extensão C-terminal estabiliza a estrutura quaternária e aumenta a solubilidade do complexo proteína / substrato. Como chaperonas moleculares, as sHSPs protegem a estrutura e atividade de proteínas, mas podem contribuir com o mal funcionamento destas quando alteradas (SUN e MacRAE, 2005). De maneira geral, as sHSPs protegem os tumores através de mecanismos independentes como, por exemplo, estresse oxidativo, proteção contra desnaturação de proteínas, atividade anti-apoptótica e por supressão do sistema imune (LAUDANSKI e WYCZECZOWSKA, 2006).

A Hsp20 é ubíqua e parece estar envolvida em várias funções incluindo a regulação da vasodilatação (contração de músculo liso), agregação de plaquetas (FAN *et al.*, 2005), transporte de glicose, proteção cardíaca e prevenção da apoptose. A fosforilação da Hsp20 está correlacionada com o relaxamento do

músculo liso (CHERNIK *et al.*, 2007). Rembold e colaboradores (2001), estudando a fosforilação da Hsp20 em carótidas de suínos, observaram que o valor do pI da proteína não fosforilada era próximo de 6,3. Além dessa forma, diferentes tecidos contêm formas de Hsp20 com pI 6,0, 5,9 e 5,7 (GUSEV *et al.*, 2005). Comparando estes dados é possível supor que as formas com pI 6,0 e 5,9 representam a Hsp20 fosforilada contendo sítios fosforilados em diferentes partes da molécula, enquanto a forma com pI 5,7 representa a proteína duplamente fosforilada (GUSEV *et al.*, 2005). O pI observado para esta proteína neste trabalho foi 6,35 aproximadamente, o que leva a crer que a proteína identificada não estava fosforilada. Acredita-se que muitas das funções atribuídas à Hsp20 possam ser explicadas pela associação com a proteína 14-3-3 (CHERNIK *et al.*, 2007), que será discutida posteriormente. Em 2007, Chernik e colaboradores, em estudo sobre a interação entre as proteínas Hsp20 e 14-3-3 gama, observaram que apenas a forma fosforilada da Hsp20 interage com a proteína 14-3-3 formando um complexo e aumentando sua atividade como chaperona. Sob condições fisiológicas a Hsp20 não aparece fosforilada e, portanto, não interage com a proteína 14-3-3.

Entre outras chaperonas e enzimas de dobramento, a endoplasmina e a proteína dissulfeto isomerase têm se mostrado ligantes às proteínas recém sintetizadas no retículo endoplasmático e costumam formar complexos multiprotéicos (MEUNIER *et al.*, 2002). A maioria das proteínas glicose-reguladas (onde são incluídas a endoplasmina e a proteína dissulfeto-isomerase) são chaperonas moleculares do retículo endoplasmático que auxiliam no dobramento protéico e são também proteínas  $\text{Ca}^{2+}$ -ligante. A chamada família de proteínas glicose-reguladas abrange uma variedade de chaperonas do retículo endoplasmático que podem ser induzidas em células cultivadas em meio deficiente em glicose. Condições de estresse geralmente levam ao acúmulo de proteínas não dobradas no retículo endoplasmático (LEE, 2001). A endoplasmina, ou proteína glicose-regulada de 94 kDa, pertence a família das Hsp90 e atua no processamento e transporte de proteínas secretadas (MINAFRA *et al.*, 2006) e é codificada pelo gene *HSP90B1*, localizado no cromossomo 12 (12q24.2-q24.3). A MM observada para esta proteína foi de aproximadamente 102 kDa enquanto a teórica é de 92,7 kDa. Mesmo sendo essa diferença aceitável, cabe dizer que ela provavelmente é ainda menor. O fato de o marcador de massa utilizado possuir 97 kDa como a maior

MM prejudica a estimativa da MM de proteínas acima desse valor. Observando a posição da proteína no gel (Figura 5, banda 164) percebe-se que ela apresenta MM muito próxima de 97 kDa.

Formação de pontes dissulfeto nativas é um processo complexo. As ligações devem ser não somente formadas (oxidação), mas também ligações incorretas devem ser desfeitas (redução) ou rearranjadas (isomerização) (ELLGAARD e RUDDOCK, 2005). A proteína dissulfeto isomerase (PDI) é essencial no retículo endoplasmático como catalítica do processo de dobramento e também como chaperona. Introduz ligações dissulfeto em proteínas (atividade de oxidase) e catalisa o rearranjo de ligações incorretas (atividade de isomerase) (WILKINSON e GILBERT, 2004). Segundo Janiszewski e colaboradores (2005), a PDI pode ainda estar envolvida na regulação da atividade da NADPH oxidase em células musculares lisas vasculares. Codificada pelo gene *P4HB*, localizado no cromossomo 17 em 17q25, a PDI é constitutivamente expressa na maioria dos tecidos e órgãos e é uma das proteínas mais abundantes no retículo endoplasmático. A sua alta concentração celular é devida tanto à sua alta expressão como à sua grande estabilidade (WILKINSON e GILBERT, 2004).

Tanto a PDI (banda 268) como a endoplasmina (banda 164) obtiveram percentual de cobertura de seqüência abaixo de 20%. O *score* obtido na identificação da PDI (103), a grande proximidade entre os valores teóricos e experimentais de massa e *pI* e o bom percentual de pareamento das massas obtidas por espectrometria (42%) foram considerados suficientes para aceitar a identificação. No caso da endoplasmina, o percentual de cobertura de seqüência abaixo de 20% pode ser devido ao tamanho da proteína (MM observada de 102 kDa). Considerando isto, o valor de *score* (72) satisfatório e a pequena variação entre a MM e o *pI* teóricos e experimentais, a identificação foi aceita.

#### 7.4.5. Crescimento Celular e Regulação da Proliferação

A categoria das proteínas envolvidas no **crescimento celular e regulação da proliferação (5%)** é composta por duas proteínas, a 14-3-3 beta / alfa e a 14-3-3 gama, cada uma correspondente a uma banda (460 e 453 respectivamente). O gene



*YWHAB* (cromossomo 20 em 20q13.1) codifica para a proteína 14-3-3 beta / alfa e o gene *YWHAG* (cromossomo 7 em 7q11.23) para a 14-3-3 gama.

As proteínas 14-3-3 exibem uma notável conservação de seqüências entre as espécies e compõem um grupo de proteínas multifuncionais que se ligam e modulam a função de várias proteínas celulares (FU *et al.*, 2000). Presente em praticamente todos os tecidos, este grupo de proteínas é codificado por sete genes em mamíferos ( $\alpha/\beta$ ,  $\epsilon$ ,  $\eta$ ,  $\gamma$ ,  $\tau/\theta$ ,  $\zeta/\delta$  e  $\sigma$ ). Mais de 200 alvos protéicos da 14-3-3 são descritas na literatura. Interagem predominantemente com proteínas fosforiladas nos resíduos Serina e Treonina, porém também são capazes de se ligar à proteínas não fosforiladas (JIN *et al.*, 2004). Através dessas interações a 14-3-3 controla a progressão do ciclo celular, regula a atividade de várias proteínas quinases, afeta a apoptose, coordena a sinalização celular e a dinâmica do citoesqueleto (CHERNIK *et al.*, 2007). O alto nível de conservação e natureza abundante desta proteína pode refletir sua fundamental importância na biologia das células eucarióticas (FU *et al.*, 2000).

#### 7.4.6. Detoxificação e Proteínas redox

Proteínas relacionadas à **detoxificação e proteínas redox (4%)** são enzimas que realizam funções importantes no catabolismo celular e proteção contra estresses metabólicos (MINAFRA *et al.*, 2006). Duas proteínas são representantes desta classe: glutathione S-transferase P (banda 479) e a peroxirredoxina 2 (banda 481).

As glutathione S-transferases (GSTs) representam o principal grupo de enzimas detoxificantes (HAYES e PULFORD, 1995). As GSTs, além de detoxificar xenobióticos eletrofílicos, por exemplo, químicos carcinogênicos e quimioterápicos, também inativam metabólitos secundários formados durante o estresse oxidativo. Estas enzimas também estão envolvidas na biossíntese de testosterona, prostaglandinas e progesterona, assim como na degradação de tirosina (HAYES *et al.*, 2004). Múltiplas isoenzimas GST, citosólicas e ligadas à membrana, estão presentes em todas as espécies de eucariotos, cada uma delas com propriedades catalíticas, assim como não catalíticas, distintas. As enzimas citosólicas são

codificadas por no mínimo cinco famílias de genes distantemente relacionadas (classe GST-alfa, mu, pi, sigma e teta) (HAYES e PULFORD, 1995). A glutathione S-transferase P é codificada pelo gene polimórfico *GSTP1* (classe GST-pi) localizado em 11q13 e que codifica proteínas ativas e funcionalmente diferentes, envolvidas no metabolismo de xenobióticos (NCBI, 2010).

As peroxiredoxinas constituem a família de peroxidases homodiméricas que reduzem peróxido de hidrogênio ( $H_2O_2$ ) à água e hidroperóxidos alquil à álcool (LOW *et al.*, 2007). Recentemente, o sistema de peroxiredoxinas tem recebido muita atenção devido à sua alta eficiência como antioxidante. Alguns membros desta família também afetam a diferenciação celular e apoptose, sendo encontrados super expressos no câncer de mama (MINAFRA *et al.*, 2006). A peroxiredoxina 2 é a terceira proteína mais abundante em eritrócitos humanos (SCHRODER *et al.*, 2000) e uma das peroxidoredoxinas citoplasmáticas predominantes nas células de mamíferos (PESKIN *et al.*, 2007). Esta proteína pode ter um efeito proliferativo, desempenhando um papel no desenvolvimento e progressão do câncer, e é codificada pelo gene *PRDX2* localizado em 19p13.2. Transcritos variantes que codificam isoformas distintas da proteína já foram observados para esse gene (NCBI, 2010).

#### 7.4.7. Proteínas Associadas à Membrana com Múltiplas Atividades

As anexinas I e V, identificadas a partir das bandas 422 e 433 respectivamente, representam a classe das **proteínas associadas à membrana com múltiplas atividades (5%)**. A família das anexinas é composta por proteínas membrana-ligantes que compartilham propriedades estruturais e atividades biológicas associadas com processos relacionados à membrana (GENDEREN *et al.*, 2008), atuando como proteínas de dobramento (scaffolding) ou de transporte transmembrana (bridging) (MINAFRA *et al.*, 2006). Por definição, uma proteína anexina deve ser capaz de se ligar à fosfolípidos negativamente carregados na presença de  $Ca^{2+}$  (GERKE e MOSS, 2002). A seqüência N-terminal de cada membro desta família é única, com variação de tamanho e aminoácidos, e acredita-se que a função específica e atividade biológica das anexinas sejam associadas a esta

seqüência (LIM e PERVAIZ, 2007). Estas proteínas estão envolvidas em interações com o citoesqueleto, inibição de fosfolipase, sinalização intracelular e anticoagulação (MINAFRA *et al.*, 2006).

As anexinas I e V (assim como a II e IV) têm sido encontradas na superfície celular e na circulação, indicando que são liberadas no meio extracelular (GENDEREN *et al.*, 2008).

Codificada pelo gene *ANXA1* localizado no cromossomo 9 (9q12-q21.20), a anexina I, descrita no final da década de 1970, foi o primeiro membro caracterizado desta família (BLACKWELL *et al.*, 1980). Entre as diversas funções às quais é atualmente associada estão: atividade antiinflamatória (inibe a transmigração de leucócitos, bloqueando a resposta inflamatória); participação na regulação da proliferação celular, diferenciação e apoptose (LIM e PERVAIZ, 2007); participação, através do complexo anexina I – profilina na regulação do citoesqueleto associado à membrana e; participação em vias de biossíntese (GERKE e MOSS, 2002).

Codificada pelo gene *ANXA5* localizado no cromossomo 4 (4q26-q28; 4q28-q32), a anexina V extracelular tem função anticoagulante (inibe a formação do complexo protrombinase), e também atua na apoptose, desempenhando seu papel na superfície celular. No geral, a anexina V pode desempenhar várias funções na apoptose, na inflamação e na homeostase e trombose (GENDEREN *et al.*, 2008).

#### 7.4.8. Degradação Protéica

A proteína *tripartite motif-containing protein 13* (banda 262) é a representante da classe de proteínas associadas à **degradação protéica (2%)**. Esta proteína é codificada pelo gene *TRIM13*, pertencente a família TRIM, que está localizado no cromossomo 13 em 13q14 (NCBI, 2010). A presença do domínio RING (presente nas ubiquitinas E3 ligase) sugere a função das proteínas desta família no processo de ubiquitinação de proteínas danificadas ou em excesso. Essa sinalização faz com que as proteínas sejam encaminhadas para os proteossomos (MERONI e ROUX, 2005). Genes pertencentes a esta família estão associados a diversos processos, como desenvolvimento e crescimento celular, e envolvido também em diversas doenças humanas (REYMOND *et al.*, 2001). Esta identificação não obteve melhora

com a purificação peptídica e foi considerada duvidosa por apresentar *score* 58 e cobertura de seqüência de apenas 15%. Houve também variação na massa acima de 20%. Apesar de três isoformas originadas de *splicing* alternativo do gene *TRIM13* terem sido relatadas, nenhuma tem massa mais próxima da observada.

#### 7.4.9. Proteínas com Outras Funções

A classe das **proteínas com outras funções (9%)** é composta por onze bandas correspondentes a quatro proteínas e suas isoformas: cadeia alfa-1 do colágeno VI (bandas 139, 140, 142, 144, 145 e 146), glicoproteína alfa-1B (bandas 185, 189 e 197), a ATPase transicional do retículo endoplasmático (banda 319) e a ubiquitina E3 ligase RAD 18 (banda 375).

As proteínas da superfamília dos colágenos são responsáveis pela manutenção da integridade de vários tecidos. O colágeno VI é uma glicoproteína grande composta por três cadeias diferentes, alfa-1 (VI), alfa-2 (VI) e alfa-3 (VI) (BALDOCK *et al.*, 2003), sendo a cadeia alfa-1 (VI) codificada pelo gene *COL6A1* localizado em 21q22.3 (NCBI, 2010). A principal propriedade atribuída ao colágeno VI é a de formar uma rede microfibrilar, que funciona como substrato para a adesão celular, conectando as fibras de colágeno, nervos, células musculares e vasos sanguíneos à matriz que os envolve (KUO, *et al.*, 1997; LAMPE e BUSHBY, 2005). Os adipócitos, um dos tipos celulares predominantes no estroma do tecido mamário, representam a mais abundante fonte de colágeno VI, mesmo este sendo expresso por vários outros tipos celulares (IYENGAR *et al.*, 2005). A expressão aumentada de colágeno VI estromal está correlacionada com a tumorigênese e progressão maligna (IYENGAR *et al.*, 2005). As seis isoformas identificadas da cadeia alfa-1 do colágeno VI (FIGURA 13) apresentam um deslocamento de pI principalmente no sentido ácido de 5,29 até 5,1, considerando que o pI teórico para esta proteína é de 5,26. Foi possível observar outras marcações seqüenciais às bandas identificadas, e que provavelmente sejam outras isoformas da cadeia alfa-1 do colágeno VI, mas que não foram detectadas pelo programa de análise em mais de duas amostras. A variação das massas experimentais de todas as isoformas em relação a MM teórica (109 kDa) foi em média de 15,2%, mas pode ter sido acentuada pelo fato de não

haver um marcador de massa acima de 97 kDa. Isso prejudica na estimativa da massa de proteínas acima desse valor.

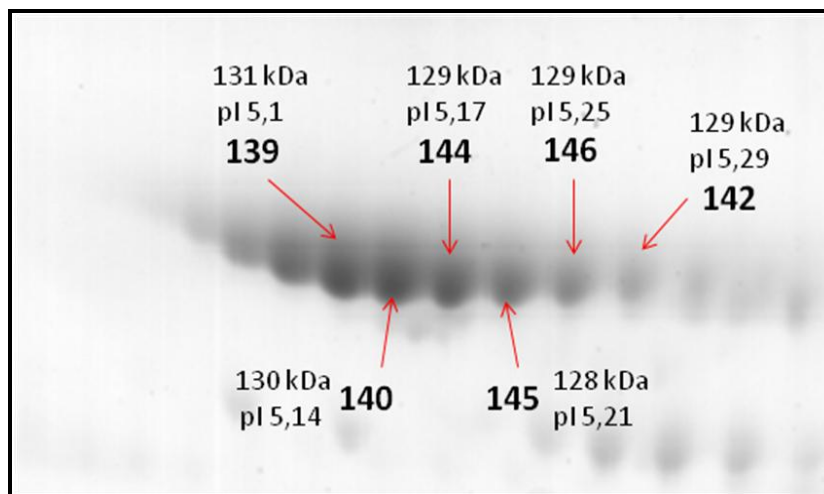


FIGURA 13 – DETALHE DAS BANDAS DA CADEIA ALFA-1 DO COLÁGENO VI

FONTE: O autor (2010)

LEGENDA: Os dados de pI e massa experimentais estão demonstrados para as bandas 139, 140, 142, 144, 145 e 146.

Enquanto os polipeptídeos na conformação correta deixam o retículo endoplasmático através da via de secreção, as proteínas danificadas, desdobradas ou desagregadas são marcadas (ubiquitinadas) e retidas. Estas proteínas são degradadas, necessitando serem transportadas novamente para o citosol por um processo chamado retro-translocação. A ATPase transicional do retículo endoplasmático (TERA) é uma das proteínas que compõem a maquinaria de retro-translocação melhor estabelecida (YE *et al.*, 2004). Esta proteína pertence à família AAA (*ATPases Associated with different cellular Activities*), que está associada a vários processos celulares incluindo fusão de membranas, biogênese de organelas, degradação protéica e regulação do ciclo celular (MEYER *et al.*, 2000). A proteína TERA é codificada pelo gene *VCP* localizado no cromossomo 9 em 9p13.3 (NCBI, 2010).

A proteína codificada pelo gene *A1BG*, localizado no cromossomo 19 em 19q13.4, é uma glicoproteína plasmática de função desconhecida, chamada de glicoproteína alfa-1B (NCBI, 2010). Esta proteína possui similiaridade de seqüência com várias regiões de alguns membros da família das imunoglobulinas. Dados obtidos por Stehle Jr. e colaboradores (2007), através de estudos com

camundongos, sugerem que a expressão da glicoproteína alfa-1B muda dependendo do sexo e da idade. As 3 bandas identificadas como glicoproteína alfa-1B (Figura 9) apresentaram variação de massa acima de 30% em relação ao esperado teórico (54,8 kDa), contudo, o *score* e o percentual de cobertura das bandas 189 e 197 foram muito bons (Tabela 6). As variações de massa e pI podem ser devido à modificações pós-traducionais.

A ubiquitina E3 ligase RAD 18, codificada pelo gene *RAD 18* (3p25-p24) está envolvida no reparo pós-replicação do DNA danificado por UV completando as lacunas da fita não-molde (NCBI, 2010). A via ubiquitina-protossomo é a maior para a degradação de proteínas e envolve múltiplas etapas de reações catalizadas por uma cascata enzimática, incluindo as ubiquitinas E3 ligase (SUN, 2006). Estas enzimas, através do domínio RING se ligam a proteínas específicas do substrato e promovem a transferência de ubiquitina (HERSHKO e CIECHANOVER, 1998). Existem pelo menos 1000 ubiquitinas E3 ligase no genoma humano (SUN, 2006). Com *score* 57, 19% de cobertura de seqüência e dados de MM e pI com grande variação (24,5% e 2,34), esta identificação da ubiquitina E3 ligase RAD 18 (banda 375) foi considerada duvidosa. Nenhum relato de *splicing* alternativo é relatado para este gene. Possui dez sítios de fosforilação, mas de acordo com simulação realizada (*Phosphosite Plus* – <http://www.phosphosite.org>) mesmo com todos estes resíduos fosforilados a variação de pI fica em uma unidade de pH.

## 8. CONCLUSÃO

O tecido mamário é bastante complexo devido à influência de diversos fatores ao longo da vida, sendo difícil estabelecer um perfil proteômico. Neste trabalho foram identificadas diversas proteínas já descritas na literatura, além de prováveis isoformas que se constituem em importante base para futuras pesquisas buscando evidenciar as modificações pós traducionais responsáveis pela variação observada nos géis.

Diversas proteínas identificadas neste trabalho são apontadas na literatura como diferencialmente expressas em relação a tumores mamários, tornando nossos dados importantes tanto no auxílio à caracterização do tecido não tumoral como para estudos subseqüentes sobre expressão diferencial com o tecido neoplásico.

A implantação da metodologia proteômica no Laboratório de Citogenética Humana e Oncogenética adiciona mais uma ferramenta importante às pesquisas sobre o câncer de mama, permitindo uma abordagem integrada entre estudos que já vem sendo realizados no laboratório (citogenéticos, gênicos e do transcriptoma). Este será um dos trabalhos que servirão como base para o desenvolvimento de novas frentes de pesquisa avaliando diretamente os produtos protéicos.

A presença das proteínas identificadas neste trabalho é corroborada pelos dados da literatura como presentes no tecido mamário não tumoral, sejam secretadas por ele ou presentes no micro-ambiente. A extensão da análise, avaliando todas as proteínas expressas, é necessária para a obtenção do mapa proteômico, etapa essencial na caracterização do tecido mamário.

## REFERÊNCIAS

AAPC: Associação de Assistência e Prevenção ao Câncer. Disponível em <[http://www.aapcan.org.br/sobre\\_cancer.php](http://www.aapcan.org.br/sobre_cancer.php)> Acesso em: março de 2010.

ABE, K. Therapeutic potential of neurotrophic factors and neural stem cells against ischemic brain injury. **J. Cereb. Blood Flow Metab.**, v. 20, p. 1393-1408, 2000.

ALAOUI-JAMALI, M. A.; XU, Y. Proteomic technology for biomarker profiling in cancer: an update. **Journal of Zhejiang University SCIENCE B.** v. 7, n. 6, p. 411-420, 2006.

ALBERTS, B.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. **Biologia Molecular da Célula**, 4. ed. São Paulo: Artmed, São Paulo, 2004.

ALLRED, D. C.; BROWN, P.; MEDINA, D. The origins of estrogen receptor alphapositive and estrogen receptor alpha-negative human breast cancer. **Breast Cancer Res.**, v. 6, p. 240–245, 2004.

ALTAMURA, C.; SQUITTI, R.; PASQUALETTI, P.; GAUDINO, C.; PALAZZO, P.; TIBUZZI, F.; LUPOI, D.; CORTESI, M.; ROSSINI, M.; VERNIERI, F. Ceruloplasmin/Transferrin system is related to clinical status in acute stroke. **Stroke**, v. 40, p. 1282-1288, 2009.

AUNG, T. T.; CHANDANA, R. S.; D'SILVA, K. J.; DIMITROV, N. V. The role of vitamin D in breast cancer. **Oncol. Rev.**, v. 3, p.19–25, 2009.

AZAD, N. F.; RASOOL, N.; ANNUNZIATA, C. M.; MINASIAN, L.; WHITELEY, G.; KOHN, E. C. Proteomics in clinical trials and practice – present uses and future promise. **Molecular & Cellular Proteomics.** v. 5, p. 1819-1829, 2006.

BALDOCK, C.; SHERRATT, M. J.; SHUTTLEWORTH, C. A.; KIELTY, C. M. The supramolecular organization of collagen VI microfibrils. **J. Mol. Biol.**, n. 330, p. 297–307, 2003.

BALLANTYNE, C. M.; NAMBI, V. Apolipoprotein A-I and high-density lipoprotein: is this the beginning of the era of noninvasive angioplasty? **Journal of the American College of Cardiology**, v. 44, n. 7, p. 1436-1438, 2004.



BALMAIN, A.; GRAY, J.; PONDER, B. The genetics and genomics of cancer. **Nature Genetics Supplement**, v. 33, p. 238-247, 2003.

BÁNKFALVI, A.; LUDWIG, A.; HESSELLE, B.; BUERGER, H.; BUCHWALOW, I. B.; BOECKER, W. Different proliferative activity of the glandular and myoepithelial lineages in benign proliferative and early malignant breast diseases. **Mod Pathol.**, v. 17, p. 1051-1061, 2004.

BARTSCH, H. Studies on biomarkers in cancer etiology and prevention: a summary and challenge of 20 years of interdisciplinary research. **Mutat. Res.** v. 462, p. 255-279, 2000.

BERTUCCI, F.; BIRNBAUM, D.; GONÇALVES, A. Proteomics of breast cancer: principles and potential clinical applications. **Molecular & Cellular Proteomics.** v. 5, p. 1772-1786, 2006.

BICKERSTAFF, M. C. M.; BOTTO, M.; HUTCHINSON, W. L.; HERBERT, J.; TENNENT, G. A.; BYBEE, A.; MITCHELL, D. A.; COOK, H. T.; BUTLER, P. J. G.; WALPORT, M. J.; PEPYS, M. B. Serum amyloid P component controls chromatin degradation and prevents antinuclear autoimmunity. **Nature Medicine**, v. 5, n 6, 1999.

BISCH, P. M. Proteômica. In: LUÍS MIR. (Org.). **Genômica**. São Paulo: Editora Atheneu, 2004, p. 139-162.

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

BUDEL, V. M. Anatomia cirúrgica da mama. In: RIETJENS, M.; URBAN, C. A. **Cirurgia da Mama Estética e Reconstructora**. Rio de Janeiro: Revinter, 2007. p. 360-363.

CALVO, K. R.; LIOTTA, L. A.; PETRICOIN, E. F. Clinical Proteomics: from biomarker Discovery and cell signaling to individualized personal therapy. **Bioscience Reports**, v. 25, n. 1/2, p. 107-125, 2005.

CAÑAS, B.; PIÑEIRO, C.; CALVO, E.; LÓPEZ-FERRER, D.; GALLARDO, J. M. Trends in sample preparation for classical and second generation proteomics. **Journal of Chromatography A**. v. 1153, p. 235–258, 2007.

CEBRIAN, A.; PHAROAH, P. D.; AHMED, S.; ROPERO, S.; FRAGA, M. F.; SMITH, P. L.; CONROY, D.; LUBEN, R.; PERKINS, B.; EASTON, D. F.; DUNNING, A. M.; ESTELLER, M.; PONDER, B. A. J. Genetic variants in epigenetic genes and breast cancer risk. **Carcinogenesis**, v. 27, n. 8, p. 1661-1669, 2006.

CHEN, M. H. S.; YIP, G. W. C.; TSE, G. M. K.; MORIYA, T.; LUI, P. C. W.; ZIN, M. L.; BAY, B. H.; TAN, P. H. Expression of basal keratins and vimentin in breast cancers of young women correlates with adverse pathologic parameters. **Modern Pathology**, v. 21, p. 1183-1191, 2008.

CHERNIK, I. S.; SEIT-NEBI, A. S.; MARSTON, S. B.; GUSEV, N. B. Small heat shock protein Hsp20 (HspB6) as a partner of 14-3-3 $\gamma$ . **Molecular and Cellular Biochemistry**, n. 295, p. 9-17, 2007.

CLARKE, M.F.; FULLER, M. Stem cells and cancer: two faces of eve. **Cell**, v. 124, n. 6, p. 1111-1115, 2006.

COLLINS, J. A.; BLAKE, J. M.; CROSIGNANI, P. G. Breast cancer risk with postmenopausal hormonal treatment. **Human Reprod. Update**, v. 11, n. 6, p. 545-560, 2005.

COURTILLOT, C.; PLU-BUREAU, G.; BINART, N.; BALLEYGUIER, C.; SIGAL-ZAFRANI, B.; GOFFIN, V.; KUTTENN, F.; KELLY, P. A.; TOURAINE, P. Benign breast diseases. **J. Mammary Gland. Biol. Neoplasia**, v. 10, p. 325-335, 2005.

CRESPI, B.; SUMMERS, K. Evolutionary biology of cancer. **Trends in Ecology and Evolution**, v. 20, n. 10, p. 545-552, 2005.

DAVIDSON, W. S.; THOMPSON, T. B. The structure of apolipoprotein A-I in high density lipoproteins. **Journal of Biological Chemistry**, v. 282, n. 31, p. 22249-22253, 2007.

DAVIS, M. A.; HANASH, S. High-throughput genomic technology in research and clinical management of breast cancer: Plasma-based proteomics in early detection and therapy. **Breast Cancer Research**, v. 8, p. 217, 2006.

DENG, S. S.; XING, T. Y.; ZHOU, G. Y.; XIONG, R. H.; LU, Y. G.; WEN, B.; LIU, S. Q.; YANG, H. J. Comparative proteome analysis of breast cancer and adjacent normal breast tissues in human. **Genomics, Proteomics & Bioinformatics**, v. 4, n. 3, p. 165-172, 2006.

DOBRYSYCKA, W. Biological functions of haptoglobin--new pieces to an old puzzle (ABSTRACT). **Eur. J. Clin. Chem. Clin. Biochem.**, v. 35, n 9, p 647-654, 1997.

DUFLOTH, R. M.; CARVALHO, S.; HEINRICH, J. K.; SHINZATO, J. Y.; SANTOS, C. C.; ZEFERINO, L. C.; SCHMITT, F. Analysis of *BRCA1* and *BRCA2* mutations in Brazilian breast cancer patients with positive family history. **Sao Paulo Medical Journal**, v. 123, p. 192-197, 2005.

DONTU, G.; AL-HAJJ, M.; ABDALLAH, W. M.; CLARKE, M. F.; WICHA, M. S. Stem cells in normal breast development and breast cancer. **Cell proliferation**, v. 36 (Suppl. 1), p. 59-72, 2003.

DUPONT, A.; CORSEAUX, D.; DEKEYZER, O.; DROBECQ, H.; GUIHOT, A. L.; SUSEN, S.; VINCENTELLI, A.; AMOUYEL, P.; JUDE, B.; PINET, F. The proteome and secretome of human arterial smooth muscle cells. **Proteomics Journal**, v. 5, p. 585-596, 2005.

ELLGAARD, I.; RUDDOCK, L. W. The human protein disulphide isomerase family: substrate interactions and functional properties. **European Molecular Biology Organization Reports**, v. 6, n. 1, p. 28-32, 2005.

FAN, G. C.; REN, X.; QIAN, J.; YUAN, Q.; NICOLAOU, P.; WANG, Y.; JONES, K.; CHU, G.; KRANIAS, E. G. Novel cardioprotective role of a small heat-shock protein, hsp20, against ischemia/reperfusion injury. **Circulation**, v. 111, p. 1792-1799, 2005.

FU, H.; SUBRAMANIAN, R. R.; MASTERS, S. C. 14-3-3 Proteins: Structure, Function, and Regulation. **Annu. Rev. Pharmacol. Toxicol.**, n. 40, p. 617-647, 2000.

GARNIS, C.; BUYS, T. P.; LAM, W. L. Genetic alteration and gene expression modulation during cancer progression. **Molecular Cancer**. v. 3, 9, 2004.

GE HEALTHCARE Handbook. **2D – Electrophoresis: Principles and Methods**. Suécia: Editora Elanders Tofters, 2005. 162 p.

GERWUZ, H.; ZHANG, X. H.; LINT, T. F. Structure and function of the pentraxins. **Current Opinion in Immunology**, v. 7, p. 54-64, 1995.

GILLEY, D.; TANAKA, H.; HERBERT, B. S. Telomere dysfunction in aging and cancer. **The International Journal of Biochemistry & Cell Biology**. v. 37, p. 1000-1013, 2005.

GOOME, P. T.; BERTOLINI, J. Therapeutic potential of vitamin D-binding protein. **TRENDS in Biotechnology**, v.22, n. 7, p. 340-345, 2004.

GOOME, P. T.; McCANN, K. B. Transferrin: structure, function and potential therapeutic actions. **Drug Discovery Today**, v. 10, n. 4, p. 267-273, 2005.

GRIFFITHS, A. J. F.; MILLER, J. H.; SUZUKI, D. T.; LEWONTIN, R. C.; GELBART, W. M. **Introdução à genética**. 7. ed. New York: Editora Guanabara Koogan S. A., 2002.

GUO, Y.; FU, Z.; EYK, J. E. V. A proteomic primer for the clinician. **Proceedings of the American Thoracic Society**. v. 4, p. 9-17, 2007.

GUSEV, N. B.; BUKACH, O. V.; MARSTON, S. B. Structure, Properties, and Probable Physiological Role of Small Heat Shock Protein with Molecular Mass 20 kD (Hsp20, HspB6). **Biochemistry (Moscow)**, v. 70, n. 6, p. 629-637, 2004.

HEASMAN, S. J.; RIDLEY, A. J. Mammalian Rho GTPases: new insights into their functions from *in vivo* studies. **Nature Reviews. Molecular Cell Biology**, v. 9(9), p. 690-701, 2008.

HAYES, J. D.; FLANAGAN, J. U.; JOWSEY, I. R. Glutathione transferases. **Annu. Rev. Pharmacol. Toxicol.**, v. 45, p. 51-88, 2005.

HAYES, J. D.; PULFORD, D. J. The glutathione S-transferase supergene family: regulation of GST and the contribution of the isoenzymes to cancer chemoprotection and drug resistance (ABSTRACT). **Crit. Rev. Biochem. Mol. Biol.**, v. 30, n. 6, p. 445-600, 1995.

HENNIGHAUSEN, L.; ROBINSON, G. W. Signaling pathways in mammary gland development. **Developmental Cell**, v. 1, p. 467-475, 2001.

HERSHKO, A.; CIECHANOVER, A. The ubiquitin system. **Annu. Rev. Biochem**, v. 67, p. 425-479, 1998.

HONDERMARCK, H.; DOLLÉ, L.; YAZIDI-BELKOURA, I. E.; VERCOUTTER-EDOUART, A. S.; ADRIAENSSENS, E.; LEMOINE, J. Functional proteomics of breast cancer for signal pathway profiling and target discovery. **Journal of Mammary Gland Biology and Neoplasia**, v. 7, n. 4, p. 395-405, 2002.

HONDERMARCK, H. Breast cancer – when proteomics challenges biological complexity. **Molecular & Cellular Proteomics**. v. 2, p. 281-291, 2003.

HONDERMARCK, H.; TASTET, C.; YAZIDI-BELKOURA, I. E.; TOILLON, R. A.; LE BOURHIS, X. Proteomics of breast cancer: the quest for markers and therapeutic targets. **Journal of Proteome Research**. v. 7, n. 4, p. 1403-1411, 2008.

HUGHES, L.E.; MANSEL, R.E.; WEBSTER, D.J.T. Benign disorders and diseases of the breast—concepts and clinical management. **W.B. Saunders**, v. 2, p. 7–20, 2000.

HUTCHINSON, W. L.; HOHENESTER, E.; PEPYS, M. B. Human serum amyloid P component is a single uncomplexed pentamer in whole serum. **Molecular Medicine**, v. 6, n. 6, p. 482–493, 2000.

IARC: International Agency for Research on Cancer. Disponível em <<http://www.iarc.fr>> Acesso em: março de 2010.

IBCC: INSTITUTO BRASILEIRO DE CONTROLE DO CÂNCER. Disponível em <<http://www.ibcc.org.br>> Acesso em: março de 2010.

INCA/MS. **Falando sobre Câncer de Mama**, 66p. Rio de Janeiro, 2000.

INCA/MS: Instituto Nacional de Câncer do Ministério da Saúde. Estimativa 2010 – Incidência de câncer no Brasil. Rio de Janeiro. Disponível em <<http://www.inca.gov.br>> Acesso em: março de 2010.

INGENBLEEK, Y.; YOUNG, V. Transthyretin (prealbumin) in health and disease: Nutritional implications. **Annu. Rev. Nutr.**, v. 14, p. 495-533, 1994.

IYENGAR, P.; ESPINA, V.; WILLIAMS T. W.; LIN, Y.; BERRY, D.; JELICKS, L. A.; LEE, H.; TEMPLE, K.; GRAVES, R.; POLLARD, J.; CHOPRA, N.; RUSSELL, R. G.; SASISEKHARAN, R.; TROCK, B. J.; LIPPMAN, M.; CALVERT, V. S.; PETRICOIN III, E. F.; LIOTTA, L.; DADACHOVA, E.; PESTELL, R. G.; LISANTI, M. P.; BONALDO, P.; SCHERER, P. E. Adipocyte-derived collagen VI affects early mammary tumor progression in vivo, demonstrating a critical interaction in the tumor/stroma microenvironment. **Journal of Clinical Investigation**, v. 115, n. 5, p. 1163–1176, 2005.

JANISZEWSKI, M.; LOPES, R. L.; CARMO, A. O.; PEDRO, M. A.; BRANDES, R. P.; SANTOS, C. X. C.; LAURINDO, F. R. M. Regulation of NAD(P)H oxidase by associated protein disulfide isomerase in vascular smooth muscle cells. **Journal of Biological Chemistry**, v. 280 (49), p. 40813-40819, 2005.

JAEGER, J. L.; SHIMIZU, N.; GITLIN, J. D. Tissue-specific ceruloplasmin gene expression in the mammary gland. **Biochem. Journal**, n. 280, p. 671-677, 1991.

JIN, J.; SMITH, F. D.; STARK, C.; WELLS, C. D.; FAWCETT, J. P.; KULKARNI, S.; METALNIKOV, P.; O'DONNELL, P.; TAYLOR, P.; TAYLOR, L.; ZOUGMAN, A.; WOODGETT, J. R.; LANGE BERG, L. K.; SCOTT, J. D.; PAWSON, T. Proteomic, functional, and domain-based analysis of in vivo 14-3-3 binding proteins involved in cytoskeletal regulation and cellular organization. **Current Biology**, v. 14, p. 1436-1450, 2004.

JOHNSON, K.C. Accumulating evidence on passive and active smoking and breast cancer risk. **International Journal of Cancer**, v. 117, p. 619-628, 2005.

JOHNSON-THOMPSON, M. C.; GUTHRIE J. Ongoing research to identify environmental risk factors in breast carcinoma. **Cancer**, v. 88, p. 9-1224, 2000.

KIERNAN, U. A.; NEDELKOV, D.; TUBBS, K. A.; NIEDERKOFER, E. E.; NELSON, R. W. Proteomic characterization of novel serum amyloid P component variants from human plasma and urine. **Proteomics**, v. 4, p. 1825–1829, 2004.

KIM, H.; COPE, M. B.; HERRING, R.; ROBINSON, G.; WILSON, L.; PAGE, G. P.; BARNES, S. 2D difference gel electrophoresis of prepubertal and pubertal rat mammary gland proteomes. **Journal of Proteome Research**, p. 4638–4650, n, 7, 2008.

KUMAR, Y.; KHACHANE, A.; BELWAL, M.; DAS, S.; SOMSUNDARAM, K.; TATU, U. ProteoMod: A new tool to quantitate protein post-translational modifications. **Proteomics**, v. 4, p. 1672–1683, 2004.

KUO, H. J.; MASLEN, C. L.; KEENE, D. R.; GLANVILLE, R. W. Type VI collagen anchors endothelial basement membranes by interacting with type IV collagen. **Journal of Biological Chemistry**, v. 272, n. 42, p. 26522-26529, 1997.

LAMPE, A. K.; BUSHBY, K. M. D. Collagen VI related muscle disorders. **J. Med. Genet.**, v. 42, p. 673-685, 2005.

LASH, T.L.; ASCHENGRAU, A. Alcohol drinking and risk of breast cancer. **The Breast Journal**, v. 6, n. 6, p. 396-399, 2000.

LEE, A. S. The glucose-regulated proteins: stress induction and clinical applications. **TRENDS in Biochemical Sciences**, v. 16, n. 8, p. 504-510, 2001.

LIS, H.; SHARON, N. Protein glycosylation: Structural and functional aspects. **Eur. J. Biochem.**, v. 218, p. 1-27, 1993.

LOW, F. M.; HAMPTON, M. B.; PESKIN, A. V.; WINTERBOURN, C. C. Peroxiredoxin 2 functions as a noncatalytic scavenger of low-level hydrogen peroxide in the erythrocyte. **Blood**, v. 109, n. 6, p. 2611-2617, 2007.

LÖNNERDAL, B. Bioactive proteins in human milk: mechanisms of action. **Journal of Pediatrics**, v. 156, n. 2, p. S26-30, 2010.

LÖSTER, K.; KANNICHT, C. 2-Dimensional Electrophoresis: Detection of glycosylation and influence on spot pattern. In: **Post-translational Modifications of Proteins: Tools for Functional Proteomics: Methods in Molecular Biology**, v. 446, p. 199-214. 2. ed., Totowa: Human Press, 2008.

MA, J.; LIAO, X. L.; LOU, B.; WU, M. P. Role of apolipoprotein A-I in protecting against endotoxin toxicity. **Acta Biochimica et Biophysica Sinica**, v. 36, n. 6, p. 419-424, 2004.

MACEDO, M. F.; de SOUSA, M. Transferrin and the transferrin receptor: of magic bullets and other concerns (ABSTRACT). **Inflammation & Allergy Drug Targets**, v. 7, n. 1, p. 41-52, 2008.

MAEKAWA, M.; TANIGUCHI, T.; ISHIKAWA, J.; SUGIMURA, H.; SUGANO, K.; KANNO, T. Promoter hypermethylation in cancer silences *LDHB*, eliminating lactate dehydrogenase isoenzymes 1-4. **Clinical Chemistry**, v. 49, n. 9, p. 1518-1520, 2003.

MAREEL, M.; LEROY, A. Clinical, cellular, and molecular aspects of cancer invasion. **Physiol. Rev.**, v. 83, p. 337-376, 2003.

MARTIN, D. B.; NELSON, P. S.; From genomics to proteomics: techniques and applications in cancer research. **Trends in Cell Biology**. v. 11, n. 11, p. 60-64, 2001.

MAYR, U.; MAYR, M.; YIN, X.; BEGUM, S.; TARELLI, E.; WAIT, R.; XU, Q. Proteomic dataset of mouse aortic smooth muscle cells. **Proteomics**, v. 5, p. 4546-4557, 2005.

McKEAN, P. G.; VAUGHAN, S.; GULL, K. The extended tubulin superfamily. **Journal of Cell Science**, v. 114, p. 1723-1733, 2001.

MERONI, G.; ROUX, G. D. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. **BioEssays**, v. 27, p. 1147-1157, 2005.

MEYER, H. H.; SHORTER, J. G.; SEEMANN, J.; PAPPIN, D.; WARREN, G. A complex of mammalian ufd1 and npl4 links the AAA-ATPase, p97, to ubiquitin and nuclear transport pathways. **European Molecular Biology Organization Journal**, v. 19, n. 10, p. 2181-2192, 2000.

MILANESE, T. R.; HARTMANN, L. C.; SELLERS, T. A.; FROST, M. H.; VIERKANT, R. A.; MALONEY, S. D.; PANKRATZ, V. S.; DEGNIM, M. C.; VACHON, C. N.; REYNOLDS, C. A.; THOMPSON, R. A.; MELTON III, L. J.; GOODE, E. L.; VISSCHER, D. W. Age-Related Lobular Involution and Risk of Breast Cancer. **J. Nat. Cancer Inst.**, v. 98, n. 22, p. 1600-1607, 2006.

MINAFRA, I. P.; CANCEMI, P.; FONTANA, S.; MINAFRA, L.; FEO, S.; BECCHI, M.; FREYRIA, A. M.; MINAFRA, S. Expanding the protein catalogue in the proteome reference map of human breast cancer cells. **Proteomics**, v. 6, p. 2609-2625, 2006.

MORCH, L. S.; STAHLBERG, C.; LOKKEGAARD, E.; HUNDRUP, Y. A.; OBEL, E. B.; GRONBAEK, M. Amount of alcohol, drinking pattern and breast cancer among Danish nurses. **American Journal of Epidemiology**, v. 161, p. 54-54, 2005.

MOSESSON, M. W. Fibrinogen and fibrin structure and functions. **Journal of Thrombosis and Haemostasis**, v. 3, p. 1894-1904, 2005.

MOSESSON, M. W.; SIEBENLIST, K. R.; MEH D. A. The structure and biological features of fibrinogen and fibrin. **Annals of the New York Academy of Sciences**, v. 936, p. 11-30, 2001.



NANDI, S.; GUZMAN, R. C.; YANG, J. Hormones and mammary carcinogenesis in mice, rats, and humans: A unifying hypothesis. **Proc. Natl. Acad. Sci. USA**, 92, p. 3650-3657, 1995.

NCBI: National Center for Biotechnology Information. Disponível em <<http://www.ncbi.nlm.nih.gov/>>. Acesso em: abril de 2010.

MUTI, P. The role of endogenous hormones in the etiology and prevention of breast cancer: The epidemiological evidence. **RRCR**, v. 166, p. 245-256, 2005.

NAVARRETE, M. A.; MAIER, C. M.; FALZONI, R.; QUADROS, L. G.; LIMA, G. R.; BARACAT, E. C.; NAZÁRIO, A. C. Assessment of the proliferative, apoptotic and cellular renovation indices of the human mammary epithelium during the follicular and luteal phases of the menstrual cycle. **Breast Cancer Research**, v. 7, p. 306-313, 2005.

NUSSBAUM, R. L.; MCINNES, R. R.; WILLARD, H. F.; **THOMPSON & THOMPSON: Genética Médica**. 6. ed., Rio de Janeiro: Guanabara Koogan, 2002.

OMS: Organização Mundial De Saúde. Disponível em <<http://www.who.int/en/>> Acesso em: março de 2010.

ONG, S. E.; PANDEY, A. An evaluation of the use of two-dimensional gel electrophoresis in proteomics. **Biomolecular Engineering**, v. 8, p. 195-205, 2001.

OU, K.; YU, K.; KESUMA, D.; HOOI, M.; HUANG, N.; CHEN, W.; LEE, S. Y.; GOH, X. P.; TAN, L. K.; LIU, J.; SOON, S. Y.; RASHID, S. B. A.; PUTTI, T. C.; JIKUYA, H.; ICHIKAWA, T.; NISHIMURA, O.; SALTO-TELLEZ, M.; TAN, P. Novel breast cancer biomarkers identified by integrative proteomic and gene expression mapping. **Journal of Proteome Research**, v. 7, n. 4, p. 1518-1528, 2008.

PAGE, M. J.; AMESS, B.; TOWNSEND, R. R.; PAREKH, R.; HERATH, A.; BRUSTEN, L.; ZVELEBIL, M. J.; STEIN, R. C.; WATERFIELD, M. D.; DAVIES, S. C.; O'HARE, M. J. Proteomic definition of normal human luminal and myoepithelial breast cells purified from reduction mammoplasties. **Cell Biology**, v. 96, n. 22, p. 12589-12594, 1999.

PARKIN, D. M.; BRAY, F. I.; DEVESA, S. S. Cancer burden in the year 2000. The global picture. **Eur. J. Cancer**, v. 37, p. 4-66, 2001.

PATSTON, P. A. Serpins and other serine protease inhibitors. **Immunol. Today**, n. 21, p. 354, 2000.

POSADAS, E. M.; SIMPKINS, F.; LIOTTA, L. A.; MACDONALD, C.; KOHN, E. C. Proteomic analysis for the early detection and rational treatment of cancer—realistic hope? **Annals of Oncology**. v. 16, p. 16-22, 2005.

QIN, Z., KREPLAK, L., BUEHLER, M. J. Hierarchical structure controls nanomechanical properties of vimentin intermediate filaments. **PLoS ONE**, v. 4, n. 10, 2009.

RAMOS, C. M. P.; BORA, P. S. Functional characterization of acetylated Brazil nut (*Bertholletia excelsa* HBK) kernel globulin. **Ciênc. Tecnol. Aliment.**, n. 24, v. 1, p. 134-138, 2004.

REMBOLD, C. M.; O'CONNOR, M.; CLARKSON, M.; WARDLE, R. L.; MURPHY, R. A. *Signal Transduction in Smooth Muscle* Selected Contribution: HSP20 phosphorylation in nitroglycerin- and forskolin-induced sustained reductions in swine carotid media tone. **Journal of Applied Physiology**, v. 91, p. 1460-1466, 2001.

REYMOND, A.; MERONI, G.; FANTOZZI, A.; MERLA, G.; CAIRO, S.; LUZI, L.; RIGANELLI, D.; ZANARIA, E.; MESSALI, S.; CAINARCA, S.; GUFFANTI, A.; MINUCCI, S.; PELICCI, P. G.; BALLABIO, A. The tripartite motif family identifies cell compartments. **European Molecular Biology Organization Journal**, v. 20, n. 9, p. 2140-2151, 2001.

REYNOLDS, P.; HURLEY, S.; GOLDBERG, D. E.; ANTON-CULVER, A.; BERNSTEIN, L.; DEAPEN, D.; HORN-ROSS, P. L.; PEEL, D.; PINDER, R.; ROSS, R. K.; WEST, D.; WRIGHT, W. E.; ZIOGAS, A. Active smoking, household passive smoking, and breast cancer: evidence from the california teachers study. **Journal of the National Cancer Institute**. V. 96, n. 1, p. 29-37, 2004.

RITCHIE, R. F.; PALOMAKI, G. E.; NEVEUX, L. M.; NAVOLOTSKAIA, O.; LEDUE, T. B.; CRAIG, W. Y. Reference distributions for the negative acute-phase serum proteins, albumin, transferrin and transthyretin: a practical, simple and clinically relevant approach in a large cohort. **Journal of Clinical Laboratory Analysis**, v. 13, p. 273–279, 1999.

ROCHA, T. L. *et al.* Eletroforese bidimensional e análise de proteomas. **Comunicado Técnico Embrapa**. 11 p., 2005.

RONNOV-JENSEN, L.; PETERSEN, O. W.; BISSELL, M. J. Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. **Physiological Reviews**, v. 76, p. 69-125, 1996.

ROSOSCHANSKY, J.; SOARES, C. T. A importância das margens cirúrgicas no tratamento conservador do câncer de mama, In: RIETJENS, M.; URBAN, C. A. **Cirurgia da Mama Estética e Reconstructora**. Rio de Janeiro: Revinter, 2007. p. 360-363.

SCHRODER, E.; LITTLECHILD, J. A.; LEBEDEV, A. A.; ERRINGTON, N.; VAGIN, A. A.; ISUPOV, M. N. Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 angstrom resolution. **Structure**, v. 8, p. 605-615, 2000.

SHRINIVASAN, A.; POONGOTHAI, A. R. CHANDRASEKHAR, S. R.; SRINIVASULU, M.; VISHNUPRIYA, S. Serum lactate dehydrogenase (LDH) levels in breast cancer. **Indian Journal of Human Genetics**, v. 5, n. 2, p. 21-26, 1999.

SMITH, T. R.; MILLER, M. S.; LOHMAN, K. K.; CASE, L. D.; HU, J. J. DNA damage and breast cancer risk. **Carcinogenesis**, v. 24, p. 883-889, 2003.

SOKOLOV, A. V.; PULINA, M. O.; ZAKHAROVA, E. T.; SUSOROVA, A. S.; RUNOVA, O. L.; KOLODKIN, N. I.; VASILYEV, V. B. Identification and isolation from breast milk of ceruloplasmin–lactoferrin complex. **Biochemistry (Moscow)**, v. 71, n. 2, p. 160-166, 2006.

SOMIARI, R. I.; SOMIARI, S.; RUSSELL, S.; SHRIVER, C. D. Proteomics of breast carcinoma. **Journal of Chromatography B**, v. 815, p. 215-225, 2005.

SRINIVAS, P. R.; KRAMER, B. S.; SRIVASTAVA, S. Trends in biomarker research for cancer detection. **Lancet Oncol.** v. 2, p. 698–704, 2001.

STEHLE JUNIOR, J. R.; WEEKS, M. E.; LIN, C.; WILLINGHAM, M.C.; HICKS, A. M.; TIMMS, J. F.; CUI, Z. Mass spectrometry identification of circulating alpha-1-B glycoprotein, increased in aged female C57BL/6 mice. **Biochimica et Biophysica Acta**, n. 1770, p. 79–86, 2007.

SUN, Y. E3 ubiquitin ligases as cancer targets and biomarkers. **Neoplasia**, v. 8, n. 8, p. 645 – 654, 2006.

SUN, Y.; MacRAE, T. H. The small heat shock proteins and their role in human disease. **FEBS Journal**, v. 262, p. 2613-2627, 2005.

TERRY, P.; MILLER, A.; ROHAN, T. Cigarette smoking and breast cancer risk: a long latency period? **International Journal of Cancer**, v. 100, p. 723-728, 2002.

TOILLON, R. A.; LAGADEC, C.; PAGE, A.; CHOPIN, V.; SAUTIÈRE, P. E.; RICORT, J. M.; LEMOINE, J.; ZHANG, M.; HONDERMARCK, H.; LE BOURHIS, X. Proteomics demonstration that normal breast epithelial cells can induce apoptosis of breast cancer cells through insulin-like growth factor-binding protein-3 and maspin. **Molecular and Cellular Proteomics**, v. 6, p. 1239-1247, 2007.

TROYANOVSKY, S. M.; LEUBE, R. E.;FRANKE, W. W. Characterization of the human gene encoding cyokeratin 17 and its expression pattern. **Eur. J. Cell Biol.** v. 59, n. 1, p 127-137, 1992. (ABSTRACT).

VAN DER VUSSE, G. J. Albumin as fatty acid transporter. **Drug Metab. Pharmacokinet.**, v. 24, n. 4, p. 300–307, 2009.

VAN GENT, D.; SHARP, P.; MORGAN, K.; KALSHEKER, N. Serpins: structure, function and molecular evolution. **The International Journal of Biochemistry & Cell Biology**, n. 35, p 1536-1547, 2003.

VERBOVEN, C.; RABIJNS, A.; MAEYER, M. D.; BAELEN, H. V.; BOUILLON, R.; RANTER, C. D. A structural basis for the unique binding features of the human vitamin D-binding protein. **Nature Structural Biology**, v. 9, n. 2, p. 131-136, 2002.

WASSEL, J. Haptoglobin: function and polymorphism (ABSTRACT). **Clinical Laboratory**, n. 46, p. 547-552, 2000.

WEISEL, J. W. Fibrinogen and fibrin. **Advances in Protein Chemistry**, n. 70, p. 247-299, 2005.

WILKINS, M. R.; WILLIAMS, K. L.; APPEL, R. D.; HOCHSTRASSER, D. F.(Eds.). **Proteome research: New frontiers in functional genomics**. Berlin: Editora Springer-Verlag, 1997. 243 p.

WILKINSON, B.; GILBERT, H. F. Protein disulfide isomerase **Biochimica et Biophysica Acta 1699**, p. 35-44, 2004.

WESTERMEIER, R. Sensitive, quantitative, and fast modifications for coomassie blue staining of polyacrylamide gels. **Practical Proteomics**, v. 6, suppl. 2, p. 61-64, 2006.

WOGAN, G. N. Molecular epidemiology in cancer risk assessment and prevention: recent progress and avenues for future research. **Environ. Health Perspect.**, v. 98, p. 167-178, 1992.

WHITE, P.; COOKE, N. The multifunctional properties and characteristics of vitamin D-binding protein. **Trends Endocrinol. Metab.**, v. 11, n. 8, p. 320-327, 2000.

WULFKUHLE, J. D.; LIOTTA, L. A.; PETRICOIN, E. F. Proteomic applications for the early detection of cancer. **Nature Reviews – Cancer**, v. 3, p. 267-275, 2003.

YE, Y.; SHIBATA, Y.; YUN, C.; RON, D.; RAPAPORT, T. A. A membrane protein complex mediates retro-translocation from the ER lumen into the cytosol. **Nature**, v. 429, p. 841-847, 2004.

ZHANG, B.; ZHANG, Y.; DAGHER, M. C.; SHACTER, E. Rho GDP dissociation inhibitor protects cancer cells against drug-induced apoptosis. **Cancer Research**, v. 65(14), p. 6054-6072, 2005.

ZHU, K.; ZHAO, J.; LUBMAN, D. M.; MILLER, F. R.; BARDER, T. J. Protein pI shifts due to posttranslational modifications in the separation and characterization of proteins (ABSTRACT). **Analytical Chemistry**, v. 77, n. 9, p. 2745-2755, 2005.

## APÊNDICE

**Quadro com o nome em inglês das proteínas identificadas, suas respectivas bandas e a tolerância de erro utilizada para identificação.**

Nome da proteína (Código Swiss-Prot)	ID	Tolerância (ppm)
Collagen alpha-1(VI) chain (CO6A1_HUMAN)	139	200
	140	100
	142	50
	144	100
	145	100
	146	50
Alpha-1B-glycoprotein (A1BG_HUMAN)	185	100
	189	50
	197	50
Serotransferrin (TRFE_HUMAN)	192	150
	195	50
	198	50
	201	100
	196	100
Serum albumin (ALBU_HUMAN)	230	100
	271	50
	A	100
Antithrombin-III (ANT3_HUMAN)	260	50
Alpha-1-Antitrypsin (A1AT_HUMAN)	267	100
	272	50
	283	50
	617	200
	618	200
Vitamin D-binding protein (VTDB_HUMAN)	286	100
Vitamin D-binding protein (VTDB_HUMAN)	285 (mistura)	100
Alpha-1-Antitrypsin (A1AT_HUMAN)		100
Tubulin beta-2C chain (TBB2C_HUMAN)	292	150
Tubulin beta chain (TBB5_HUMAN)	298	50
ATP synthase subunit beta, mitochondrial (ATPB_HUMAN)	319	100
Fibrinogen gamma chain (FIBG_HUMAN)	320	50
	317	100
Actin, cytoplasmic 1 (ACTB_HUMAN); Actin, cytoplasmic 2 (ACTG_HUMAN)	359	100
		100
	362	100
		100
367	150	
	150	
Actin, aortic smooth muscle (ACTA_HUMAN); Actin, alpha cardiac muscle (ACTC_HUMAN); Actin, alpha skeletal muscle (ACTS_HUMAN)	361	100
		100
		100

Haptoglobin (HPT_HUMAN)	369	100
	379	50
	386	150
Glycerol-3-phosphate dehydrogenase [NAD+] (GPDA_HUMAN)	415	100
L-lactate dehydrogenase B chain (LDHB_HUMAN)	417 (mistura)	50
Glycerol-3-phosphate dehydrogenase [NAD+] (GPDA_HUMAN)		50
Annexin A5 (ANXA5_HUMAN)	433	100
Annexin A1 (ANXA1_HUMAN)	422	100
Serum amyloid P-component (SAMP_HUMAN)	459	150
Apolipoprotein A-I (APOA1_HUMAN)	478	50
Glutathione S-transferase P (GSTP1_HUMAN)	479	100
Peroxiredoxin-2 (PRDX2_HUMAN)	481	50
Vimentin (VIME_HUMAN)	355	100
	378	50
	276	100
	346	150
	351	150
	275	200
Creatine kinase B-type (KCRB_HUMAN)	364 (mistura)	50
Serum albumin (ALBU_HUMAN)		50
Tropomyosin alpha-4 chain (TPM4_HUMAN)	441	50
14-3-3 protein gamma (1433G_HUMAN)	453	100
14-3-3 protein beta/alpha (1433B_HUMAN)	460	50
Rho GDP-dissociation inhibitor 1 (GDIR1_HUMAN)	470	100
Heat shock protein beta-6 (HSPB6_HUMAN)	492	50
Transthyretin (TTHY_HUMAN)	514	50
Endoplasmin (ENPL_HUMAN)	164	50
Tripartite motif-containing protein 13 (TRI13_HUMAN)	262	100
Protein disulfide-isomerase (PDIA1_HUMAN)	268	50
Transitional endoplasmic reticulum ATPase (TERA_HUMAN)	170	50
Fibrinogen beta chain (FIBB_HUMAN)	392	100
	394	50
Tropomyosin beta chain (TPM2_HUMAN)	411	50
Phosphoenolpyruvate carboxykinase, cytosolic [GTP] (PCKGC_HUMAN)	491	100
Ceruloplasmin (CERU_HUMAN)	156	100
	157	120
E3 ubiquitin-protein ligase RAD18 (RAD18_HUMAN)	375	120
Keratin, type I cytoskeletal 17 (K1C17_HUMAN)	333	100
Actin-related protein 3B (ARP3B_HUMAN)	497	50

## ANEXOS

ANEXO 1	Interface do programa de busca MASCOT.....	100
ANEXO 2	Exemplo de identificação protéica.....	101
ANEXO 3	Dados complementares da identificação.....	102

## ANEXO 1 - Interface do programa de busca MASCOT

**MASCOT Peptide Mass Fingerprint**

<b>Your name</b>	<input type="text" value="Gustavo"/>	<b>Email</b>	<input type="text"/>
<b>Search title</b>	<input type="text" value="banda 433 (CP632N)"/>		
<b>Database(s)</b>	<input type="text" value="SwissProt"/> NCBI contaminants MSDB	<b>Enzyme</b>	<input type="text" value="Trypsin"/>
		<b>Allow up to</b>	<input type="text" value="1"/> missed cleavages
<b>Taxonomy</b>	<input type="text" value="..... Homo sapiens (human)"/>		
<b>Fixed modifications</b>	<input type="text" value="Carbamidomethyl (C)"/>	> <	<input type="text" value="Acetyl (K)"/> <input type="text" value="Acetyl (N-term)"/> <input type="text" value="Acetyl (Protein N-term)"/> <input type="text" value="Amidated (C-term)"/> <input type="text" value="Amidated (Protein C-term)"/> <input type="text" value="Ammonia-loss (N-term C)"/> <input type="text" value="Biotin (K)"/> <input type="text" value="Biotin (N-term)"/> <input type="text" value="Carbamyl (K)"/> <input type="text" value="Carbamyl (N-term)"/> <input type="text" value="Carboxymethyl (C)"/>
	Display all modifications <input type="checkbox"/>		
<b>Variable modifications</b>	<input type="text" value="Oxidation (M)"/>	> <	
<b>Protein mass</b>	<input type="text"/> kDa	<b>Peptide tol. ±</b>	<input type="text" value="100"/> ppm
<b>Mass values</b>	<input checked="" type="radio"/> MH <sup>+</sup> <input type="radio"/> M <sub>r</sub> <input type="radio"/> M-H <sup>-</sup>	<b>Monoisotopic</b>	<input checked="" type="radio"/> Average <input type="radio"/>
<b>Data file</b>	<input type="text" value="G:\espectrometria de massa\C-"/> <input type="button" value="Procurar..."/>		
<b>Query</b> NB Contents of this field are ignored if a data file is specified.	<input type="text"/>		
<b>Decoy</b>	<input type="checkbox"/>	<b>Report top</b>	<input type="text" value="AUTO"/> hits
<input type="button" value="Start Search ..."/>		<input type="button" value="Reset Form"/>	



## ANEXO 2 - Exemplo de identificação protéica

**MASCOT SCIENCE Mascot Search Results**

User : Gustavo  
 Email : gustavoogu@hotmail.com  
 Search title : banda 433 (CP632N)  
 MS data file : peaklist.xml  
 Database : SwissProt 2010\_04 (516081 sequences; 181677051 residues)  
 Taxonomy : Homo sapiens (human) (20280 sequences)  
 Timestamp : 15 Apr 2010 at 03:07:01 GMT  
 Top Score : 193 for **ANXA5\_HUMAN**, Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2

### Mascot Score Histogram

Protein score is  $-10 \cdot \log(P)$ , where P is the probability that the observed match is a random event.  
 Protein scores greater than 56 are significant ( $p < 0.05$ ).

Protein Score	Number of Hits
~20	5
~193	1

### Concise Protein Summary Report

Format As:  [Help](#)

Significance threshold  $p <$   Max. number of hits

- [ANXA5\\_HUMAN](#) Mass: 35971 Score: **193** Expect:  $1e-15$  Matches: 16  
 Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2  
[DOC2B\\_HUMAN](#) Mass: 46377 Score: 52 Expect: 0.14 Matches: 5  
 Double C2-like domain-containing protein beta OS=Homo sapiens GN=DOC2B PE=1 SV=1  
[AN34C\\_HUMAN](#) Mass: 58016 Score: 37 Expect: 4 Matches: 6  
 Ankyrin repeat domain-containing protein 34C OS=Homo sapiens GN=ANKRD34C PE=2 SV=1  
[AKT1\\_HUMAN](#) Mass: 56050 Score: 36 Expect: 5.5 Matches: 4  
 RAC-alpha serine/threonine-protein kinase OS=Homo sapiens GN=AKT1 PE=1 SV=2

## ANEXO 3 - Dados complementares da identificação

**Mascot Search Results****Protein View**

Match to: ANXA5\_HUMAN Score: 193 Expect: 1e-15  
 Annexin A5 OS=Homo sapiens GN=ANXA5 PE=1 SV=2  
 Found in search of peaklist.xml

Nominal mass ( $M_r$ ): 35971; Calculated pI value: 4.94  
 NCBI BLAST search of [ANXA5\\_HUMAN](#) against nr  
 Unformatted [sequence string](#) for pasting into other applications

Taxonomy: [Homo sapiens](#)

Fixed modifications: Carbamidomethyl (C)  
 Variable modifications: Oxidation (M)  
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P  
 Number of mass values searched: 30  
 Number of mass values matched: 16  
 Sequence Coverage: 52%

Matched peptides shown in **Bold Red**

1 MAQVLRGTVT **DFPGFDERAD** AETLRKAMK**G** LGTDEESILT LLTSRSNAQR  
 51 QEISAAFRTL FGRDLLDDLK SELTGKFEKL IVALMKPSRL YDAYELKHAL  
 101 KGAGTNEKVL **TEIIASRTP**E ELRAIK**QVYE** EEYGSSELEDD VVGDTSGYYQ  
 151 **RMLVLLQAN** RDPDAGIDEA **QVEQDAQALF** QAGELK**WGTD** EEK**FITIFGT**  
 201 **RSVSHLRKVF** DKY**MTISGFQ** IEETIDRETS GNLEQLLAV VKSIRSIPAY  
 251 LAETLYYAMK **GAGTDDHTLI** RVMVSR**SEID** LFNIRKEFRK **NFATSLYSMI**  
 301 **KGDTSGDYKK** ALLLLCGEDD

Show predicted peptides also

Sort Peptides By

Residue Number  Increasing Mass  Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Sequence
7 - 18	1340.6056	1339.5983	1339.6045	-5	0	R.GTVTDFPGFDER.A
30 - 45	1704.7703	1703.7630	1703.8941	-77	0	K.GLGTDEESILTLLTSR.S
80 - 89	1143.7041	1142.6968	1142.6845	11	0	K.LIVALMKPSR.L Oxidation (M)
90 - 97	1014.5415	1013.5342	1013.5069	27	0	R.LYDAYELK.H
109 - 117	1001.6797	1000.6725	1000.5917	81	0	K.VLTEIIASR.T
127 - 151	2888.2336	2887.2264	2887.2308	-2	0	K.QVYEEYGSSELEDDVVGDTSGYYQR.M
152 - 161	1156.7366	1155.7293	1155.6798	43	0	R.MLVLLQANR.D
152 - 161	1172.6965	1171.6893	1171.6747	12	0	R.MLVLLQANR.D Oxidation (M)
162 - 186	2658.2590	2657.2518	2657.2457	2	0	R.DPDAGIDEAQVEQDAQALFQAGELK.W
194 - 201	954.6252	953.6180	953.5335	89	0	K.FITIFGTR.S
213 - 227	1802.7736	1801.7663	1801.8556	-50	0	K.YMTISGFQIEETIDR.E
213 - 227	1818.7821	1817.7748	1817.8506	-42	0	K.YMTISGFQIEETIDR.E Oxidation (M)
261 - 271	1155.6401	1154.6329	1154.5680	56	0	K.GAGTDDHTLIR.V
277 - 285	1106.6189	1105.6116	1105.5768	32	0	R.SEIDLFNIR.K
291 - 301	1274.6274	1273.6202	1273.6376	-14	0	K.NFATSLYSMIK.G
291 - 301	1290.6328	1289.6255	1289.6326	-5	0	K.NFATSLYSMIK.G Oxidation (M)

No match to: 831.5469, 833.1635, 869.4998, 871.1216, 907.4598, 1128.6058, 1144.5649, 1182.5380, 1354