

INSTITUTO OSWALDO CRUZ

Mestrado em Biologia Celular e Molecular

**CARACTERIZAÇÃO PROTEICA DO FLUIDO ASPIRADO DE MAMA DE
PACIENTES BRASILEIRAS COM CÂNCER DE MAMA UNILATERAL**

GISELLE VILLA FLOR BRUNORO

**Rio de Janeiro
2010**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Ficha catalográfica elaborada pela
Biblioteca de Ciências Biomédicas/ ICICT / FIOCRUZ - RJ

B898

Brunoro, Giselle Villa Flor.

Caractetização proteica do fluido aspirado de mama de pacientes
brasileiras com câncer de mama unilateral / Giselle Villa Flor Brunoro. –
Rio de Janeiro, 2010.

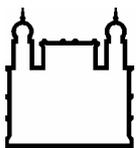
xix, 104 f. : il. ; 30 cm.

Dissertação (mestrado) – Instituto Oswaldo Cruz, Pós-Graduação em
Biologia Celular e Molecular, 2010.

Bibliografia: f. 88-98.

1. Câncer. 2. Mama. 3. Proteômica. 4. NAF. I. Título.

CDD 370.711



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ
Pós-Graduação em Biologia Celular e Molecular

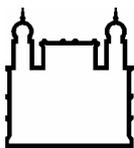
GISELLE VILLA FLOR BRUNORO

Caracterização proteica do fluido aspirado de mama de pacientes brasileiras com câncer de mama unilateral.

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Ciências.

Orientadora: Prof^ª. Dr^ª. Ana Gisele da Costa Neves Ferreira

RIO DE JANEIRO
2010



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Pós-Graduação em Biologia Celular e Molecular

AUTOR: Giselle Villa Flor Brunoro

CARACTERIZAÇÃO PROTEICA DO FLUIDO ASPIRADO DE MAMA DE PACIENTES BRASILEIRAS COM CÂNCER DA MAMA UNILATERAL.

ORIENTADORA: Prof^ª. Dr^ª. Ana Gisele da Costa Neves Ferreira

Aprovada em: 07 / 05 / 2010

EXAMINADORES:

Prof(a). Dr(a).	Patricia Torres Bozza - Presidente
Prof(a). Dr(a).	Claudia Vitoria de Moura Gallo
Prof(a). Dr(a).	Gilberto Barbosa Domont
Prof(a). Dr(a).	Luciana Pizzatti Barboza
Prof(a). Dr(a).	Richard Hemmi Valente

Rio de Janeiro, 07 de maio de 2010

Dedicatória

Dedico este trabalho aos meus pais,
Jose e Edelza, que sempre estiveram ao meu lado.
Pelo amor incondicional a mim dedicado.

Agradecimentos

Agradeço primeiramente aos quatro pesquisadores que conduziram este projeto de dissertação. À Dra. Ana Gisele, pesquisadora competente e orientadora que conduziu de forma segura este projeto. Agradeço também por ter compartilhado comigo a sua experiência e conhecimento. Ao Dr. Dante, idealizador, pensador e entusiasta. Por toda a sua inspiração contagiante, experiência e conhecimento que o fazem um verdadeiro cientista. Ao Dr. Jonas Perales, chefe do laboratório e grande incentivador da pesquisa científica. Por toda a sua dedicação para a realização deste projeto, garantindo os recursos necessários, e pela confiança em mim depositada. À Dra. Claudia Gallo, brilhante pesquisadora, por me inspirar e me acolher na vinda ao Rio de Janeiro para a realização desta importante etapa da minha vida em busca do crescimento científico. Agradeço pela sua valiosa contribuição na revisão e avaliação desta dissertação.

Aos demais membros da banca examinadora que aceitaram com carinho este convite, a Dra. Patricia Torres Bozza, Dr. Gilberto Barbosa Domont, Dra. Luciana Pizzatti Barboza e Dr. Richard Hemmi Valente.

Ao Dr. Júlio Paixão, por toda a contribuição e suporte na coleta das amostras no Instituto Fernandes Figueira.

Ao Dr. Napoleão Leão, chefe do Serviço de Ginecologia do Hospital da Lagoa, por abrir as portas e acolher este projeto com tanto entusiasmo.

A toda equipe de enfermeiras, dentistas e serventes do Hospital da Lagoa pelo auxílio e carinho na coleta de amostras.

A cada paciente que mesmo num momento difícil aceitou participar do projeto. Por me ensinar que o sofrimento pode ser fonte de força.

À Dra. Leila Lopes Bezerra e toda sua equipe, pelo grande auxílio na digitalização das imagens no scanner de fluorescência.

Ao Dr. Carlos Bloch, competente pesquisador do Laboratório de Espectrometria de massas da EMBRAPA, e a sua equipe. Pela generosidade e por me acolher em seu laboratório.

Ao grupo do *Protein Research Group*, Departamento de Bioquímica e Biologia Molecular, da Universidade do Sul da Dinamarca, Odense, Dinamarca.

Ao Dr. Luis Claudio, por toda a contribuição na coleta de amostras.

Ao Dr. Donat Alexander, pelos importantes ensinamentos e colaborações.

Ao Dr. André Ferreira e Dra. Daniela Beghini, que além de serem profissionais excepcionais e contribuidores deste trabalho, são grandes amigos. Por todo o carinho, amizade e colaboração diária.

À Dra. Monique, por todo o auxílio na análise de imagens.

Aos demais membros do Laboratório de Toxinologia, Surza, Karina, André, Thayssa, Carolina, Caroline, Luma, Larissa, Viviane, Graziela, Jayron e Paulo, pela agradável convivência e por toda contribuição e carinho durante este trabalho.

À Lidiane, antiga membro do Laboratório, pelos primeiros ensinamentos.

Aos antigos membros do laboratório, em especial a Patrícia, Simone, Ileana, Frank, Marcela, pelo carinho e acolhida.

A todas as pessoas de outros laboratórios, em especial a Marina, Silvia, Valéria, Evelyn, Karen, Juliana, Rubem, Nadja e Tatiana Paula, por todo o carinho.

Aos funcionários da pós-graduação do Instituto Oswaldo Cruz, pela atenção e boa vontade demonstrada ao longo de todo o curso.

Aos funcionários do Pavilhão Ozório de Almeida, pelo suporte laboratorial.

Às plataformas tecnológicas de Espectrometria de massas e de Eletroforese bidimensional e fracionamento – RJ, por todo o suporte laboratorial para a realização deste trabalho.

Ao Setor de Produção e Tratamento de Imagens do IOC/Fiocruz, pela disponibilidade em fotografar os géis de zimografia deste trabalho.

À Rede Proteômica do Rio de Janeiro, por todo o apoio.

Às agências de fomento, CNPq, CAPES, FAPERJ e ao PDTIS/Fiocruz, pelo financiamento deste projeto e pela bolsa de mestrado.

À Daniel Ribeiro e família, amigos e grandes incentivadores nos primeiros passos deste trabalho.

Às amigas Daisy e Lívia, pelo incentivo, carinho, apoio e amizade infinita! Vocês estão no meu coração!

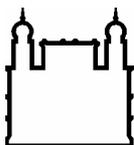
Aos amigos Wellington e Michelle, pelas conversas diárias.

Ao meu namorado Philipe, por toda a compreensão e paciência. Por me ouvir, acalmar, incentivar e enxugar cada lágrima (que não foram poucas!). Pela presença, companheirismo e amor. Amo você!

À minha irmã Michelle que mesmo distante me apoiou incondicionalmente. Por todas as orações e carinho. Te amo, maninha!

Aos meus pais, Jose e Edelza, preciosidades da minha vida e meus amores! Por mais palavras que eu possa escrever, elas não preencheriam a magnificência de seu apoio incansável e incentivo constante. As palavras não podem descrever o que seus olhos podem ler em meus olhos. No rosto de vocês, reflete não só a imensa felicidade desta conquista, mas também a imagem de quem acreditou, em forma de confiança e dedicação, sendo amigos e companheiros. Vocês pacientemente dividiram comigo essa luta, me encorajando a persistir e prosseguir. Dedico este trabalho a vocês!

Ao meu Deus e à minha Mãezinha do céu, agradeço por todos os momentos em que pensei não resistir mas lá estavam para me amparar, proteger, guiar e amar, fazendo enxergar que com fé tudo posso naquele que me fortalece. Pelo dom da vida, por mais uma etapa vencida, pelo que passou e pelo futuro. Receba Senhor minha alegria e minha gratidão.



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

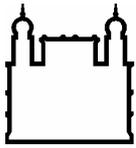
CARACTERIZAÇÃO PROTEICA DO FLUIDO ASPIRADO DE MAMA DE PACIENTES BRASILEIRAS COM CÂNCER DE MAMA UNILATERAL

RESUMO

DISSERTAÇÃO DE MESTRADO

Giselle Villa Flor Brunoro

O câncer de mama é a principal causa de morte por câncer em mulheres no mundo. De acordo com o Instituto Nacional do Câncer, 7.470 novos casos são estimados em 2010 apenas no Rio de Janeiro. A maioria dos casos de câncer de mama são originados nas células ductais mamárias. Elas secretam o fluido aspirado de mama (NAF), que contém proteínas intimamente associadas ao microambiente tumoral. O objetivo deste trabalho foi utilizar a abordagem proteômica como ferramenta para caracterizar e comparar o NAF das duas mamas de pacientes brasileiras com câncer da mama unilateral. A concentração média de proteínas das amostras controle, tumor benigno ou câncer de mama de 13 mulheres foi aproximadamente $115 \pm 60 \mu\text{g}/\mu\text{L}$. A atividade proteolítica do NAF foi visualizada por zimografia das amostras de duas pacientes. Observamos que o padrão de bandas gelatinolíticas diferiu substancialmente entre as pacientes, mas não entre as duas mamas do mesmo indivíduo. As enzimas proteolíticas se apresentaram sensíveis à inibição por agentes quelantes de metais (EDTA e orto-fenantrolina). As amostras de NAF também foram submetidas à 2D-PAGE em condições otimizadas (IPG 4-7 e aplicação de amostra por reidratação *in gel*) e poucos *spots* proteicos muito abundantes foram detectados e identificados por espectrometria de massas MALDI-TOF/TOF como albumina, imunoglobulinas, zinco- α 2-glicoproteína e proteína induzida por prolactina. Para enriquecer o NAF em suas proteínas menos abundantes, testamos estratégias de depleção usando a coluna de múltipla afinidade Hu-6 da Agilent ou proteína L agarose. Este passo de pré-fracionamento introduziu variabilidade adicional importante nas amostras e por este motivo não foi empregado em etapas posteriores do estudo. Catorze amostras de NAF das mamas pareadas de mulheres com câncer unilateral foram sistematicamente comparadas por 2D-DIGE. Foram observadas diferenças qualitativas substanciais entre as pacientes, prejudicando as comparações quantitativas entre as amostras. Dois grandes perfis proteicos distintos foram definidos por 2D-DIGE. O perfil A caracterizou-se pela presença de dois grupos de *spots* ácidos típicos ($\text{pI} \approx 4,5$ e $\text{MM} 29,5$ e 18 kDa), ausentes no perfil B. A presença de glico- e fosfo-proteínas foi determinada em amostras de ambos os perfis após revelação dos géis 2D com ProQ- Emerald ou Diamond, respectivamente. As amostras dos dois perfis se mostraram muito mais ricas em glicoproteínas do que em fosfoproteínas. Este trabalho constituiu a primeira caracterização proteômica sistemática do NAF de pacientes brasileiras com câncer de mama unilateral. Nossos resultados mostraram uma variabilidade importante do conteúdo de proteínas nas amostras individuais de NAF, sugerindo a necessidade de uma estratificação mais uniforme das pacientes. Por fim, o emprego de tecnologias proteômicas mais sensíveis e acuradas deverá ser necessário para possibilitar a detecção e identificação de proteínas menos abundantes no NAF, que poderão constituir candidatos a biomarcadores do câncer de mama.



Ministério da Saúde

FIOCRUZ
Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

CARACTERIZAÇÃO PROTEICA DO FLUIDO ASPIRADO DE MAMA DE PACIENTES BRASILEIRAS COM CÂNCER DE MAMA UNILATERAL

ABSTRACT

DISSERTAÇÃO DE MESTRADO

Giselle Villa Flor Brunoro

Breast cancer is the main cause of cancer deaths worldwide in women. According to the Instituto Nacional do Câncer, 7.470 new cases are estimated for 2010 in Rio de Janeiro. Most breast cancer cases originate from mammary ductal cells that secrete the nipple aspirate fluid (NAF), which contains proteins associated with the breast cancer microenvironment. The aim of this study was to use the proteomic approach as a reliable tool to characterize NAF samples from Brazilian patients with unilateral breast cancer. The mean concentration of NAF proteins from healthy, benign or cancer breast samples from 13 women was approximately $115 \pm 60 \mu\text{g}/\mu\text{L}$. The proteolytic activity of NAF samples from both breasts of two patients was visualized by SDS-PAGE zymography. The patterns of gelatinolytic bands differed substantially between patients but were very similar within the same patient. The proteolytic activity was sensible to inhibition by metal quelators (EDTA and orthophenantrolin). In this study, NAF samples were submitted to 2D-PAGE under optimized conditions (IPG 4-7 and sample in gel rehydration) and a few very abundant protein spots were detected. They were identified by MALDI-TOF/TOF MS as albumin, immunoglobulins, zinc- α 2-glycoprotein and prolactin-induced protein. To enrich NAF for less abundant proteins, we tested protein depletion using Agilent Hu-6 column or Protein L Agarose. These pre-fractionation steps introduced important additional variability in the samples and were not further used. Fourteen NAF samples from tumorous and non-tumorous breasts of women with unilateral cancer were systematically compared by 2D-DIGE. Substantial qualitative individual differences were observed, impairing proper quantitative comparisons. Two main 2D-DIGE protein profiles were observed: profile A was characterized by two groups of acidic spots ($\text{pI} \approx 4,5$ and $\text{MM} 29,5$ and 18 kDa), not detected in profile B. The presence of glyco- and phosphoproteins was determined after staining NAF 2D gels with ProQ-Emerald or Diamond, respectively. The samples from both profiles were much richer in glycoproteins than phosphoproteins. This work was the first systematic proteomic characterization of NAF from Brazilian patients with unilateral breast cancer. The results highlight important variability in the protein content of NAF samples, suggesting that a more uniform patient stratification and the use of more sensible and accurate proteomic technologies will be needed for the detection and identification of breast cancer biomarker candidates.

Índice

1 – Introdução	1
1.1. Carcinogênese	2
1.2. A mama	4
1.3. Classificação histológica dos carcinomas de mama	6
1.4. Epidemiologia do câncer de mama	7
1.5. Fatores de risco e métodos diagnósticos	8
1.6. O fluido aspirado de mama (NAF)	10
1.7. Proteômica	11
1.8. Principais técnicas proteômicas	12
1.8.1. Cromatografia Líquida	12
1.8.2. Eletroforese bidimensional	12
1.8.3. Eletroforese diferencial em gel bidimensional (2D-DIGE)	14
1.8.4. Espectrometria de massas	15
1.9. Justificativa do trabalho	17
2 – Objetivos	19
3 - Material e métodos	21
3.1. Coleta das amostras de fluido de mama (NAF)	22
3.2. Dosagem de proteínas pelo ácido bicinchonínico (BCA)	22
3.3. Precipitação de proteínas pelo ácido tricloroacético (TCA)	23
3.4. Eletroforese unidimensional	23
3.5. Zimografia	24
3.6. Depleção das proteínas mais abundantes do NAF	24
3.6.1. Cromatografia de afinidade usando anticorpos policlonais monoespecíficos	24
3.6.2. Cromatografia de afinidade usando Proteína L agarose	25
3.7. Eletroforese bidimensional convencional (2D-PAGE)	25
3.8. Revelação dos géis	27
3.8.1. Revelação dos géis por Coomassie R-250	27
3.8.2. Revelação dos géis por Coomassie coloidal G-250	27
3.8.3. Impregnação dos géis por nitrato de prata	28
3.8.4. Detecção específica de glico- e fosfo-proteínas	28

3.9. Eletroforese de fluorescência diferencial (2D-DIGE)	29
3.9.1. Dosagem de proteínas pelo 2D-Quant kit	29
3.9.2. Análise por 2D-DIGE	30
3.9.3. Digitalização e análise das imagens	30
3.10. Identificação de proteínas por espectrometria de massas	31
3.10.1. Digestão trípica em gel de poliacrilamida	31
3.10.2. Análise por MALDI-TOF/TOF MS	32
3.11. Análise preliminar do NAF por LC-MS/MS	33
3.11.1. Digestão trípica em solução	33
3.11.2. Identificação das proteínas no LTQ-Orbitrap XL	33
4 - Resultados e discussão	35
4.1. Coleta das amostras de NAF e dosagem de proteína	36
4.2. Análise do NAF por eletroforese unidimensional	40
4.3. Análise zimográfica do NAF	44
4.4. Otimização das condições de 2D-PAGE para as amostras de NAF	47
4.5. Caracterização do conteúdo proteico do NAF por 2D-PAGE e espectrometria de massas	50
4.6. Depleção de proteínas abundantes do NAF por afinidade com anticorpos	55
4.7. Depleção de proteínas abundantes do NAF por afinidade com proteína L	60
4.8 – Análise de amostras pareadas de NAF por DIGE	64
4.9. Análise do conteúdo de glicoproteínas do NAF	70
4.10. Análise do conteúdo de fosfoproteínas do NAF	74
4.11. Perspectivas futuras: caracterização do NAF por LC-MS/MS	78
5 – Conclusões	85
6 - Referências bibliográficas	88
7 – Anexos	99
7.1. Anexo I	100
7.2. Anexo II	103

Lista de Tabelas

Tabela 4.1: Características clínico-patológicas das amostras do estudo	39
Tabela 4.2: Identificações por MALDI TOF/TOF das bandas de eletroforese unidimensional de NAF	43
Tabela 4.3: Identificações por MALDI TOF/TOF dos spots de 2D-PAGE das duas mamas pareadas de uma paciente	53
Tabela 4.4: Identificações por MALDI TOF/TOF dos spots da 2D-PAGE da fração do NAF que interage com a coluna de múltipla afinidade da Agilent	57
Tabela 4.5: Identificações por MALDI TOF/TOF das bandas de eletroforese unidimensional da fração do NAF que interage com a Proteína L	62
Tabela 4.6: Amostras pareadas de mamas controle e mamas com câncer utilizadas na 2D-DIGE	65
Tabela 4.7: Identificações do NAF de quatro pacientes por LC-MS/MS	81

Lista de Figuras

Figura 1.1: Diagrama esquemático da anatomia da mama	5
Figura 1.2: Representação esquemática de um experimento utilizando a técnica de fluorescência diferencial em 2D-DIGE	15
Figura 4.1: Concentração média de proteínas nas amostras dos grupos controle, tumor benigno ou câncer	38
Figura 4.2: Gráfico demonstrando a concentração proteica das mamas pareadas de sete pacientes com câncer de mama unilateral	38
Figura 4.3: SDS-PAGE do NAF de indivíduos dos grupos controle e tumor benigno após normalização por volume de amostra ou quantidade de proteína	42
Figura 4.4: Zimografias de amostras de NAF coletadas de mamas doentes e sadias de duas pacientes	46
Figura 4.5: 2D-PAGE de uma amostra de NAF de indivíduo normal com intervalos de pH de 3-10NL e 4-7	49
Figura 4.6: 2D-PAGE de uma amostra de NAF de indivíduo normal, utilizando os métodos de aplicação de amostra de <i>cup loading</i> e reidratação <i>in gel</i>	49
Figura 4.7: 2D-PAGE das amostras das mamas sadia e doente de uma paciente com carcinoma ductal infiltrante unilateral	52
Figura 4.8: 2D-PAGE da fração de NAF da mama controle de indivíduo com câncer (P2d) que interagiu com a coluna de múltipla afinidade da Agilent	56
Figura 4.9: SDS-PAGE das frações de NAF de uma amostra controle (P5e) eluídas da coluna de proteína L agarose	61
Figura 4.10: 2D-PAGE da mesma amostra de NAF de um indivíduo controle (P5d) após a depleção das imunoglobulinas por proteína L (A) e sem depleção	63
Figura 4.11: Imagens da amostra de padrão interno marcado com Cy2 dos sete géis usados no experimento de DIGE	66

Figura 4.12: Imagens de DIGE das amostras de NAF classificadas no perfil A	67
Figura 4.13: Imagens de DIGE das amostras de NAF classificadas no perfil B	68
Figura 4.14: 2D-PAGE de NAF das mamas da paciente P29 que apresenta o perfil A corado com <i>Pro-Q Emerald</i> e com Coomassie G-250	72
Figura 4.15: 2D-PAGE de NAF das mamas da paciente P34 que apresenta o perfil B corado com <i>Pro-Q Emerald</i> e com Coomassie G-250	73
Figura 4.16: 2D-PAGE de NAF das mamas da paciente P29 que apresenta o perfil A corado com <i>Pro-Q Diamond</i> e com Coomassie G-250	76
Figura 4.17: 2D-PAGE de NAF das mamas da paciente P34 que apresenta o perfil B corado com <i>Pro-Q Diamond</i> e com Coomassie G-250	77

Lista de abreviaturas

- %C – porcentagem da proporção de *cross-linker*
- %T – porcentagem do peso total de monômero
- 2D-PAGE - eletroforese bidimensional em gel de poliacrilamida
- AGC - *Automatic Gain Control*
- AMP – adenosina monofosfato
- APMSF – *4-(amidinophenyl)methanesulphonyl fluoride*
- ASB14 - Amidosulfobetaina-14
- BCA – ácido bicinchonínico
- BFB – azul de bromofenol
- BRCA – breast cancer gene*
- BSA - albumina sérica bovina
- BVA - *Biological Variation Analysis*
- CDI - carcinoma ductal infiltrante
- CDIS - carcinoma ductal *in situ*
- CDK - quinase dependente de ciclina
- CHAPS - 3-[(3-colamido propil) dimetil amônio]-1-propanosulfonato
- CID - *collision-induced dissociation*
- CIS - carcinoma *in situ*
- CLI - carcinoma lobular infiltrante
- CLIS - carcinoma lobular *in situ*
- cm - centímetro
- CONEP - Comissão Nacional de Ética em Pesquisa
- Cys - cisteína
- Da - Dalton
- DIA - *Differential In-gel Analysis*
- DIGE - eletroforese diferencial em gel bidimensional
- DTT - ditioneitol
- E-64 – [trans-epoxysuccinyl-L-leucylamido-(4-guanidino)butane]
- EDTA – ácido Etilenodiamino tetra-acético
- EGFR - receptor de fator de crescimento epitelial

Em - emissão
ER - receptor de estrógeno
Ex - excitação
FWHM - *full width at half maximum*
GCDFP-15 - *gross cystic disease fluid protein -15*
HER2 - receptor tipo 2 do fator de crescimento epidérmico humano
His - histidina
IAA - iodoacetamida
IARC - Agência Internacional para Pesquisa em Câncer
ICAT - *isotope coded affinity tag*
IEF – focalização isoelétrica
IGA – imunoglobulina A
IgD – imunoglobulina D
IgE – imunoglobulinaE
IGF-1R - receptor do fator de crescimento de insulina
IgG – imunoglobulina G
IL - interleucina
INCA - Instituto Nacional de Câncer
IPG - gradiente imobilizado de pH
iTRAQ - isobaric tags relative and absolute quantification
kDa - kiloDalton
kV - quilovolt
LC - cromatografia líquida
LH – hormônio luteinizante
LTQ- *trap* linear
M - molar
m/z - massa/carga
mA - miliampere
MALDI - *matrix-assisted laser desorption ionization*
MAPK - proteína quinase mitógeno ativada
Met – metionina
mg - miligramas
min - minuto

mL - mililitros
mM – milimolar
MM - padrão de massa molecular
MMP - metaloproteases de matriz
mRNA – ácido ribonucléico mensageiro
MS – espectrometria de massas
µg - microgramas
µM - micromolar
µL - microlitros
NAF - *Nipple Aspirate Fluid*
nESI – *nano electron spray ionization*
NF-κB - fator nuclear kappa B
nL - nanolitro
nm - nanômetro
PAGE - eletroforese em gel de poliacrilamida
PBS - solução salina tamponada com fosfato
pH – potencial de hidrogênio
PI3K - fosfatidilinositol 3-quinase
PMF - *peptide mass fingerprinting*
pmol - picomol
PSD - *post-source decay*
qsp – quantidade suficiente para
SDS – dodecil sulfato de sódio
seg – segundo
SELDI - *surface-enhanced laser desorption ionization*
TCA – ácido tricloroacético
TF – antígeno Thomsen-Friedenreich
TFA – ácido trifluoroacético
Tn – antígeno Tn
TNF-α - fator de necrose tumoral alfa
TNM - tumor primário, linfonodos e metástases
TOF - *time-of-flight*
Trp – triptofano

UICC - *Union Internationale Contre Cancer*

V - volt

VhS – volt hora por passo

WHO - *World Health Organization*

ZAG - zinco- α 2-glicoproteina

1 – INTRODUÇÃO

1.1 – Carcinogênese

Os cânceres são causados pelo acúmulo de mutações em genes que controlam o crescimento (via regulação de proliferação e apoptose) e diferenciação das células. Como resultado, há crescimento celular exagerado e descontrolado, com invasão do tecido adjacente. Estas mutações afetam genes importantes para a manutenção do ciclo celular, levando à perda de função antitumoral (genes supressores de tumor ou antioncogenes) ou ao ganho de função pró-tumoral (oncogenes). Por exemplo, a proteína codificada pelo antioncogene p53 reconhece o mau emparelhamento da dupla fita de DNA e inibe a progressão do ciclo celular por inativação do complexo CDK (quinase dependente de ciclina) - ciclina. Quando o gene sofre mutações, a proteína perde sua função de monitorar a replicação do DNA e a frequência geral de mutações gênicas aumenta. Outro exemplo é o bcl-2, um gene que regula negativamente a apoptose. Quando fundido por translocação com um acentuador transcricional, mantém-se ativo constitutivamente, aumentando o tempo de vida da célula, durante o qual são acumuladas mutações promotoras da proliferação (Griffiths, 2002).

As maquinarias de proliferação e morte programada das células devem estar subordinadas às necessidades da população de células. Para coordenar estes sistemas, sinais precisam ser transmitidos entre as células. As vias intercelulares de sinalização são compostas pelas moléculas sinalizadoras, seus receptores e pelo sistema de transdução de sinal. A comunicação é estabelecida por mecanismos de fosforilação, interação alostérica e interação de subunidades proteicas. As moléculas sinalizadoras podem ser secretadas na circulação como sinais endócrinos (ex.: hormônios) ou localmente pelas células adjacentes (como sinais parácrinos)(Griffiths, 2002). Uma molécula sinalizadora relacionada ao câncer de mama bem conhecida é o estrógeno. Ele pode ativar o receptor de estrógeno (ER) pelas vias genômica e não-genômica e promover a proliferação de células malignas na mama (Yamaguchi, 2007). Pela via genômica de sinalização, o estrógeno se liga ao ER presente no núcleo celular. Este receptor se dimeriza, funcionando como um fator de transcrição que se liga a elementos de resposta ao estrogênio (ERE) nas regiões promotoras de genes específicos. Pela via não-genômica, o estrógeno pode atuar como co-ativador e interage com outros fatores de transcrição (ex.: NF-κB) por interação proteína-proteína. Adicionalmente, o ER ligado ao estrógeno pode se associar a receptores de fatores de crescimento (IGF-1R - receptor do fator de crescimento de insulina; EGFR - receptor de fator

de crescimento epitelial; ou HER2 - receptor tipo 2 do fator de crescimento epidérmico humano) na face citoplasmática da membrana plasmática. O estrógeno induz a formação de complexos (ex.: ER + IGF-1R + proteína adaptadora Shc) que ativam por fosforilação o receptor de fator de crescimento que, por sua vez, induz a transdução de sinal através de sinais específicos (ex.: proteínas G e AMP cíclico) levando a ativação de quinases como a MAPK (proteína quinase mitógeno ativada) e a PI3K (fosfatidilinositol 3-quinase). O estrógeno pode ser produzido a partir do andrógeno, por ação da enzima aromatase. Esta enzima pode ser expressa localmente no tecido tumoral, principalmente em mulheres pós-menopausa, como resultado da comunicação entre o tumor e o estroma através da produção de fatores como TNF- α (fator de necrose tumoral- α), IL-6 e IL-11 (interleucinas 6 e 11), citado por Yamaguchi (2007) (Ali e Coombes, 2002; Yamaguchi, 2007).

O câncer progride histologicamente da seguinte forma: no tecido normal, células epiteliais diferenciadas e estratificadas são separadas pela membrana basal do compartimento estromal, composto de matriz (fibras de colágeno e fibronectina), fibroblastos, capilares sanguíneos e poucos leucócitos. A homeostase do tecido é mantida por interações entre estas células e o estroma. No estado pré-maligno ou displásico, ocorre hiperplasia do tecido, caracterizada por células epiteliais pouco diferenciadas, fibroblastos ativados, maturação de mais capilares e aumento da quantidade de leucócitos no estroma. A progressão para um carcinoma está associada à proliferação desordenada de células indiferenciadas invadindo o estroma através da membrana basal, que é degradada. A matriz extracelular é remodelada pela degradação das cadeias de colágeno, a angiogênese é mantida, os fibroblastos se diferenciam em miofibroblastos e há aumento de células inflamatórias (linfócitos, neutrófilos, macrófagos), resultando na expressão de fatores de crescimento, componentes de matriz, proteases e citocinas (Mueller e Fusenig, 2004).

A comunicação entre as células tumorais e o estroma ativado é muito importante para a progressão do tumor (Liotta e Kohn, 2001). As células tumorais ativam o microambiente estromal através da secreção de fatores de crescimento (fator de crescimento de fibroblasto básico, fator de crescimento endotelial vascular, fator de crescimento derivado de plaquetas, fator de crescimento de transformação) e proteases (metaloproteases de matriz, como a MMP1, MMP3, MMP9, entre outras) que agem de forma autócrina e parácrina. Simultaneamente, estas células iniciam a secreção de componentes de matriz pró-migratórios (por exemplo, as integrinas) e pró-invasivos (tenascina), seus respectivos

receptores e reduzem a expressão de inibidores de proteases. O desequilíbrio entre as proteases e seus inibidores resulta na degradação da matriz que, como consequência, expõe fragmentos reativos de colágeno e fibronectina e mobiliza fatores de crescimento de matriz. Todas essas moléculas induzem angiogênese e recrutamento de mais células inflamatórias e fibroblastos que, por sua vez, secretam mais fatores de crescimento e proteases, amplificando o sinal e contribuindo para o recrutamento de mais células, proliferação e ativação do estroma (LaRocca *et al.*, 2004; Mueller e Fusenig, 2004).

1.2 – A mama

A glândula mamária é uma glândula túbulo-alveolar similar às glândulas sudoríparas organizada em seis a dez sistemas ductais principais. Cada sistema drena para um ducto excretor principal chamado de seio lactífero. As ramificações sucessivas destes ductos originam ramos terminais diferenciados em lóbulos (Figura 1.1). Estes lóbulos são constituídos por alvéolos e formam as unidades funcionais da mama. Os lóbulos e os ductos são compostos por células epiteliais cujas principais funções são produzir, secretar e conduzir o leite materno ao mamilo. O epitélio mamário é revestido por uma camada achatada de células contráteis (células mioepiteliais) anexada à membrana basal, que tem papel na manutenção estrutural dos lobos e ductos. O estroma da mama consiste em tecido fibroconjuntivo denso, misturado com tecido adiposo contendo fibroblastos, fibras elásticas e adipócitos (Franco, 1997; Contran, 2000).

Tradicionalmente, a mama é considerada uma glândula de secreção apócrina, caracterizada pela secreção por decapitação da região apical da célula (Mazoujian *et al.*, 1983). Entretanto, este aspecto vem sendo contestado, argumentando-se que, na realidade, a secreção se dá por exocitose, assim como nas glândulas écrinas (canal excretor se abre num poro diretamente na superfície da pele) (Gesase e Satoh, 2003). A presença do epitélio apócrino seria resultado da metaplasia ou diferenciação incompleta do epitélio mamário normal (Mazoujian *et al.*, 1983; Petrakis *et al.*, 1993).

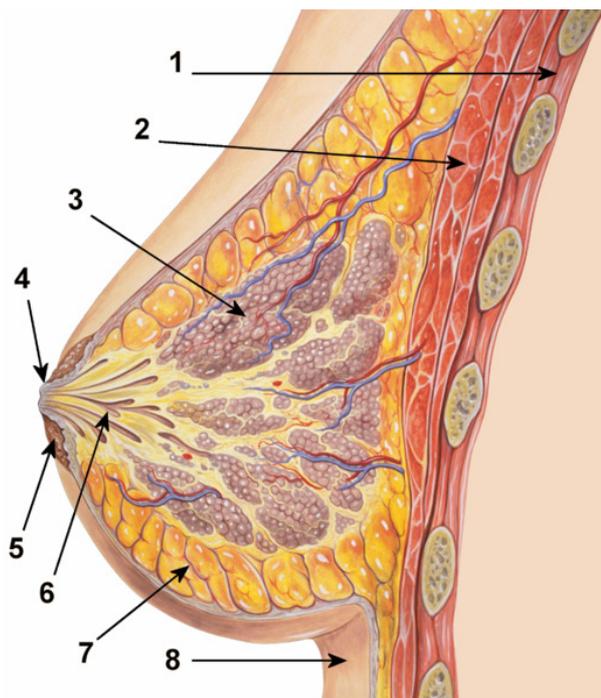


Figura 1.1: Diagrama esquemático da anatomia da mama: 1 - Caixa torácica; 2 - Músculo peitoral; 3 - Lóbulos; 4 - Mamilo; 5 - Aréola; 6 - Ductos; 7 - Tecido adiposo; 8 - Pele. Fonte: Wikipédia. Disponível em: <http://pt.wikipedia.org/wiki/Seios>, consultado em 24 de março de 2010.

O desenvolvimento da glândula mamária ocorre principalmente por influência de esteróides ovarianos e da prolactina, sendo plenamente completado no final da primeira gravidez. Durante o ciclo menstrual ocorrem mudanças hormonais e ele pode ser dividido em 3 fases: a folicular, com início no primeiro dia da menstruação, quando há produção de estrógeno pelo ovário; em seguida, vêm a ovulatória que começa com o pico de estrogênio e ocorre o pico do LH quando ocorre a ovulação; e a fase lútea, que se inicia 14 dias antes da próxima menstruação e é caracterizada pela produção de progesterona (Gordon, 2001). A relação entre as possíveis modificações estruturais da mama e o ciclo menstrual é controversa, sendo alvo de diversos estudos. Alguns autores defendem que, excetuando-se os períodos de gravidez e pós-parto, não haveria verdadeiras modificações do sistema ductal-alveolar, caracterizado por elementos glandulares esparsos, tecidos adiposo e conjuntivo abundantes e lóbulos organizados em pequenos ductos com fundo cego não funcionais. Segundo estas pesquisas, as modificações observadas na segunda fase do ciclo menstrual ocorreriam principalmente devido ao acúmulo anormal de líquido no espaço intersticial e à congestão vascular. No entanto, para outros autores, a mama apresentaria lóbulos quiescentes na fase

folicular do ciclo menstrual (primeira metade). A influência do estrogênio e dos níveis crescentes de progesterona resultariam no aumento do número de alvéolos nos lóbulos, na vacuolização das células e no aumento pronunciado do volume estromal. A diminuição dos níveis de estrogênio e progesterona na menstruação induziria apoptose das células epiteliais, diminuição do edema e regressão do tamanho do lóbulos (Contran, 2000).

Na gravidez, a mama se torna morfológicamente madura, com o desenvolvimento dos alvéolos através da proliferação celular, invasão e diferenciação. Após o parto, cada ducto terminal se torna funcionalmente ativo e origina as glândulas secretoras verdadeiras. Ao fim do período de amamentação, os lóbulos regredem e atrofiam, reduzindo o tamanho total da mama (Franco, 1997; Hondermarck, 2003).

1.3 – Classificação histológica dos carcinomas de mama

Os tipos histológicos mais comuns de carcinoma de mama são divididos em duas categorias: os carcinomas não infiltrantes (*in situ*) e os infiltrantes. Inicialmente, os carcinomas *in situ* foram classificados em ductal ou lobular porque, aparentemente, poderiam afetar estas duas regiões distintamente. Embora estes termos ainda sejam utilizados, atualmente acredita-se que todos os tumores de mama se originam no sistema ductal-alveolar terminal (Contran, 2000).

O carcinoma *in situ* (CIS) caracteriza-se pelo pleomorfismo e coloração escurecida nuclear (aumento da quantidade de cromatina), com alteração da orientação celular do epitélio, mas sem degradação da membrana basal e invasão do estroma. Os CISs abrangem de 15 a 30% dos cânceres de mama, podendo ser classificados em: carcinoma ductal *in situ* (CDIS) e carcinoma lobular *in situ* (CLIS). O CDIS é um tumor formado por uma população de células malignas que se disseminam no sistema ductal, mas que não possuem a capacidade de degradar a membrana basal e invadir o estroma e tecidos distantes (metástases). O CLIS é um tipo raro (1 a 6% dos carcinomas de mama), caracterizado pela proliferação de células em um ou mais ductos terminais ou alvéolos, com a manutenção da arquitetura subjacente e dos lóbulos. Os CISs podem progredir para o tipo invasivo em até 35% dos casos (Franco, 1997; Contran, 2000).

Os carcinomas infiltrantes são os mais frequentes, correspondendo a 70-85% de todos os casos de câncer de mama. Eles podem ser subdivididos em mais de 10 tipos, sendo os mais

comuns o carcinoma ductal infiltrante (CDI) e o carcinoma lobular infiltrante (CLI). Outros tipos, como o medular, o mucinoso, o tubular e o papilar são pouco frequentes, correspondendo a 1-2,4% dos casos infiltrantes (Contran, 2000; Hondermarck, 2003). Aproximadamente 76% dos carcinomas infiltrantes são classificados como CDIs. Eles se caracterizam por aumento do estroma fibroso, conferindo ao tumor uma consistência dura, com nódulos bem delimitados. As células malignas invadem o estroma, formando ninhos, cordões, túbulos ou massas anastomosantes. Elas podem exibir morfologia variada, desde células pequenas com núcleo hipercromático, até células gigantes com núcleos irregulares. Os CLIs compõem apenas 8% dos carcinomas infiltrantes de mama, mas são relevantes, pois tendem a ser bilaterais (20% dos casos), multicêntricos e difusos (não palpáveis). Eles são representados por grupos de alvéolos ocupados por células vacuolizadas sem a orientação celular habitual e núcleos redondos ou ovais sem hipercromatismo (Contran, 2000; Li *et al.*, 2005).

1.4 – Epidemiologia do câncer de mama

Segundo o relatório *World Cancer Report 2008* da Agência Internacional para Pesquisa em Câncer (IARC)/ WHO (*World Health Organization*), apenas no ano de 2008 estimaram 12,4 milhões de casos novos de câncer, com 7,6 milhões de óbitos por câncer no mundo. Dentre os tipos com maior incidência, o câncer de mama apareceu com 1,29 milhões de novos casos, sendo o segundo tipo de câncer mais frequente, e o mais comum entre as mulheres.

De acordo com pesquisa realizada em 2002 (Parkin *et al.*, 2005), o câncer de mama é a principal causa de morte por câncer entre as mulheres de todo o mundo. Ele é também o tipo de câncer mais prevalente nesta população (17,9% dos casos), compreendendo 4,4 milhões de pacientes com sobrevida estimada de até 5 anos. Só no ano de 2002, foram 636 mil novos casos nos países desenvolvidos e 514 mil nos países em desenvolvimento. Na África Sub-Saariana, China e em alguns países do leste asiático, a incidência é muito baixa (20/100.000). Na América do Norte e do Sul, incluindo o Brasil, Austrália e norte e oeste europeus, essas taxas são as mais elevadas (80 a 90/100.000). A cada ano, cerca de 22% dos casos novos de câncer em mulheres são de mama (WHO, 2009). O Instituto Nacional de Câncer (INCA) estimou 49.290 novos casos da doença no Brasil para o ano de 2010, com a segunda maior incidência para o estado do Rio de Janeiro, onde são esperados 7.470 novos casos (INCA, 2009).

1.5 – Fatores de risco e métodos diagnósticos

O risco de desenvolvimento do câncer de mama está relacionado a fatores bem estabelecidos (INCA, 2009; WHO, 2009). A menarca precoce aumenta o risco, pelo maior acúmulo de ciclos menstruais (ovarianos). A cada ano de atraso na menarca, o risco diminui em aproximadamente 15%, aumentando em 3% por ano de atraso na idade da menopausa. A paridade e a idade da primeira gestação (< 30 anos) também são relevantes, uma vez que mulheres que já pariram têm maiores níveis de prolactina e menores quantidades da globulina ligadora de esteróides sexuais quando comparadas a mulheres nulíparas. A lactação também exerce efeito protetor, atribuído à supressão da função ovulatória na amamentação. O risco aumenta em função da duração da terapia de reposição hormonal e mulheres que fazem uso corrente de anticoncepcionais orais têm até 25% mais chances de desenvolver a doença. Lesões benignas, como a doença fibrocística da mama ou fibroadenoma, aumentam em 2 a 3% o risco do câncer. A história familiar e fatores genéticos também estão associados ao maior risco de desenvolvimento de câncer de mama. Mulheres que apresentam mutação nos genes *BRCA1* e *BRCA2* têm 85% de chance de desenvolver a doença antes dos 70 anos de idade. A amamentação, prática de atividade física, alimentação saudável e a manutenção do peso corporal estão associadas a um menor risco (INCA, 2009; WHO, 2009).

A prevenção do câncer de mama ainda é um desafio devido à diversidade dos fatores de risco e às características genéticas relacionadas à sua etiologia. No Brasil, a mamografia é o principal método de rastreamento dos cânceres mamários. Segundo as recomendações do Ministério da Saúde, baseadas no *Documento de Consenso para Controle do Câncer de Mama* de 2004, mulheres a partir dos 40 anos devem realizar anualmente o exame clínico da mama e aquelas entre de 50 a 69 anos devem fazer uma mamografia, no máximo, a cada dois anos. O exame clínico deve ser realizado em todas as mulheres que procuram o serviço de saúde, independente da faixa etária, como parte do atendimento à saúde da mulher. As mulheres de alto risco (com história familiar de câncer de mama em parentes de primeiro grau ou com diagnóstico histopatológico de lesão proliferativa da mama) devem fazer o exame clínico da mama e a mamografia, anualmente, a partir de 35 anos (INCA, 2004).

O câncer de mama, quando diagnosticado em estágios precoces e prontamente tratado, apresenta um bom prognóstico. No entanto, no Brasil as taxas de mortalidade continuam

elevadas, possivelmente devido ao diagnóstico tardio da doença (INCA, 2009). A maioria das mortes por câncer de mama ocorre em consequência de doença disseminada, processo iniciado ainda na fase de promoção do tumor e, portanto, muito antes das técnicas de imagem serem capazes de detectá-lo (Varnum *et al.*, 2003). Em geral, os métodos de detecção e rastreamento de tumores de mama mais utilizados são o exame clínico e a mamografia, procedimentos que falham em diagnosticar cerca de 40% dos casos precoces (Sauter *et al.*, 1997; Sauter *et al.*, 2002).

A mamografia é um método de imagem que utiliza raios-X potencialmente carcinogênicos e que detecta tumores com poucos milímetros, mas que já contêm centenas de milhões de células (Hondermarck, 2003). Além disso, sabe-se que a angiogênese, responsável pela disseminação e metástase do câncer, já está presente em tumores com até dois milímetros de diâmetro. Desta forma, o diagnóstico pela mamografia, mesmo que precoce em termos de imagem, pode ser considerado tardio em relação à história biológica do tumor, constituindo uma importante limitação do método. Adicionalmente, para confirmar a presença de malignidade na mama, a avaliação histológica e/ou citológica deve ser realizada por métodos invasivos, como biópsia ou aspiração com agulha (Mannello *et al.*, 2008).

Na clínica, o grau e o estadiamento do tumor são importantes fatores adicionais para a escolha do tratamento e a predição do prognóstico. A classificação de um câncer nos graus I a III se baseia no nível de diferenciação das células tumorais e no número de mitoses presentes no tumor, variáveis supostamente relacionadas à agressividade da neoplasia. Quanto maior o grau numérico, menor a diferenciação das células (Elston e Ellis, 1991). Embora a gradação seja útil, a correlação entre o aspecto histológico e o comportamento biológico não é precisa. A classificação TNM do *Union Internationale Contre Cancer* (UICC) é o sistema de avaliação dos cânceres mais utilizado no mundo. Ele leva em conta a extensão do tumor primário (T), o comprometimento regional dos linfonodos (N) e a ausência ou presença de metástases (M). O estadiamento da doença neoplásica tem grande importância na escolha da melhor forma de tratamento para o paciente e no seu prognóstico (Contran, 2000; WHO, 2009).

Fatores possivelmente associados ao crescimento e ao potencial metastático dos tumores mamários ou relacionados à sensibilidade e/ou à resistência a agentes terapêuticos são alvo de muitos estudos. No microambiente tumoral, os níveis de estrógeno e de fatores de

crescimento que influenciam a atividade de receptores de estrógeno (ER) são regulados pelas interações entre o tumor e o estroma. A análise deste microambiente pode ajudar a elucidar os mecanismos de comunicação das células neoplásicas com o estroma e prever a resposta individual à terapia endócrina. Um biomarcador de uso recomendado na rotina clínica é o receptor de estrógeno. Quando ausente, pode estar relacionado à agressividade dos tumores (menos diferenciados) e, quando presente, seleciona pacientes que potencialmente respondem ao tratamento com drogas anti-estrógenos, citado por Yamaguchi (2007) (Yamaguchi, 2007; Hondermarck *et al.*, 2008).

Sabemos que os tumores são heterogêneos e que somente a análise histopatológica habitual não é suficiente para a tomada de decisões terapêuticas e prognósticas. Portanto, o desenvolvimento de novas técnicas mais sensíveis e específicas e menos invasivas, capazes de detectar mais precocemente o tumor é muito desejado, pois poderia selecionar melhor o paciente, otimizando os custos e a eficiência dos procedimentos diagnósticos e terapêuticos (Nystrom *et al.*, 2002; Alexander *et al.*, 2004).

1.6 – O fluido aspirado de mama (NAF)

Mais de 95% dos cânceres de mama nascem de elementos epiteliais que revestem o sistema ductal mamário (Schnitt e Guidi, 2004). O fluido mamário (NAF, *Nipple Aspirate Fluid*) é a secreção dos epitélios lobular e ductal de mulheres que estão fora do período de lactação e gravidez. O NAF pode apresentar as cores branca, amarela, verde, marrom, preta e transparente. As colorações escuras podem estar relacionadas ao uso do tabaco e às altas concentrações dos produtos de oxidação do colesterol e lipídeos; as cores mais claras estariam associadas à lactose, citado por Petrakis (1988) (Petrakis *et al.*, 1988; Djuric *et al.*, 2005). O fluido é composto por proteínas (incluindo lactalbumina, imunoglobulinas e fatores de crescimento), glicídios (lactose) e lipídios (ácidos graxos, colesterol e seus produtos de oxidação e hormônios, incluindo estrógenos, andrógenos, progesterona e prolactina) (Petrakis, 1993; Malatesta *et al.*, 2000; Chatterton *et al.*, 2004).

O NAF é um fluido muito interessante para o estudo do câncer de mama, pois é secretado especificamente pelas células do sistema ductal mamário, intimamente relacionadas com o microambiente de desenvolvimento e progressão tumoral, sem a contaminação de outros fluidos intersticiais ou de outros tipos celulares da glândula mamária. Outra vantagem é a sua coleta não invasiva, de baixo custo e que pode ser repetida diversas vezes (Varnum *et*

al., 2003; Alexander *et al.*, 2004; Hondermarck *et al.*, 2008). A mama sendo um órgão pareado também possibilita a análise comparativa das duas mamas da mesma paciente com doença unilateral, onde a mama contralateral sadia funciona como controle negativo (Kuerer *et al.*, 2004). Os níveis proteicos do NAF são relativamente altos (média de 170 mg/ml), tornando possível o uso das técnicas proteômicas mesmo partindo de pequenos volumes de amostra (Harding *et al.*, 2000). Alguns estudos indicam que o NAF pode conter marcadores potenciais para o diagnóstico e prognóstico do câncer de mama, e a proteômica é uma importante ferramenta para a identificação dessas moléculas.

1.7 – Proteômica

O termo proteoma surgiu pela primeira vez na literatura em 1995, sendo definido como o conjunto de proteínas expressas por um organismo, tecido ou célula (Wasinger *et al.*, 1995). Atualmente, a proteômica envolve a identificação, a caracterização (incluindo determinação de modificações pós-traducionais), a quantificação e a localização das proteínas expressas pelo sistema, num dado momento, sob condições determinadas. Além disso, as análises proteômicas também podem incluir o estudo das estruturas, funções e interações de proteínas (Graham *et al.*, 2005). Portanto, didaticamente, a proteômica pode ser dividida em três grandes setores: a) a proteômica dos perfis, que tenta estabelecer o perfil de expressão das proteínas entre diferentes amostras; b) a proteômica funcional, utilizada no estudo das funções de proteínas, suas modificações pós-traducionais e suas interações; c) a proteômica estrutural, que analisa a estrutura terciária de proteínas e de seus complexos (Mitulovic e Mechtler, 2006).

A proteômica pode contribuir de modo importante para o estudo de diversas doenças, aumentando nossa compreensão acerca de seus mecanismos moleculares e identificando novos biomarcadores para diagnóstico e prognóstico, assim como novos alvos terapêuticos (Hanash, 2003). Biomarcadores são moléculas usadas especificamente para medir a progressão de uma doença, uma condição ou um tratamento. Eles podem auxiliar tanto no diagnóstico, monitoramento e prognóstico de doenças, assim como na verificação da eficácia de determinadas intervenções terapêuticas. Fluidos biológicos, como plasma, urina e saliva, vêm sendo muito estudados como fontes de biomarcadores. Normalmente eles são obtidos com mais facilidade, de forma mais barata e menos invasiva do que as amostras de tecido (ex.: biópsias). Entretanto, por apresentar um intervalo dinâmico muito grande,

dificultando a detecção de proteínas menos abundantes, a utilização do plasma humano como amostra destas pesquisas vem sendo fortemente criticada (Good *et al.*, 2007; Lescuyer *et al.*, 2007). Neste contexto, o fluido mamário (NAF) pode constituir uma opção alternativa interessante, permitindo a identificação de marcadores tumorais relevantes, que podem ser úteis na avaliação do risco, diagnóstico, monitoramento do tratamento e detecção de recidivas do câncer, além de ajudar a compreender melhor a sua patogênese (Kuerer *et al.*, 2002).

1.8. Principais técnicas proteômicas

1.8.1. Cromatografia Líquida

Para caracterizar os proteomas, em particular dos fluidos humanos, são necessários métodos robustos, com grande poder de resolução e alta sensibilidade. A cromatografia líquida e a eletroforese bidimensional (2D-PAGE) são duas técnicas proteômicas muito utilizadas para separar proteínas de misturas complexas (Qian *et al.*, 2006). De modo complementar, a espectrometria de massas identifica e quantifica estas proteínas, além de permitir a caracterização detalhada das modificações pós-traducionais presentes (Diamandis, 2004).

A cromatografia líquida (LC) permite a separação de misturas complexas de acordo com diferentes características físico-químicas das moléculas (volume, carga, hidrofobicidade, solubilidade, ligação específica) (Lohaus *et al.*, 2007). Na cromatografia multidimensional, duas colunas cromatográficas são utilizadas em sequência, na maior parte das vezes combinando um trocador catiônico forte com uma coluna de fase reversa. A utilização de colunas cromatográficas longas (até 80 cm), com diâmetros internos pequenos (até 15 µm) e fluxos muito lentos (\approx 100-300 nL/min) melhora a sensibilidade de detecção, aumentando o leque de proteínas identificadas. Para análises quantitativas, podem ser usados diferentes métodos de marcação isotópica das proteínas antes da cromatografia, como o ICAT (*Isotope coded affinity tag*) e o iTRAQ (*isobaric tags relative and absolute quantification*) (Somari *et al.*, 2005; Qian *et al.*, 2006; Lohaus *et al.*, 2007). Como veremos pelos resultados desta dissertação, esta abordagem é bastante promissora para análise de fluidos biológicos humanos, em particular o NAF.

1.8.2. Eletroforese bidimensional

A 2D-PAGE pode resolver milhares de proteínas simultaneamente, produzindo um mapa característico da amostra na condição analisada [44]. A 2D-PAGE em condições desnaturantes foi originalmente desenvolvida por O'Farrell (1975) e é uma das técnicas mais eficientes para resolver misturas protéicas complexas. Ela permite a separação das cadeias polipeptídicas intactas de uma amostra, fornecendo informações sobre as mudanças nos níveis de expressão proteica e sobre a presença de modificações pós-traducionais. Ela também permite o isolamento das proteínas para análises subsequentes por espectrometria de massas (O'Farrell, 1975; Görg *et al.*, 2004).

O alto poder de resolução da eletroforese bidimensional se deve ao fato de que ela utiliza, em sequência, dois métodos de separação baseados em propriedades físico-químicas independentes. A técnica acopla a focalização isoelétrica (IEF) na primeira dimensão, seguida da eletroforese em gel de poliacrilamida (SDS-PAGE) na segunda dimensão. Nestas duas etapas, as proteínas são separadas de acordo com o seu ponto isoelétrico e seu volume molecular, respectivamente (Görg *et al.*, 2004).

Na focalização isoelétrica, as proteínas são separadas em função de sua carga líquida. Como são moléculas anfotéricas, elas podem atuar como ácido ou como base, liberando ou recebendo prótons (íons H^+), respectivamente. Desta forma, dependendo da composição de aminoácidos e do pH em que se encontram, as proteínas apresentarão carga líquida positiva ou negativa. Ao submeter as amostras a uma diferença de potencial num gel de poliacrilamida contendo um gradiente imobilizado de pH (fitas de focalização isoelétrica), os polipeptídeos migram em direção ao pólo de carga oposta à sua. Quando atingem o pH no qual sua carga líquida é zero (ponto isoelétrico), param de migrar, sendo considerados "focalizados" (Lefkovits, 2003).

Na segunda dimensão, os polipeptídios focalizados são equilibrados num tampão contendo SDS (detergente aniônico que carrega negativamente as cadeias polipeptídicas), em condições redutoras, e aplicados no gel de SDS-PAGE, onde são separados de acordo com seus volumes moleculares (Choe *et al.*, 2006). As proteínas podem ser facilmente visualizadas através de métodos de coloração, como o azul de Coomassie, ou pela impregnação por nitrato de prata. No entanto, a 2D-PAGE apresenta uma série de limitações, tais como dificuldade de automação, capacidade limitada de aplicação de amostra no gel, dificuldade de isolar proteínas com extremos de pI ou de volume molecular e capacidade limitada de detectar proteínas pouco abundantes ou muito hidrofóbicas (Lohaus *et al.*, 2007). A introdução de agentes caotrópicos (ex.: tiouréia) e detergentes (ex.:

CHAPS, ASB14) mais poderosos, assim como o advento dos corantes fluorescentes mais sensíveis (ex.: Sypro e CyDyes) contribuíram para avanços importantes desta técnica nos últimos anos (Görg *et al.*, 2004).

1.8.3 - Eletroforese diferencial em gel bidimensional (2D-DIGE)

A eletroforese de fluorescência diferencial em gel bidimensional (DIGE) é um método que permite a análise proteômica múltipla através da utilização de marcadores fluorescentes espectralmente independentes. Esta técnica foi inicialmente desenvolvida por Ünlu *et al.* (1997), onde somente eram utilizados 2 fluoróforos, o Cy3 e o Cy5 (Unlu *et al.*, 1997). Alban *et al.* (2003) descreveram o uso de um terceiro fluoróforo, o Cy2, que é utilizado para marcar o padrão interno (Alban *et al.*, 2003). Este padrão consiste numa amostra composta pela mistura de partes iguais de todas as amostras do estudo e tem como funções facilitar a sobreposição das imagens dos diferentes géis e normalizar as abundâncias relativas dos *spots* proteicos, minimizando a variação experimental. Antes da eletroforese, as duas amostras que se quer comparar são marcadas com fluoróforos (Cy3 e Cy5) e separadas no mesmo gel bidimensional, juntamente com o padrão interno (Cy2) (Figura 1.2). Os fluoróforos possuem massas moleculares e cargas equivalentes e se ligam covalentemente a resíduos de lisina. As cargas positivas originalmente associadas a estes resíduos na amostra protéica são substituídas por igual carga dos grupamentos amino quaternários das moléculas do corante. Desta forma, as proteínas marcadas migram normalmente até o seu ponto isoelétrico e só apresentam uma pequena variação (≈ 450 Da) em relação à sua massa original. A 2D-DIGE, além de reduzir o número de géis produzidos num experimento, simplifica a análise e a comparação entre as amostras, diminuindo a variabilidade experimental e tornando mais confiável o pareamento entre os géis e a quantificação dos *spots* proteicos (Gharbi *et al.*, 2002; Alban *et al.*, 2003). Adicionalmente, o 2D-DIGE é um método de marcação muito sensível (1 ng/spot) e que permite avaliar um grande intervalo dinâmico linear (até cinco ordens de grandeza) (Görg *et al.*, 2004).

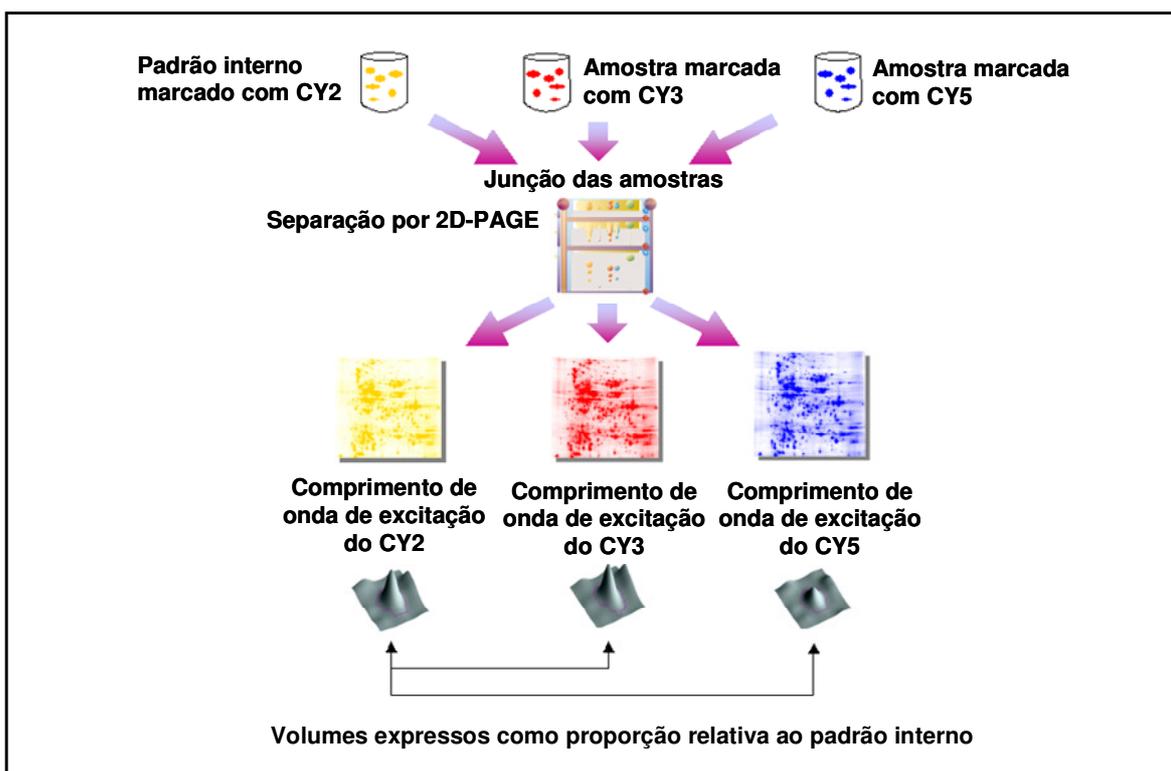


Figura 1.2: Representação esquemática de um experimento utilizando a técnica de fluorescência diferencial em 2D-DIGE. Fonte: Manual Ettan DIGE System 18-1173-17 adaptado por Albuquerque LM, Dissertação (mestrado), 2007.

1.8.4 – Espectrometria de massas

A espectrometria de massas é uma técnica analítica utilizada na identificação de compostos desconhecidos, na quantificação de substâncias conhecidas e na elucidação de propriedades estruturais e químicas das moléculas (Gross, 2004). Com o desenvolvimento de técnicas para volatilização e ionização de biomoléculas, grandes avanços tecnológicos na instrumentação e a disponibilidade crescente de sequências proteicas e nucleotídicas depositadas nos bancos de dados, a espectrometria de massas se tornou uma tecnologia central na proteômica. Trata-se de uma técnica rápida, acurada e extremamente sensível, requerendo pouca quantidade de amostra (femtomoles). Um espectrômetro de massas consiste numa fonte de ionização, um ou mais analisadores de massas que medem a razão massa/carga (m/z) dos analitos ionizados e um detector que registra o número de íons de cada valor de m/z (Aebersold e Mann, 2003).

Nesta dissertação, utilizamos o MALDI-TOF/TOF (*matrix-assisted laser desorption ionization time-of-flight/time of flight*) como principal instrumento para identificação das proteínas do NAF. Ele é um equipamento relativamente simples, sensível e muito utilizado para análise de proteínas. Neste método, a ionização dos digestos proteicos ocorre por dessorção a laser auxiliada por uma matriz, constituída por ácidos orgânicos aromáticos, que absorve a maior parte da energia do laser. A amostra co-cristalizada com a matriz é irradiada por pulsos de laser e o aquecimento dos cristais provoca a sublimação de ambas as moléculas. O processo de ionização não é totalmente compreendido, mas a teoria mais aceita é que a transferência de prótons entre a matriz e o analito ocorra na fase gasosa. As moléculas de amostra ionizadas são então aceleradas por um campo elétrico até a entrada de um tubo de vôo, no qual percorrem uma trajetória definida até o detector de íons. A velocidade de migração de cada íon neste tubo será proporcional à sua relação massa/carga (Aebersold e Mann, 2003; Görg *et al.*, 2004).

A estratégia de PMF (*peptide mass fingerprinting*) identifica proteínas através da determinação das massas do conjunto de seus fragmentos, na maior parte das vezes, peptídeos trípticos (Aebersold e Mann, 2003). Esta abordagem funciona bem desde que a proteína tenha sequência conhecida depositada em bancos de dados (como o NCBI nr e o SwissProt). As massas monoisotópicas de seus peptídeos determinadas experimentalmente são comparadas com as massas teóricas dos peptídeos obtidas pela digestão *in silico* de proteínas depositadas nos bancos de dados. Para isto, vários algoritmos de busca podem ser utilizados, por exemplo o *Mascot* (Perkins *et al.*, 1999) ou *ProFound* (Zhang e Chait, 2000). A confiança na identificação da proteína dependerá da qualidade do espectro obtido no experimento, incluindo o número de peptídeos obtidos, a acurácia dos valores de m/z determinados e a presença de picos contaminantes (Hjerno, 2007).

De modo complementar, a sequência de aminoácidos dos diferentes peptídeos pode ser obtida após sua fragmentação em analisadores sequenciais (em *tandem*), como o TOF/TOF. O primeiro analisador TOF é utilizado para selecionar os íons que podem ser fragmentados, preferencialmente em suas ligações peptídicas, por colisões entre o íon precursor e um gás inerte ocorrendo na célula de colisão, um processo chamado de dissociação induzida por colisão (CID). Também é possível fragmentar o íon precursor sem a presença do gás inerte. Esta é a chamada fragmentação de íons metaestáveis pela técnica de PSD ou *post-source decay* (Aebersold e Mann, 2003; Yates, 2004). O segundo analisador TOF mede a massa dos

íons gerados pela fragmentação do íon precursor selecionado. Os espectros de MS/MS podem ser usados diretamente pelas ferramentas de busca (algoritmos *Mascot*, *Sequest*) e/ou utilizados para determinar a identidade e a ordem dos resíduos de aminoácidos medindo as diferenças (referentes às massas) entre os diferentes íons representantes da fragmentação do íon precursor (sequenciamento *denovo*). As listas de massas são comparadas contra os bancos de dados para identificar os peptídeos fragmentados (Steen e Mann, 2004).

1.9 – Justificativa do trabalho

A maior parte dos estudos proteômicos em câncer de mama analisou amostras de biópsia ou linhagens celulares cancerosas (Somari *et al.*, 2005). Os tecidos mamário e tumoral provenientes de biópsia são heterogêneos, pois contêm uma mistura de diferentes tipos celulares em proporções variáveis, incluindo fibroblastos, células epiteliais, células mioepiteliais, adipócitos, macrófagos, células endoteliais e fibras nervosas. Devido à vascularização dos tumores, o tecido tumoral contém também componentes do sangue, o que complica ainda mais a análise proteômica deste tipo de amostra (Hondermarck, 2003). Em estudos quantitativos de expressão diferencial, toda esta heterogeneidade tecidual dos carcinomas se mostrou um problema importante. Atualmente, a técnica de microdissecção a laser vem sendo cada vez mais empregada para selecionar populações celulares mais homogêneas (Wulfkühle *et al.*, 2002).

Especificamente quanto à análise proteômica do NAF, poucos trabalhos foram publicados. Utilizando 2D-PAGE tradicional e PMF (*Peptide Mass Fingerprinting*), Alexander *et al.* (2004), identificaram 41 proteínas distintas, duas delas (GCDFP-15 e glicoproteína α -1 ácida) possivelmente relacionadas com câncer de mama. Kuerer *et al.* (2002) observaram diferenças qualitativas nos padrões de expressão entre as duas mamas de 3 pacientes com carcinoma unilateral invasivo por 2D-PAGE. Numa avaliação do NAF de mulheres sadias por cromatografia líquida seguida de espectrometria de massas, Varnum *et al.* (2003) identificaram 64 proteínas, 15 das quais haviam sido previamente reportadas como alteradas em amostras de câncer de mama. Em sua maioria, os trabalhos empregaram a técnica de SELDI para detectar a expressão diferenciada de picos proteicos (não identificados) nos espectros de NAF de mulheres com câncer e mulheres normais (Coombes *et al.*, 2003; Pawlik *et al.*, 2005; Sauter, 2005). Esta abordagem tem recebido críticas em

função da baixa reprodutibilidade e da impossibilidade de identificação das proteínas detectadas como diferenciais.

A relação entre as variações na composição do NAF e o câncer de mama ainda é pouco compreendida (Djuric *et al.*, 2005). O fluido aspirado de mama pode vir a se tornar, junto com o exame clínico e os exames de imagem da mama, uma importante ferramenta no auxílio do diagnóstico precoce do câncer de mama. A utilização das técnicas proteômicas no estudo do NAF é fundamental para compreender melhor as características deste fluido e para detectar e identificar candidatos a biomarcadores. Estes candidatos podem ser muito úteis para diagnosticar precocemente a doença, identificar novos fatores prognósticos e auxiliar na escolha e no monitoramento do tratamento.

2 – OBJETIVOS

Objetivo geral

Este trabalho teve como objetivo geral a caracterização do conteúdo proteico de amostras de NAF de pacientes brasileiras de câncer de mama unilateral, utilizando técnicas de análise proteômica. Este conhecimento será fundamental para que, eventualmente, possamos identificar biomarcadores para o diagnóstico e/ou prognóstico da doença, assim como para o monitoramento de seu tratamento.

Objetivos específicos:

- Padronização do método de coleta do NAF;
- Otimização das técnicas proteômicas para análise do NAF, tais como métodos de depleção de proteínas mais abundantes e eletroforese bidimensional;
- Determinação do perfil proteico de NAF de ambas as mamas da paciente utilizando DIGE, comparando a mama afetada com a não afetada;
- Análise do conteúdo de metaloproteases, glico- e fosfo-proteínas do NAF de pacientes com câncer de mama unilateral.

3 - MATERIAL E MÉTODOS

3.1. Coleta das amostras de fluido de mama (NAF)

As amostras foram colhidas das duas mamas de pacientes com câncer de mama unilateral atendidas no Serviço de Mastologia do Instituto Fernandes Figueira, Fiocruz/RJ e no Ambulatório de Ginecologia do Hospital da Lagoa (Ministério da Saúde). As pacientes convidadas a participar do estudo apresentavam carcinoma ductal ou lobular infiltrante unilateral e não foram previamente submetidas à qualquer cirurgia mamária ou quimioterapia neo-adjuvante. O diagnóstico foi confirmado por biopsia percutânea com agulha grossa (Parker *et al.*, 1993). Para padronização inicial das condições experimentais, também foram coletadas amostras de NAF de mamas com tumor benigno e de mamas saudáveis de mulheres voluntárias não relacionadas às pacientes e também atendidas nos ambulatórios citados acima. Todas as mulheres assinaram um termo de consentimento informado (Anexo I). Este projeto de pesquisa foi aprovado pela CONEP (Comissão Nacional de Ética em Pesquisa), licença nº. 13498, em 19/12/06 (Anexo II).

O procedimento de coleta de NAF das voluntárias foi realizado após desinfecção do mamilo com álcool a 70% (v/v), seguida por 5 minutos (min) de massagem da mama, na direção parede torácica - mamilo. Após a aplicação de compressa morna por 5 min, a mama foi massageada mais uma vez e submetida à aspiração do fluido com bomba de aspirar leite, comumente utilizada em lactantes, até a visualização das gotículas de fluido, adaptado por Pawlik *et al.* (2005) (Pawlik *et al.*, 2005). O NAF foi coletado com micropipeta automática e o volume de cada amostra foi registrado. Imediatamente após a coleta, as amostras foram diluídas 10 vezes em solução salina tamponada com fosfato (PBS pH 7,4) e centrifugadas a 250 x g, por 10 min, a 4 °C. O sobrenadante do NAF foi coletado e armazenado a -80 °C.

3.2. Dosagem de proteínas pelo ácido bicinchonínico (BCA)

A concentração proteica das amostras foi determinada pelo método do BCA (Smith *et al.*, 1985), utilizando o kit *Bicinchoninic Acid Protein Assay Kit* (Sigma-Aldrich, EUA). Antes da dosagem, as amostras foram centrifugadas a 16.000 x g, por 10 min, a 4 °C. A curva padrão de albumina sérica bovina (BSA, 1 mg/mL) foi preparada no intervalo de 10 a 50 µg de BSA, em um volume final de 100 µL de PBS. As amostras de NAF foram preparadas no mesmo volume final. Em cada um dos tubos, foram adicionados 2 mL do reagente C, preparado com 50 partes do reagente A (ácido bicinchonínico e tartarato em tampão carbonato alcalino) e

uma parte do reagente B (sulfato de cobre). Após 30 min de incubação a 37 °C, a absorvância foi medida a 562 nm na leitora de placas *VersaMax Microplate Reader* (Molecular Devices, EUA). Para comparar as médias das concentrações entre grupos controle e câncer, foi realizado o teste T de Student não pareado utilizando o programa *GraphPad InStat 3.06*. A diferença entre as médias dos grupos foi considerada estatisticamente significativa quando $p < 0,05$ (intervalo de confiança de 95%).

3.3. Precipitação de proteínas pelo ácido tricloroacético (TCA)

Amostras de NAF foram precipitadas com TCA, de acordo com o protocolo de Stone e Williams (1993) modificado (Stone e Williams, 1993). Para cada 1 mL de amostra, foram adicionados 278 μ L de TCA 50% e 139 μ L de Triton X-100 1% (v/v). A mistura foi agitada em vórtex e incubada em gelo por até 2 horas. Em seguida, as amostras foram centrifugadas a 20.800 x g, por 10 min, a 4 °C, e o sobrenadante foi descartado. O precipitado foi então lavado duas vezes com 500 μ L de acetona 90% (v/v) gelada e deixado ao ar livre até secar.

3.4. Eletroforese unidimensional

O NAF foi analisado por eletroforese em gel de poliacrilamida em presença de SDS (SDS-PAGE) e β -mercaptoetanol, segundo a metodologia descrita por Laemmli (1970) (Laemmli, 1970). O gel de empilhamento foi feito com acrilamida a 4 %T/ 2,67 %C em tampão tris-HCl 0,5 M pH 6,8, contendo 0,4% de SDS, e o gel de separação foi preparado com acrilamida a 12 %T/ 2,67 %C em tampão tris-HCl 1,5 M pH 8,8, contendo 0,4% de SDS. O tampão de eletrodo utilizado foi tris-glicina (0,025 M de tris base e 0,192 M de glicina) pH 8,3, contendo 0,1% de SDS. As amostras [normalizadas por volume (20 μ L) ou por quantidade de proteína (20 μ g)] foram misturadas ao tampão de amostra, nas seguintes concentrações finais: 0,06 M de tris-HCl pH 6,8, 2% de SDS, 5% de β -mercaptoetanol, 10% de glicerol e 0,025% de azul de bromofenol. Os padrões de massa molecular utilizados foram fosforilase b (97 kDa), albumina sérica (66 kDa), ovalbumina (45 kDa), anidrase carbônica (30 kDa), inibidor trípico de soja (20,1 kDa) e α -lactalbumina (14,4 kDa)(GE Healthcare, EUA). Utilizamos o sistema *Mini-Protean II* da BioRad (EUA) e a corrida foi realizada a 10 mA por gel, por 30 min, seguido de 15 mA por gel até que a frente de migração atingisse o final do gel. Os géis foram então fixados e revelados por um dos métodos descritos no item 3.8.

3.5. Zimografia

Para avaliar a atividade proteolítica do NAF, foram feitos géis de zimografia utilizando como substrato gelatina a 0,1% (Sigma, EUA) co-polimerizada com gel de poliacrilamida a 12 %T, contendo 0,1% de SDS (sistema *Mini-Protean II*, BioRad, EUA). O gel de empilhamento, tampão de eletrodo e padrões de massa molecular utilizados estão descritos no item 3.4. As amostras de NAF (20 µg) foram misturadas com tampão de amostra como descrito acima, sem agente redutor. As condições elétricas utilizadas foram: 10 mA por gel por 30 min, seguido de 20 mA por gel até o final da corrida. Para renaturar as proteínas, o SDS foi removido incubando-se os géis com tampão tris-HCl 50 mM pH 7,5, contendo Triton X-100 2,5%, por 2 horas, sob agitação moderada. Em seguida, os géis foram incubados no tampão de lise tris-HCl 50 mM pH 7,5, CaCl₂ 5 mM, NaCl 15 mM, por 16-20 horas, a 37 °C, com agitação branda. Para determinar a(s) classe(s) de proteases presentes no NAF, após a incubação no tampão contendo Triton X-100, os géis foram incubados no tampão de lise contendo os seguintes inibidores: EDTA (10 mM), ortofenantrolina (10 mM), E-64 (10 µM), APMSF (100 µM) ou pepstatina (1 µM). Apenas para pepstatina, o ensaio foi feito com tampão titulado até pH 5,0 (Beynon, 1996). Os géis foram revelados por Coomassie R-250, como descrito no item 3.8.1. A digestão da gelatina foi identificada por zonas claras de hidrólise nos géis contra o fundo escuro do corante (Heussen e Dowdle, 1980; Mannello e Sebastiani, 2003). Os géis foram fotografados no Setor de Produção e Tratamento de Imagens do IOC/Fiocruz.

3.6. Depleção das proteínas mais abundantes do NAF

3.6.1. Cromatografia de afinidade usando anticorpos policlonais monoespecíficos

Para remoção das proteínas mais abundantes do NAF, verificamos inicialmente a possibilidade de empregar uma coluna de múltipla afinidade desenvolvida para remoção das seis proteínas mais abundantes do plasma (albumina, IgG, IgA, transferrina, haptoglobina e anti-tripsina)(*Multiple Affinity Removal Column* Hu-6, 4,6 X 50 mm, Agilent Technologies, Santa Clara, CA). Uma alíquota (100 µL) de NAF contendo 160 µg de proteína foi submetida à cromatografia de afinidade no sistema *ÄKTA Purifier* da GE Healthcare, de acordo com o protocolo do fabricante. Resumidamente, um tampão neutro contendo sal (tampão A #5185-5987, Agilent) foi utilizado tanto para o equilíbrio inicial da coluna quanto para a

eluição das proteínas não ligadas, ao fluxo de 0,25 mL/min. As proteínas que interagiram com a coluna foram eluídas com tampão ácido contendo uréia (tampão B #5185-5988, Agilent), a 1 mL/min. Foram coletadas frações de 0,25 mL por tubo e o perfil cromatográfico foi obtido através da leitura de absorvância a 280 nm. A fração do NAF eluída da coluna com o tampão B foi analisada por eletroforese bidimensional (item 3.7) revelada com Coomassie G250 (item 3.8.2). Os *spots* detectados foram identificados por espectrometria de massas como descrito no item 3.10.

3.6.2. Cromatografia de afinidade usando Proteína L agarose

Testamos também a remoção apenas das imunoglobulinas do NAF usando gel de agarose contendo proteína L recombinante (Pierce, EUA) (Akerstrom e Bjorck, 1989). Esta proteína da bactéria *Peptostreptococcus magnus* interage com as cadeias leves do tipo *kappa* de todas as classes de imunoglobulina, incluindo IgG, IgM, IgA, IgE e IgD. Alíquotas de NAF (360 µg de proteína/100 µL) foram aplicadas em colunas do tipo *Micro-Spin* contendo 100 µL de gel equilibrado com tampão A (fosfato de sódio 100 mM, cloreto de sódio 150 mM, pH 7,2). Após incubação por 10 min, à temperatura ambiente, o gel foi lavado com 600 µL de tampão A e as proteínas ligadas à coluna foram eluídas com 600 µL de tampão B (glicina-HCl 100 mM pH 2,7). Tanto para aplicação da amostra quanto para lavagem e eluição da coluna, utilizamos centrifugação a 100 x g. O pH das frações eluídas com tampão B foi neutralizado pela adição de 10 µL do tampão C (Tris 1M pH 9). Todas as frações tiveram seu conteúdo proteico determinado por leitura de absorvância a 280 nm no espectrofotômetro *DU 650* (Beckman, EUA). Todas as frações ligadas e não ligadas à proteína L agarose foram submetidas à eletroforese unidimensional (item 3.4) e reveladas por Coomassie G250 (item 3.8.2). As bandas reveladas na fração ligada foram excisadas, tripsinizadas e analisadas por espectrometria de massas como descrito no item 3.10. A fração do NAF que não interagiu com a coluna foi analisada por eletroforese bidimensional utilizando fitas de focalização isoeétrica de 7 cm, com intervalo de pH 4-7 e aplicação por reidratação *in gel* (item 3.7).

3.7. Eletroforese bidimensional convencional (2D-PAGE)

Para focalização isoeétrica, foram utilizadas fitas de 7 ou 18 cm, com gradiente imobilizado de pH (IPG) nos intervalos de 3-10 (não-linear/NL) ou 4-7 (GE Healthcare). As quantidades de proteína aplicadas no gel variaram em cada experimento e estão especificadas nas legendas

das figuras. A focalização isoeétrica foi feita no equipamento *IPGPhor* (GE Healthcare) e os seguintes métodos de aplicação de amostra foram testados: a) *Reidratação in gel*: amostras precipitadas com TCA (item 3.3) foram solubilizadas por 12 horas com solução de reidratação composta de 8 M de uréia, 4 % de CHAPS, 100 mM de ditioneitol (DTT), 0,002% de azul de bromofenol e 1% (v/v) de *IPG buffer* no intervalo de pH igual ao da fita. Após solubilização e centrifugação a 16.000 x *g*, por 10 min, as amostras foram aplicadas no sarcófago de cerâmica e a fita foi posicionada com o gel virado para baixo, em contato direto com a amostra. As fitas de focalização foram cobertas com 3 mL de óleo mineral para evitar evaporação e cristalização de uréia. As condições elétricas de corrida utilizadas foram as recomendadas por (Görg, 2007): 30 volts (V) por 12 horas, 200 V por 1 hora, 500 V por 1 hora, 1.000 V por 1 hora, de 1.000 a 3.500 V em 30 min e 3.500 V até alcançar 16.000 VhS (para fitas com pH 3-10 NL) ou 21.000 VhS (para aquelas com pH 4-7), considerando fitas de focalização de 7 cm. As condições de corrida para fitas de 18 cm (pH 4-7) foram as mesmas, com exceção dos dois últimos passos: 1.000 a 8.000 V em 30 min e 8.000 V até alcançar 48000 VhS. b) *Cup loading*: reidratamos separadamente a fita de focalização isoeétrica, por 12 horas, na mesma solução de reidratação descrita acima e solubilizamos as amostras precipitadas com TCA em um volume máximo de 100 µL, também com a mesma solução, porém com 9,5 M de uréia. Após a solubilização, as fitas foram posicionadas no sarcófago de *cup loading* com o gel virado para cima. Pequenos pedaços de papel de filtro grosso foram embebidos em água destilada e posicionados sobre as extremidades de cada fita, sobre os quais os eletrodos foram posicionados. O reservatório de amostra (*cup*) foi posicionado no anodo e a amostra foi nele aplicada após centrifugação a 16.000 x *g*, por 10 min. Cada fita de focalização foi coberta com 5 mL de óleo mineral e a amostra contida no reservatório foi também coberta com 20 µL de óleo. As condições elétricas de corrida utilizadas estão descritas acima, exceto o primeiro passo de reidratação a 30 V, por 12 horas, que não foi realizado previamente à aplicação da amostra.

Antes da segunda dimensão (SDS-PAGE), as fitas de IPG foram incubadas, por 15 min, com 100 mg de ditioneitol (DTT) em 10 mL de tampão de equilíbrio (50 mM tris-HCl pH 8,8, 6 M uréia, 30% (v/v) glicerol, 2% dodecil-sulfato de sódio (SDS) e 0,002% azul de bromofenol), seguido de incubação com 400 mg de iodoacetamida em outros 10 mL de tampão de equilíbrio, também por 15 min, com agitação constante. As fitas foram posicionadas sobre o gel de poliacrilamida a 12 %T/ 2,67 %C (item 3.4) e imobilizadas com agarose a 0,5%

preparada em tampão de eletrodo, a 80 °C. Para a corrida de segunda dimensão das fitas de IPG de 7 cm, utilizamos o sistema *Mini-Protean II* da BioRad (USA). A corrida foi realizada a 2,5 mA por gel, por 30 min, seguido 15 mA por gel até o final. Para as fitas de 18 cm, foi utilizado o sistema *Ettan DaltSix* (GE Healthcare) e as seguintes condições elétricas: 2,5 watts (W) por gel, por 30 min, e 100 W totais até o final. O tampão de eletrodo foi o mesmo descrito no item 3.4, porém duas vezes concentrado apenas no compartimento superior da cuba de eletroforese do *DaltSix*. Os padrões de massa molecular (GE Healthcare) foram fervidos por 5 min, misturados 1:1 (v/v) com agarose a 1% e aplicados em um pedaço de papel de filtro posicionado ao lado da fita (Görg *et al.*, 2004). Como especificado nas legendas das figuras, os diferentes géis foram revelados por um dos métodos descritos no item 3.8.

3.8. Revelação dos géis

3.8.1. Revelação dos géis por Coomassie R-250

Após a eletroforese, os géis foram fixados por 12 horas em solução aquosa de etanol 40% (v/v) e ácido acético 10% (v/v). Os géis foram então lavados por mais 30 min em solução de fixação fresca e incubados em Coomassie blue R-250 0,2% em solução de fixação, por 3 h. Os géis foram descorados utilizando a solução fixadora por 2-3 horas, seguido de incubação em água por 24 a 72 horas para completar a descoloração Fazekas de St Groth (1963) modificado (Fazekas de St Groth *et al.*, 1963). Todos os passos foram realizados sob agitação moderada. As imagens foram digitalizadas no scanner *ImageScanner III*, utilizando o programa *Labscan* (GE Healthcare).

3.8.2. Revelação dos géis por Coomassie coloidal G-250

Após eletroforese, os géis foram fixados por 30 min em solução de etanol 30% (v/v) e ácido fosfórico 1,7% (v/v). A solução foi descartada e a fixação repetida por mais duas vezes. Os géis foram lavados 3 vezes por 20 min com solução de ácido fosfórico 1,7% (v/v). No passo seguinte, incubamos os géis, por 30 min, em solução composta de ácido fosfórico 1,7% (v/v), etanol 18% (v/v), sulfato de amônio 15%, a qual adicionamos uma solução aquosa de 2% de Coomassie Blue G-250 (Sigma, USA), em volume suficiente para uma concentração final de 0,02% de corante. O ponto final da coloração foi alcançado após 24 a 72 horas de incubação

(Rabilloud e Charmont, 2000). Todos os passos foram realizados sob agitação moderada. Os géis foram escaneados com o *ImageScanner III*, utilizando o programa *Labscan* e as imagens analisadas no programa *Image Master 7.0* (GE Healthcare).

3.8.3. Impregnação dos géis por nitrato de prata

A revelação dos géis por nitrato de prata foi realizada de acordo com o método de Heukeshoven e Dernick (1985), com modificações (Heukeshoven e Dernick, 1985). Os géis foram fixados por 12 horas em solução de etanol 40% (v/v) e ácido acético 10% (v/v). A solução foi desprezada e a fixação foi repetida em um novo passo de 30 min. Em seguida, sensibilizamos os géis em solução de glutaraldeído 0,125% (v/v), etanol 30% (v/v), tiosulfato de sódio 0,2% e acetato de sódio triidratado 12%, por 30 min. Os géis foram então lavados rapidamente três vezes com água e incubados por 20 min com solução de nitrato de prata 0,25%, contendo formaldeído 0,015% (v/v). Foram realizadas mais duas breves lavagens com água, seguidas da revelação dos géis por carbonato de sódio 2,5% e formaldeído 0,007%, por 2 a 5 min. A solução reveladora foi descartada, o gel foi incubado por 10 min em uma solução de EDTA a 1,5% para deter a reação e os géis foram lavados três vezes por 5 min com água. Todos os passos foram realizados sob agitação moderada. As imagens foram digitalizadas no scanner *ImageScanner III*, utilizando o programa *Labscan* (GE Healthcare).

3.8.4. Detecção específica de glico- e fosfo-proteínas

Amostras pareadas de mamas doentes e sadias foram submetidas à 2D-PAGE em fitas de focalização isoeétrica de 7 cm, com intervalo de pH 4-7 e aplicação por reidratação *in gel* (item 3.7). Para visualizar o conteúdo de glicoproteínas do NAF, os géis foram revelados com *Pro-Q Emerald 488 Glycoprotein Gel and Blot Stain Kit* (Invitrogen, USA), segundo o protocolo do fabricante. Resumidamente, os géis foram fixados 2 vezes em solução de metanol 50% e ácido acético 5% por 60 min, lavados três vezes por 20 min com ácido acético 3% e incubados com 25 mL de solução oxidante (ácido periódico dissolvido em 3% de ácido acético) por 20 min. Os géis foram lavados três vezes por 20 min com ácido acético 3%, incubados com 25 mL de solução reveladora *Pro-Q Emerald* por 2 horas no escuro e lavados três vezes por 30 min com ácido acético 3%. As imagens foram obtidas no scanner *Typhoon Trio* da GE Healthcare, utilizando comprimentos de onda de excitação e emissão de 488 nm e 520 nm, respectivamente. Para visualizar o conteúdo de fosfoproteínas do NAF, os géis

foram revelados com o kit *Pro-Q Diamond Phosphoprotein Gel Stain* (Invitrogen, USA), também seguindo o protocolo do fabricante. Os géis foram fixados 2 vezes em solução de metanol 50% e ácido acético 10% por 30 min, lavados três vezes por 10 min com água e incubados com 40 mL de solução reveladora *Pro-Q Diamond* por 90 min no escuro. Em seguida, os géis foram descorados com acetonitrila 20%, acetato de sódio 50 mM pH 4,0 por 30 min no escuro e lavados duas vezes por 5 min com água. As imagens dos géis foram digitalizadas no scanner *Typhoon Trio* (GE Healthcare) utilizando os comprimentos de onda de excitação e emissão de 532 nm e 555 nm, respectivamente. Após digitalização das imagens de fluorescência, os géis foram revelados por Coomassie coloidal G-250 como descrito no item 3.8.2. As imagens das revelações por *Pro-Q* e por Coomassie foram sobrepostas utilizando o programa *Image Master 7.0* (GE Healthcare). Os *spots* comuns às duas imagens foram excisados para posterior tripsinização e identificação por espectrometria de massas (item 3.10).

3.9. Eletroforese de fluorescência diferencial (2D-DIGE)

3.9.1. Dosagem de proteínas pelo 2D-Quant kit

Antes da focalização isoeétrica, as amostras precipitadas com TCA (item 3.3) foram solubilizadas por 12 horas em 8 M de uréia, 4% de CHAPS e 0,002% de azul de bromofenol, com o pH ajustado para 8,5. Uma alíquota de cada amostra ressolubilizada foi dosada com o *2-D Quant Kit* da GE Healthcare, seguindo o protocolo do fabricante. Em resumo, foi preparada a curva padrão de albumina sérica bovina (BSA, 2 mg/mL) no intervalo de 10 a 50 µg de BSA. As amostras de NAF foram preparadas no mesmo volume (8 µL) e misturadas com 500 µL do precipitante e 500 µL do co-precipitante. As amostras foram centrifugadas a 14.000 x *g* por 10 min e o sobrenadante foi descartado para centrifugar mais uma vez a 16.000 x *g* por 5 min. O sobrenadante restante foi descartado e adicionamos ao *pellet* 100 µL de solução de cobre e 400 µL de água. Adicionamos, por tubo, 1 mL de reagente de trabalho, preparado com 100 partes do reagente A e uma parte do reagente B. Após 15 min de incubação à temperatura ambiente, a absorvância foi medida a 480 nm na leitora de placas *VersaMax Microplate Reader* (Molecular Devices, EUA). Por fim, cada amostra foi levada à concentração de 2,5 mg/mL antes da marcação com os *Cy dyes*.

3.9.2. Análise por 2D-DIGE

Em condições otimizadas de eletroforese bidimensional, as amostras pareadas de mamas normais e de mamas com câncer (Tabela 4.6) foram marcadas alternadamente com Cy3 ou Cy5 e um *pool* interno composto por quantidades equivalentes de todas as amostras foi marcado com Cy2, de forma a permitir a análise quantitativa por fluorescência diferencial (GE Healthcare) (Marouga *et al.*, 2005).

Cada *CyDye* (5 nmol) foi reconstituído em 5 μL de dimetilformamida (DMF) anidra (solução estoque = 1 nmol/ μL) misturando rapidamente no vórtex e centrifugando por 30 seg. A solução estoque foi diluída com DMF para a concentração final de 400 pmol/ μL . A cada amostra contendo 50 μg de proteínas (39,2 μL) foram adicionados 400 pmoles (1 μL) de *CyDye*, misturando imediatamente no vórtex, centrifugando por 30 seg e incubando por 30 min no gelo, no escuro. Para interromper a reação, adicionou-se 1 μL de solução de lisina 10 mM em cada amostra, misturando novamente no vórtex e centrifugando por 30 seg. As amostras foram incubadas por mais 15 min no gelo, no escuro. As amostras marcadas foram misturadas 1:1 (v/v) na solução de reidratação contendo 8 M uréia, 4% CHAPS, 1% (v/v) anfólitos pH 4-7, 100 mM DTT e 0,002% BFB e incubadas no gelo, por 10 min. As amostras marcadas com os 3 *Cy Dyes* foram centrifugadas, por 15 min, a 16.000 x *g* e volumes iguais de cada uma foram misturados e centrifugados, por 1 min, a 16.000 x *g*. Antes de aplicar no sarcófago, adicionamos volume de solução de reidratação suficiente para completar os 350 μL necessários para reidratar a fita de focalização de 18 cm, segundo o método de *reidratação in gel* descrito no item 3.7. As condições de corrida da primeira e da segunda dimensões também estão descritas no mesmo item.

3.9.3. Digitalização e análise das imagens

As imagens dos géis foram digitalizadas no scanner *Typhoon Trio* e analisadas pelo software *DeCyder* versão 5.0, da GE Healthcare. Os géis de DIGE foram polimerizados em placas de vidro de baixa fluorescência e escaneados utilizando os seguintes comprimentos de onda de excitação (Ex) e emissão (Em): Cy2 (Ex488 / Em520), Cy3 (Ex532 / Em580), Cy5 (Ex633 / Em670). Os géis de DIGE foram posteriormente revelados por Coomassie coloidal G-250 (item 3.8.2) para permitir a excisão dos *spots* de interesse e posterior identificação por espectrometria de massas (item 3.10).

As imagens dos géis marcados com *CyDyes* foram analisadas no software *DeCyder*, módulos *Differential In-gel Analysis* (DIA) e *Biological Variation Analysis* (BVA). No DIA foram feitas a detecção e a quantificação *spots* e a normalização de seus volumes. Para normalização, o *software* utiliza a razão logarítmica entre o volume de cada *spot* proteico e seu correspondente do padrão interno aplicado no mesmo gel. O módulo BVA faz o pareamento entre todos os géis, compara o volume normalizado dos *spots* de todas as imagens e faz a análise estatística.

Os géis corados com Coomassie coloidal G-250 foram digitalizados no scanner *ImageScanner III* (GE Healthcare).

3.10. Identificação de proteínas por espectrometria de massas

3.10.1. Digestão trípica em gel de poliacrilamida

Os *spots* de interesse revelados nos diferentes géis bidimensionais foram selecionados e submetidos à digestão trípica em gel, de acordo com Shevchenko *et al* (1996) com modificações (Shevchenko *et al.*, 1996). Primeiramente, os *spots* foram cortados com bisturis número 15 e divididos em pedaços menores. Os pedaços de gel foram descorados em 200 µL de solução de acetonitrila 50% (v/v) em bicarbonato de amônio 25 mM pH 8,0, por 12 horas. A descoloração foi repetida mais duas vezes por 15 min, agitando os pedaços de gel no vórtex. Os géis foram reduzidos com ditioneitol (DTT) 65 mM, por 30 min, a 56 °C e alquilados com iodoacetamida (IAA) 200 mM, por 30 min, à temperatura ambiente. Em seguida, foram lavados em bicarbonato de amônio 100 mM e desidratados com acetonitrila 100%. Este passo foi repetido mais uma vez e os pedaços de gel foram secos no *Speed Vac*, por 15 min. Adicionaram-se 15 µL de solução de digestão contendo tripsina (20 ng/µL em bicarbonato de amônio 40 mM pH 8,0) da Promega (EUA). Os géis foram mantidos no gelo por 45 min para que a tripsina penetrasse sem que a digestão fosse iniciada. O excesso de tripsina foi removido e foram adicionados 20 µL de bicarbonato de amônio 40 mM, deixando por 16 a 24 horas, a 37 °C. Os pedaços de gel contendo os peptídeos trípticos foram sonicados em ultrassom *Ultra Cleaner 1400* (Unique, Brasil), por 10 min, agitados por 20 seg no vórtex e os 20 µL do sobrenadante foram removidos para um novo microtubo. Adicionaram-se ao tubo de gel restante mais 30 µL de solução de ácido fórmico 5% (v/v) e acetonitrila 50% (v/v), seguido de agitação por 20 seg no vórtex, incubação por 15 min à

temperatura ambiente, sonicacão por mais 2 min e agitação por 20 seg adicionais no vórtex. A solução contendo os peptídeos foi removida e adicionada aos 20 µL anteriores. A extração dos peptídeos foi repetida mais uma vez e todos os extratos reunidos no mesmo microtubo foram concentrados até 5 µL no *Speed Vac*. Os peptídeos tripticos gerados foram dessalinizados e concentrados por micro-colunas de fase reversa (*Zip-Tip C18*, Millipore, USA), de acordo com o protocolo do fabricante, para posterior análise por espectrometria de massas.

3.10.2. Análise por MALDI-TOF/TOF MS

Os peptídeos tripsinizados e dessalinizados foram analisados no espectrômetro de massas do tipo MALDI-TOF/TOF (*4700 Proteomics Analyzer*, Applied Biosystems, EUA) no modo *reflectron* positivo. As amostras foram aplicadas na placa pelo método gota sobre gota em uma proporção de 1:1 (v/v) com a matriz, composta de ácido α -ciano-4-hidroxicinâmico a 10 mg/mL em acetonitrila 50% (v/v) e ácido trifluoacético 0,3% (v/v), e secaram à temperatura ambiente. A placa de MALDI foi calibrada para MS com o kit de padrões de massa para o analisador *4700 Proteomics Analyzer* (Applied Biosystems, USA), contendo os seguintes peptídeos: arg-bradicinina (m/z 904,468), angiotensina I (m/z 1.296,685), glu¹-fibrinopeptídeo B (m/z 1.570,677), ACTH-(1-17) (m/z 2.093,087) e ACTH-(18-39) (m/z 2.465,199). Para a calibração do MS/MS, foi utilizado o espectro de fragmentação da angiotensina I do kit de padrões de massa.

No modo MS, os espectros de massa foram obtidos selecionando uma faixa de 900 a 4.000 Da, com um foco em 2.400 Da, sendo cada espectro resultado de 2.000 tiros de laser randômicos e uniformes. A intensidade do laser foi ajustada entre 9×10^3 a 1×10^4 e a voltagem final do detector foi 1,881 kV (quilovolt). No modo MS/MS, os dez íons mais abundantes (relação sinal/ruído acima de 20) observados no espectro de MS foram selecionados e fragmentados por PSD (*post source decay*). Os espectros de MS/MS foram obtidos selecionando uma faixa de 60 Da até a massa do íon precursor, sendo cada espectro correspondente a 1.000 tiros randômicos e uniformes. A voltagem final do detector foi 2,2 kV. A lista de massas dos fragmentos foi gerada segundo os seguintes parâmetros: relação sinal/ruído acima de 2; densidade máxima de 55 picos por intervalo de 200 Da; área mínima de 10 e número máximo de 200 picos por íon precursor. A identificação das proteínas foi feita pelo algoritmo Mascot (*Matrix Science*, Inglaterra) no modo *Ion Search*. Os parâmetros

de busca utilizados foram: duas clivagens trípticas perdidas, íons com uma única carga, $\pm 0,6$ Da de erro no modo MS, $\pm 0,2$ Da de erro no modo MS/MS, modificação fixa carbamidometilação (Cys), e as seguintes modificações variáveis: oxidação (Met, Trp, His) e propionamida (Cys). As buscas foram realizadas contra um banco de dados local do NCBI (nr) com 9.993.394 entradas (19 de janeiro de 2010). Os arquivos no formato .dat gerados pelo Mascot foram utilizados para a validação estatística dos resultados no programa *Scaffold 2* da Proteome Software, USA (Keller *et al.*, 2002; Craig e Beavis, 2003; Nesvizhskii *et al.*, 2003). As identificações de proteínas foram aceitas sempre que a probabilidade calculada pelo algoritmo *Protein Prophet* (Nesvizhskii *et al.*, 2003) fosse maior ou igual a 95%.

3.11. Análise preliminar do NAF por LC-MS/MS

3.11.1. Digestão tríptica em solução

A digestão em solução do NAF bruto foi realizada segundo o protocolo de Valente R. H. (comunicação pessoal). Para cada 10 μg de proteínas de NAF bruto, foram adicionados 10 μL de solução de uréia 8 M / bicarbonato de amônio 400 mM e acrescentados 2,5 μL de ditiotreitol 100 mM, seguido de incubação a 37 °C, por 3 horas. Em seguida, as amostras foram incubadas por 15 min, no escuro, com 400 mM de iodoacetamida, à temperatura ambiente. Para diluir a uréia a 1M, adicionamos 65 μL de água milli Q. As amostras foram então incubadas com 1 μL de tripsina (0,2 $\mu\text{g}/\mu\text{L}$ em ácido acético 50 mM)(Promega, USA), por 16-24 horas, a 37 °C. A reação foi interrompida com ácido trifluoroacético (qsp 0,1% v/v). Os peptídeos trípticos foram secos no *speed vac* e enviados pelo correio para análise no LTQ Orbitrap XL (Thermo, EUA) da Universidade do Sul da Dinamarca, Odense, Dinamarca.

3.11.2. Identificação das proteínas no LTQ-Orbitrap XL

Os tripsinizados de NAF (10 μg) foram ressuspensos em 40 μL de TFA 1% (v/v) e aplicados em micro-colunas (≈ 7 mm altura) preparadas no laboratório, usando ponteiras de 10 μL e resina Poros R2 (Applied Biosystems, EUA) equilibrada nas mesmas condições (Thingholm *et al.*, 2006). Após lavagem com TFA 0,1% (v/v) para dessalinização, os peptídeos foram eluídos com 40 μL de TFA 0,1% (v/v) em acetonitrila 70% (v/v). Alíquotas de 2 μL ($< 0,5$ μg) foram secas no *speed vac*, ressuspensas em 5,5 μL de ácido fórmico 1% (v/v) e analisadas por LC-MS/MS usando o cromatógrafo EasyLC (Proxeon, Dinamarca) acoplado ao espectrômetro de

massas LTQ-Orbitrap XL (ThermoFisher Scientific, Alemanha). O nanocromatógrafo operou com uma única coluna analítica de 20 cm de comprimento e 100 µm de diâmetro interno, empacotada com resina C18 (ReproSil-Pur C18-AQ 3 µm, Dr. Maisch, Alemanha). Usamos ácido fórmico 0,1% (v/v) em água como solvente A e ácido fórmico 0,1% (v/v) em acetonitrila 90% (v/v) como solvente B. As amostras (5 µL) foram injetadas em TFA 1% (v/v) a um fluxo de 550 nL/min. Para separação dos peptídeos, usamos o fluxo de 250 nL/min e o seguinte programa de eluição: gradiente de 0-34% B em 100 min, seguido de gradiente de 34-100% B em 5 min e, finalmente, 100% B por 8 min.

A fonte de *nanospray* (Proxeon, Dinamarca) foi utilizada com voltagem de *spray* a 2.3 kV, sem auxílio de qualquer gás e com o capilar a uma temperatura de 200 °C. O espectrômetro de massas operou no modo de aquisição *data-dependent*, alternando automaticamente entre MS e MS/MS. No orbitrap, espectros *full scan MS* foram adquiridos com *Automatic Gain Control* (AGC) limitado a 1×10^6 íons e resolução de 60.000 (FWHM @ m/z 400). Para cada MS1, os cinco íons mais intensos eram isolados no *trap* linear (LTQ) para análise por MS/MS, excluindo-se os precursores monoprotonados. Para fragmentação dos íons no LTQ, utilizamos o método de dissociação induzida por colisão (*collision-induced dissociation* ou CID), com AGC limitado a 1×10^4 íons e energia de colisão (*normalized collision energy*) ajustada para 35. Todos os íons-fragmento foram analisados no *trap* linear. Os arquivos brutos com os dados do LTQ-orbitrap gerados (extensão RAW) foram convertidos pelo software BioWorks™ versão 3.3 (Thermo Electron Corporation, EUA) para o formato SRF e importados para o programa *Scaffold 2* da Proteome Software, USA (Keller *et al.*, 2002; Craig e Beavis, 2003; Nesvizhskii *et al.*, 2003). Todas as amostras de MS/MS foram analisadas usando o algoritmo *Sequest* versão 3.0 (ThermoFinnigan, EUA) e as buscas foram realizadas contra banco Swissprot com 432088 entradas (23 de fevereiro de 2010) assumindo a digestão com tripsina. Os parâmetros de busca utilizados foram: duas clivagens trípticas perdidas, $\pm 5,0$ ppm de erro no modo MS, $\pm 0,5$ Da de erro no modo MS/MS, carbamidometilação (Cys) como modificação fixa e oxidação (Met) como modificação variável. As identificações de proteínas foram aceitas sempre que a probabilidade calculada pelo algoritmo *Protein Prophet* (Nesvizhskii *et al.*, 2003) fosse maior ou igual a 95%.

4 - RESULTADOS E DISCUSSÃO

4.1. Coleta das amostras de NAF e dosagem de proteína

Mais de 95% dos cânceres de mama nascem de elementos epiteliais que revestem o sistema ductal mamário (Schnitt e Guidi, 2004). O fluido mamário (NAF) é uma amostra menos complexa do que o tecido tumoral e consiste de proteínas secretadas diretamente pelas células deste sistema ductal (Alexander *et al.*, 2004), pouco sujeito à diluição e/ou à contaminação por secreções de outros órgãos ou do plasma (Sauter *et al.*, 2004). Portanto, as proteínas presentes no NAF estão diretamente relacionadas com o microambiente de desenvolvimento e progressão do tumor. Além disso, como a mama é um órgão pareado, é possível fazer análise comparativa do NAF em uma mesma paciente com doença unilateral (Kuerer *et al.*, 2004). Salvo engano, este é o primeiro estudo proteômico sistemático realizado com amostras de NAF de pacientes brasileiras.

Em nossa experiência, a coleta do NAF, apesar de não invasiva, pode ser bastante trabalhosa e demorada. Neste estudo, conseguimos uma taxa de sucesso baixa, estimada em 33 % (1 a cada 3 tentativas). Os principais problemas enfrentados incluíram a dificuldade de fazer com que as pacientes retornassem ao ambulatório para novas tentativas de coleta, a presença de tumores bloqueando o sistema de ductos que conduzem o NAF para o meio exterior e a produção de volumes insuficientes de NAF (< 15 µL) em ambas ou em uma das mamas por algumas pacientes, impossibilitando a obtenção de amostras pareadas. Na literatura, entretanto, alguns trabalhos relatam sucesso de mais de 80% na coleta de NAF (Mitchell *et al.*, 2002; Djuric *et al.*, 2005; Higgins *et al.*, 2005). A disparidade de números se justifica já que, ao contrário do nosso trabalho, essas coletas foram realizadas, em média, ao longo de três visitas ao ambulatório do hospital, não sendo realizada coleta bilateral na mesma mulher.

No presente estudo, foram utilizadas 22 amostras de NAF, sendo oito de mamas com carcinoma ductal ou lobular infiltrante comprovado por biopsia percutânea, oito das mamas contralaterais sadias das mesmas pacientes (controles pareados com mesma numeração), quatro amostras de três doadoras saudáveis e duas amostras de mamas com tumor benigno. Estas amostras foram coletadas de 13 mulheres entre 26 e 85 anos (média 54 ± 16 anos), das quais seis estavam em pré- e sete em pós-menopausa. Somente três voluntárias nunca fizeram uso de contraceptivos e apenas uma paciente possuía história familiar em primeiro grau de câncer de mama. Nenhuma mulher fez reposição hormonal e todas tiveram pelo

menos um parto e amamentaram seus bebês. Os dados clínicos das amostras das 13 voluntárias que participaram do estudo estão detalhados na Tabela 4.1. Para identificação inequívoca das amostras, foi utilizada a seguinte codificação: Letra P (paciente) + número identificador do indivíduo + letras **e** (mama esquerda) ou **d** (mama direita).

Como o NAF é uma secreção muito viscosa e de difícil manipulação, as amostras coletadas foram imediatamente diluídas 10 vezes em PBS e centrifugadas a $250 \times g$, por 10 min, a 4°C . A baixa velocidade de centrifugação foi crítica para garantir a retirada de eventuais células epiteliais esfoliadas, sem que elas perdessem a integridade e contaminassem o NAF com seu conteúdo citoplasmático. O sobrenadante de NAF foi então coletado, aliquotado em $20 \mu\text{L}$ e imediatamente armazenado a -80°C , sem adição de inibidores de protease que pudessem interferir com as análises posteriores (ex.: mudança de carga e/ou massa detectável por 2D-PAGE) (Görg *et al.*, 2004).

As concentrações proteicas das 22 amostras de NAF utilizadas neste estudo variaram de 40 a $256 \mu\text{g}/\mu\text{L}$, com média de $115 \pm 60 \mu\text{g}/\mu\text{L}$ (Figura 4.1). A análise pelo teste T não pareado mostrou que as médias entre os grupos controle e mamas com câncer não apresentam diferença estatisticamente significativa (p valor = 0,4657). O grupo tumor benigno não foi testado em função do número muito reduzido de amostras ($n = 2$). De modo semelhante, Varnum e colaboradores (2003) relataram que a concentração protéica de NAF é bastante alta, normalmente ainda mais elevada do que a do plasma ($\approx 60 \mu\text{g}/\mu\text{L}$) (Anderson e Anderson, 2002). Analisando apenas quatro amostras, os autores mostraram concentrações proteicas variando entre 45 e $200 \mu\text{g}/\mu\text{L}$. Outro estudo relatou ainda concentração proteica média de $170 \mu\text{g}/\mu\text{L}$ no NAF (Djuric *et al.*, 2005). Estes resultados indicam que é possível obter quantidades expressivas de proteínas do NAF para a análise proteômica partindo de volumes bastante reduzidos de amostra.

As concentrações protéicas das duas mamas das sete pacientes com câncer de mama unilateral analisadas posteriormente por eletroforese diferencial em gel bidimensional podem ser visualizadas na Figura 4.2. Percebemos que quatro pacientes apresentam concentração semelhante entre a mama sadia e a doente, uma apresenta concentração proteica maior na mama com câncer e as outras duas pacientes apresentam concentração maior na mama sadia, não demonstrando relação clara da concentração proteica do NAF com o acometimento da mama.

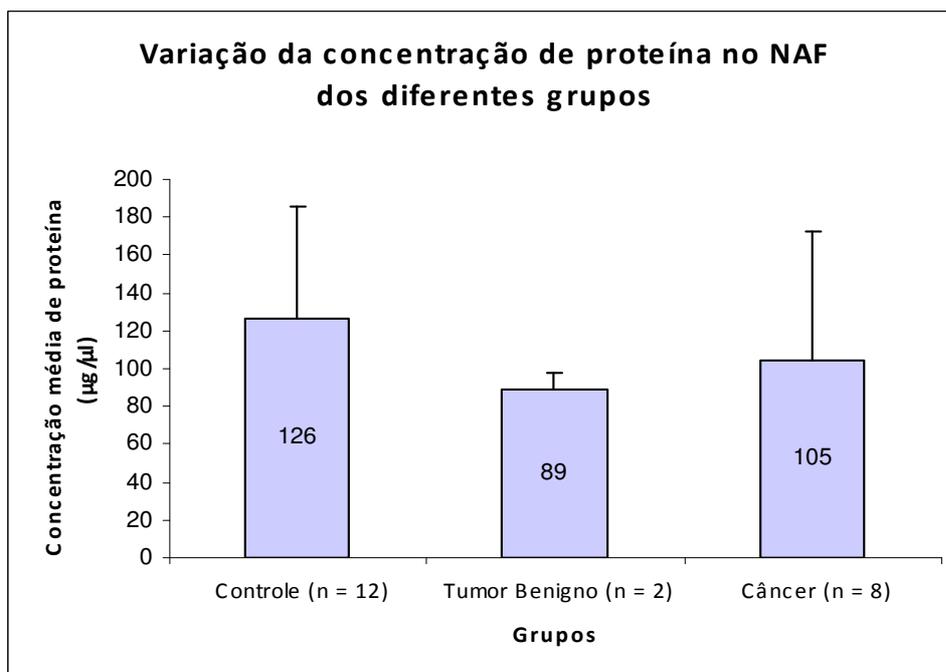


Figura 4.1: Concentração média de proteínas nas amostras dos grupos controle ($126 \mu\text{g}/\mu\text{L} \pm 60$), tumor benigno ($89 \mu\text{g}/\mu\text{L} \pm 8$) ou câncer ($104 \mu\text{g}/\mu\text{L} \pm 68$).

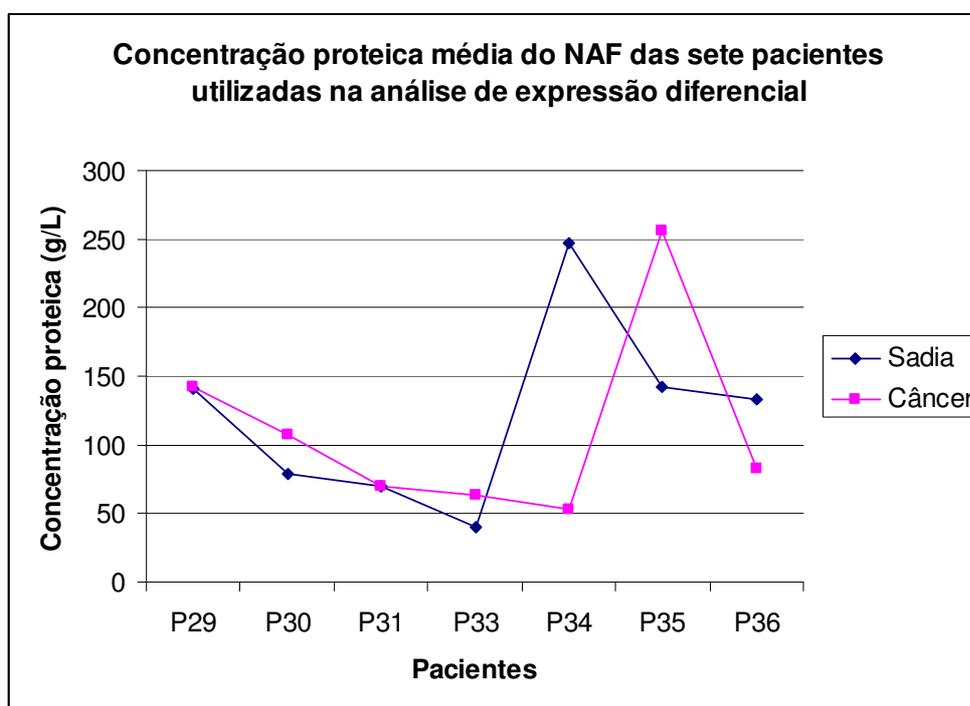


Figura 4.2: Gráfico demonstrando a concentração proteica das mamas pareadas das sete pacientes com câncer de mama unilateral utilizadas na análise posterior por eletroforese diferencial em gel bidimensional (2D-DIGE).

Tabela 4.1: Características clínico-patológicas das amostras do estudo.

Código da amostra	Grupo	Volume coletado (µL)	Concentração de proteína (µg/µL)	Idade (anos)	Menopausa	Contraceptivo	Reposição hormonal	Número de partos	Amamentação	Câncer de mama em primeiro grau na família
P2d	Controle	30	90	47	não	sim	não	3	sim	sim
P5d	Controle	16	210	26	não	não	não	pelos menos 1	sim	não
P5e	Controle	15	72	26	não	não	não	pelos menos 1	sim	não
P7d	Controle	17	132	53	sim	sim	não	3	sim	não
P11d	Controle	35	154	58	sim	sim	não	3	sim	não
P29e	Controle	21	141	48	não	sim	não	2	sim	não
P30e	Controle	16	79	64	sim	sim	não	1	sim	não
P31d	Controle	20	70	73	sim	não	não	6	sim	não
P33e	Controle	20	40	53	sim	sim	não	3	sim	não
P34e	Controle	40	246	66	sim	sim	não	7	sim	não
P35d	Controle	8	142	29	não	sim	não	4	sim	não
P36d	Controle	31	134	85	sim	não	não	3	sim	não
P4d	Tumor Benigno	30	83	53	não	sim	não	4	sim	não
P10d	Tumor Benigno	26	95	45	não	sim	não	3	sim	não
P2e	Câncer	20	63	47	não	sim	não	3	sim	sim
P29d	Câncer	13	142	48	não	sim	não	2	sim	não
P30d	Câncer	16	108	64	sim	sim	não	1	sim	não
P31e	Câncer	20	69	73	sim	não	não	6	sim	não
P33d	Câncer	16	63	53	sim	sim	não	3	sim	não
P34d	Câncer	44	53	66	sim	sim	não	7	sim	não
P35e	Câncer	12	256	29	não	sim	não	4	sim	não
P36e	Câncer	30	82	85	sim	não	não	3	sim	não

4.2. Análise do NAF por eletroforese unidimensional

As proteínas do NAF foram visualizadas por SDS-PAGE após normalização das amostras tanto por volume quanto por quantidade total de proteína. Os géis foram revelados por Coomassie G-250 coloidal e as principais bandas proteicas foram submetidas à tripsinização *in gel*, dessalinização e identificação por espectrometria de massas usando MALDI-TOF/TOF. A Figura 4.3 mostra que volumes equivalentes de NAF de quatro mulheres apresentaram perfis eletroforéticos com diferentes intensidades de banda (raias 1-4, P5d, P7d, P10d, P11d, respectivamente), indicando variabilidade quantitativa importante, como já observado através de dosagem de proteínas. Com uma exceção (raia 8), qualitativamente os padrões eletroforéticos se mostraram mais similares, com as seguintes bandas majoritárias se repetindo nas amostras das raias 5-7: albumina (banda 1 assinalada em vermelho na Figura 4.2, com 67 kDa), imunoglobulina (banda 2, 60 kDa), zinco- α 2-glicoproteína (banda 3, 46 kDa), apolipoproteína D (banda 4, 28 kDa), proteína induzida por prolactina (banda 5, 16 kDa) (Tabela 4.2). A amostra da raia 8 se mostrou bem mais rica em albumina, com a banda de proteína induzida por prolactina virtualmente ausente. As massas moleculares indicadas acima foram calculadas pelo programa *Image Master 2D Elite* (GE Healthcare), baseando-se nas mobilidades eletroforéticas dos padrões de massa molecular aplicados no mesmo gel. Os perfis eletroforéticos não sugeriram qualquer processo importante de autólise no NAF que justificasse a adoção de inibidores de protease no procedimento de preparação de amostra.

Sánchez e colaboradores (1992) analisaram amostras de NAF de 169 mulheres na pré-menopausa: 99 apresentando doenças de mama (maligna ou benigna) e 70 sem qualquer patologia mamária (Sanchez *et al.*, 1992). De acordo com o perfil das proteínas mais abundantes por SDS-PAGE, os autores propuseram a existência de duas grandes categorias. Ambas seriam encontradas nos dois grupos estudados, mas com diferentes prevalências: além de albumina e imunoglobulina, o NAF tipo I seria caracterizado pela predominância das bandas de zinco- α 2-glicoproteína (44 kDa), apolipoproteína D (24 kDa) e proteína induzida por prolactina (17 kDa) e foi observado na maioria das amostras controle (93%) e de doenças benignas (88%); o NAF tipo II seria mais rico em proteínas associadas ao leite, incluindo lactoferrina (80 kDa), α -lactalbumina (15 kDa) e lisozima C (14 kDa), sendo encontrado em 57% das amostras de pacientes com câncer de mama analisadas. De modo interessante, a expressão de lisozima em diversos tipos de tumor de mama parece estar relacionada com evoluções mais favoráveis (Vizoso *et al.*, 2001).

Em nosso estudo, as três amostras de mulheres saudáveis ou com doença benigna que mostraram perfis eletroforéticos similares na Figura 4.3 parecem se encaixar melhor na descrição do NAF tipo I (duas sendo pré-menopausa e uma pós-menopausa). Entretanto, a quarta amostra controle de mulher pós-menopausa não apresentou a banda referente à proteína induzida por prolactina, divergindo, portanto, do perfil tipo I aparentemente mais frequente em amostras normais pré-menopausa.

Outro estudo com 128 mulheres pré-menopausa com mamografias normais e que não utilizavam anticoncepcional revelou que os perfis de NAF por SDS-PAGE podiam ser classificados como tipo I (maioria dos casos), tipo II ou, ainda, uma mistura dos dois (< 10% casos). Os autores investigaram a relação entre o histórico de vida reprodutiva, a constituição corporal, a dieta e os níveis hormonais das pacientes com os diferentes perfis de NAF. Os resultados da pesquisa sugeriram que os diferentes perfis eletroforéticos das amostras de NAF de mamas saudáveis podiam sofrer influência da dieta de gordura e/ou da gordura corporal armazenada (Huang *et al.*, 2007). Esta conclusão seria compatível com a visão do tecido adiposo como um órgão endócrino capaz de sintetizar hormônios (andrógenos e estrógenos) e citocinas (interleucina-6 e TNF- α) (Bray, 2002) que poderiam regular de forma autócrina ou parácrina a expressão proteica na mama. Em cultura de células de câncer de mama, já foi mostrado que testosterona e progesterona aumentam a expressão de zinco- α 2-glicoproteína, apolipoproteína D e proteína induzida por prolactina, enquanto que o estradiol tem efeito contrário, citado por Bray (2002).

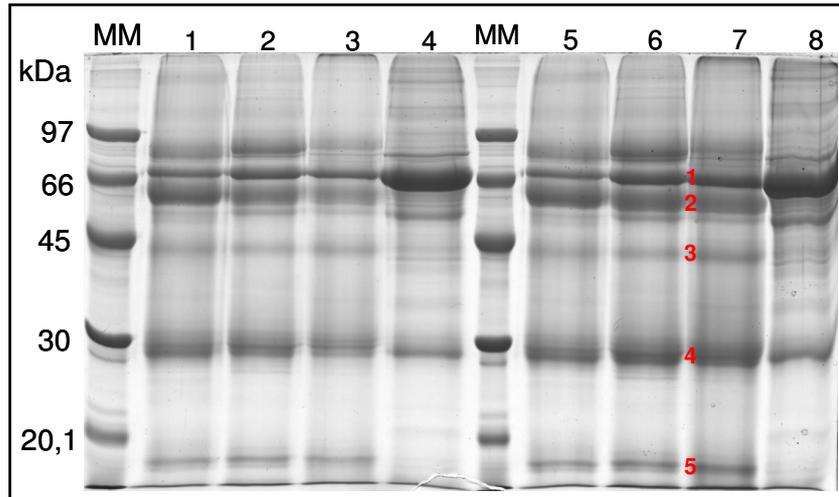


Figura 4.3: SDS-PAGE do NAF de indivíduos dos grupos controle (raias 1 e 5: P5d; raias 2 e 6: P7d) e tumor benigno (raias 3 e 7: P10d; raias 4 e 8: P11d) após normalização por volume de amostra (raias 1 a 4, 0,95 μ L NAF diluído 10x/raia) ou quantidade de proteína (raias 5 a 8, 20 μ g proteína/raia). As bandas assinaladas em vermelho e numeradas de 1 a 5 correspondem àquelas identificadas por espectrometria de massas no modo MS/MS (Tabela 2). Banda 1, albumina; banda 2, imunoglobulina; banda 3, zinco- α 2-glicoproteína; banda 4, apolipoproteína D; banda 5, proteína induzida por prolactina. Os géis foram corados com azul de Coomassie G-250 coloidal. MM, padrão de massa molecular.

Tabela 4.2: Identificações por MALDI TOF/TOF das bandas de eletroforese unidimensional de NAF.

Nome da proteína	Código de acesso no NCBI	Número da banda (posição na placa do MALDI)	Sequência	Mascore Ion score	Probabilidade do Scaffold	Precursor		Cobertura de sequência
						m/z	Carga	
Albumina	gij119626071	1 (G8)	(K)YLVEIAR(R)	46,2	95%	927,47	1	-29,06
			(K)KVPQVSTPTLVEVSR(N)	72,3	95%	1639,91	1	-15,09
			(K)FQNALLVR(Y)	59,9	95%	960,54	1	-28,39
			(R)RPCcamFSALEVDETYVPK(E)	39,1	95%	1910,91	1	-11,99
			(K)LVNEVTEFAK(T)	57,9	95%	1149,59	1	-20,13
			(R)FKDLGEEENFK(A)	51,4	95%	1226,57	1	-26,37
			(K)QNCcamELFEQLGEYK(F)	35,7	95%	1657,73	1	-16,41
			(K)AVMDDFAAFVEK(C)	48,0	95%	1342,60	1	-22,90
			(K)VFDEFKPLVEEPQNLK(Q)	28,2	88%	2045,08	1	-9,22
			(R)WLGGSQELPR(E)	81,4	95%	1213,62	1	-13,22
Imunoglobulina A	gij223069	2 (G10)	(K)SAVQGGPPER(D)	37,5	95%	940,47	1	-17,66
			(R)VPPPPPCcamCcamHPR(L)	58,0	95%	1313,61	1	-14,37
			(R)DASGATFTWTPSSGK(S)	37,8	95%	1512,67	1	-19,80
			(R)WLGGSQELPR(E)	59,0	95%	1213,63	1	-5,89
			(R)DASGATFTWTPSSGK(S)	44,9	95%	1512,68	1	-11,93
Zinco- α -glicoproteína	gij119597029	3 (G12)	(K)IDVHWTR(A)	34,5	95%	926,47	1	-14,79
			(R)AGEVQPELR(G)	45,5	95%	1127,55	1	-15,74
			(K)QKWEAEPVYVQR(A)	30,1	95%	1532,76	1	-18,82
			(K)YYDGDYIEFNK(E)	34,3	95%	1717,74	1	-19,88
			(K)IDVHWTR(A)	42,4	95%	926,49	1	2,53
Apolipoproteína D	gij4502163	4 (G18)	(R)AGEVQPELR(G)	38,7	95%	1127,57	1	4,48
			(K)AYLEEECcamPATLR(K)	29,1	95%	1451,68	1	-0,49
			(K)YYDGDYIEFNK(E)	33,7	95%	1717,77	1	-5,79
			(K)CcamPNPPVQENFDVVK(Y)	50,2	95%	1.657,71	1	-34,81
			(R)NPILPETVDSLK(N)	41,8	95%	1.423,68	1	-42,11
Proteína induzida por prolactina	gij116642259	5 (H1)	(K)NILTSNNIDV(K)	29,0	95%	1.230,62	1	-36,69
			(K)CcamPNPPVQENFDVVK(Y)	62,8	95%	1.657,71	1	-30,71
			(R)NPILPETVDSLK(N)	76,4	95%	1.423,69	1	-37,75
			(K)NILTSNNIDV(K)	57,9	95%	1.230,62	1	-37,34
			(K)MIVTDQVNCcamPK(L)	38,6	95%	1.292,55	1	-35,52
As modificações estão simbolizadas: carbamidometilação de cisteína - Ccam; oxidação de metionina - Mox.	gij116642259	5 (H2)	(R)ELGICcamPDDAAVPIKNNR(F)	38,4	95%	1.995,05	1	8,11
			(R)FYTIEILKVE(-)	38,9	95%	1.254,69	1	-5,06
			(K)YTACcamLCcamDDNPK(T)	56,3	95%	1.356,53	1	-22,84

4.3. Análise zimográfica do NAF

Para determinar se as amostras de NAF estudadas apresentavam atividade proteolítica importante, utilizamos a abordagem de análise zimográfica em géis de poliacrilamida copolimerizados com gelatina como substrato (Heussen e Dowdle, 1980). As análises foram feitas em quatro amostras de duas pacientes: duas amostras de mamas controle (P29e e P34e) pareadas com suas correspondentes das mamas com câncer (P29d e P34d). Ainda que tenhamos utilizado quantidades idênticas de proteína (20 µg), as duas pacientes apresentaram padrões distintos de hidrólise e a intensidade das bandas de degradação de gelatina entre as amostras das duas mamas de uma dessas pacientes se mostrou bastante diferente (Figura 4.4). Nos painéis A e B (Figura 4.4, raia 1), estão mostrados os resultados para a paciente P29, onde se observa uma atividade proteolítica bastante discreta localizada em uma banda de aproximadamente 66 kDa presente tanto na mama sadia quanto na doente. No caso da paciente P34, os mesmos 20 µg de proteínas do NAF foram capazes de induzir degradação intensa da gelatina em três bandas principais, estimadas em 85, 66 e 60 kDa (Figura 4.4, painéis C e D, raia 1), independente da presença de patologia na mama.

Para determinar a classe de protease(s) responsável(eis) pela atividade enzimática no NAF, após renaturação das proteínas em tampão contendo Triton X-100, os géis foram incubados no tampão de lise contendo inibidores de metaloproteases (EDTA ou ortofenantrolina), cisteíno-proteases (E-64), serino-proteases (APMSF) ou aspártico-proteases (pepstatina). Como este último ensaio teve que ser feito em pH 5 (Beynon, 1996), repetimos o controle de NAF sem inibidor nesta condição (Figura 4.4, raias 2). Em todas as amostras, as bandas gelatinolíticas desapareceram parcial ou totalmente na presença dos inibidores de metaloprotease (Figura 4.4, raias 6 e 7). Os inibidores pepstatina, E-64 e APMSF (Figura 4.4, raias 3, 4 e 5) não mostraram qualquer efeito. A inibição por agentes quelantes de cátions indica que a atividade gelatinásica do NAF se deve fundamentalmente à presença de metaloproteases neste fluido.

Também usando zimografia, Mannello e Sebastiani (2003) demonstraram que amostras de NAF de doadoras saudáveis, com doença benigna e com câncer apresentaram um perfil de gelatinases com quatro bandas principais (225, 130, 92 e 72 kDa) sensíveis à inibição por EDTA e ortofenantrolina e que foram reveladas imunologicamente como MMP-2 e MMP-9 e seus zimogênios (Mannello e Sebastiani, 2003). Apesar das nossas amostras revelarem

bandas com massas moleculares menores, é possível que elas correspondam apenas às formas ativas das MMP-2 (62-68 kDa) e MMP-9 (66-88 kDa) (Woessner e Nagase, 2000). Atualmente, estamos tentando estabelecer um protocolo de identificação destas bandas zimográficas usando espectrometria de massas. As proteases seriam tripsinizadas *in situ*, diretamente nas bandas de lise da zimografia, tornando o método de identificação muito mais direto e inequívoco do que a revelação por anticorpos (que não pode ser feita diretamente no zimograma). A limitação que deve ser superada, no entanto, se deve à presença de quantidades bastante grandes de peptídeos de gelatina nestas bandas, que suprimem o sinal dos peptídeos das metaloproteases. Uma tentativa que estamos fazendo é a extração destes peptídeos de gelatina do gel antes da tripsinização das bandas de protease. Outra possibilidade que vamos testar é submeter as bandas gelatinolíticas a um passo adicional de eletroforese (SDS-PAGE) para separar as proteases dos peptídeos de gelatina, antes da tripsinização *in gel*.

Ainda que tenhamos estudado apenas duas pacientes e que o ensaio de zimografia seja semi-quantitativo (Mannello e Sebastiani, 2003), a análise visual dos géis parece sugerir que a atividade proteolítica está aumentada na amostra de NAF da mamas com câncer de uma das pacientes (Figura 4.4, painel C). Analisando 6 pacientes, Mannello e Sebastiani (2003) também mostraram aumento de atividade proteolítica de MMP-2 e 9 no NAF de mamas com câncer. Outros estudos com soro, plasma e tecido de pacientes com câncer de mama demonstraram por zimografia e por imunohistoquímica um aumento nos níveis de MMP 9 relacionado com parâmetros clínico-patológicos (receptor de estrógeno, c-erB-2, grau histológico) (Ranuncolo *et al.*, 2003; La Rocca *et al.*, 2004; Vizoso *et al.*, 2007). Análises bioquímicas e citológicas de NAF sugerem que as células neoplásicas apresentam características de intensa biossíntese local que podem estar relacionadas ao intenso remodelamento da matriz na lesão maligna (Malatesta *et al.*, 2000; Mannello e Sebastiani, 2003; Roy *et al.*, 2009). Entretanto, ao fazermos estudos comparativos de NAF, não podemos perder a perspectiva de que em mulheres pré-menopausa o perfil eletroforético das proteínas deste fluido pode estar relacionado com o estado fisiológico da mulher, incluindo parto recente, dieta de gordura saturada e massa corporal de gordura (Vizoso *et al.*, 1992; Huang *et al.*, 2007).

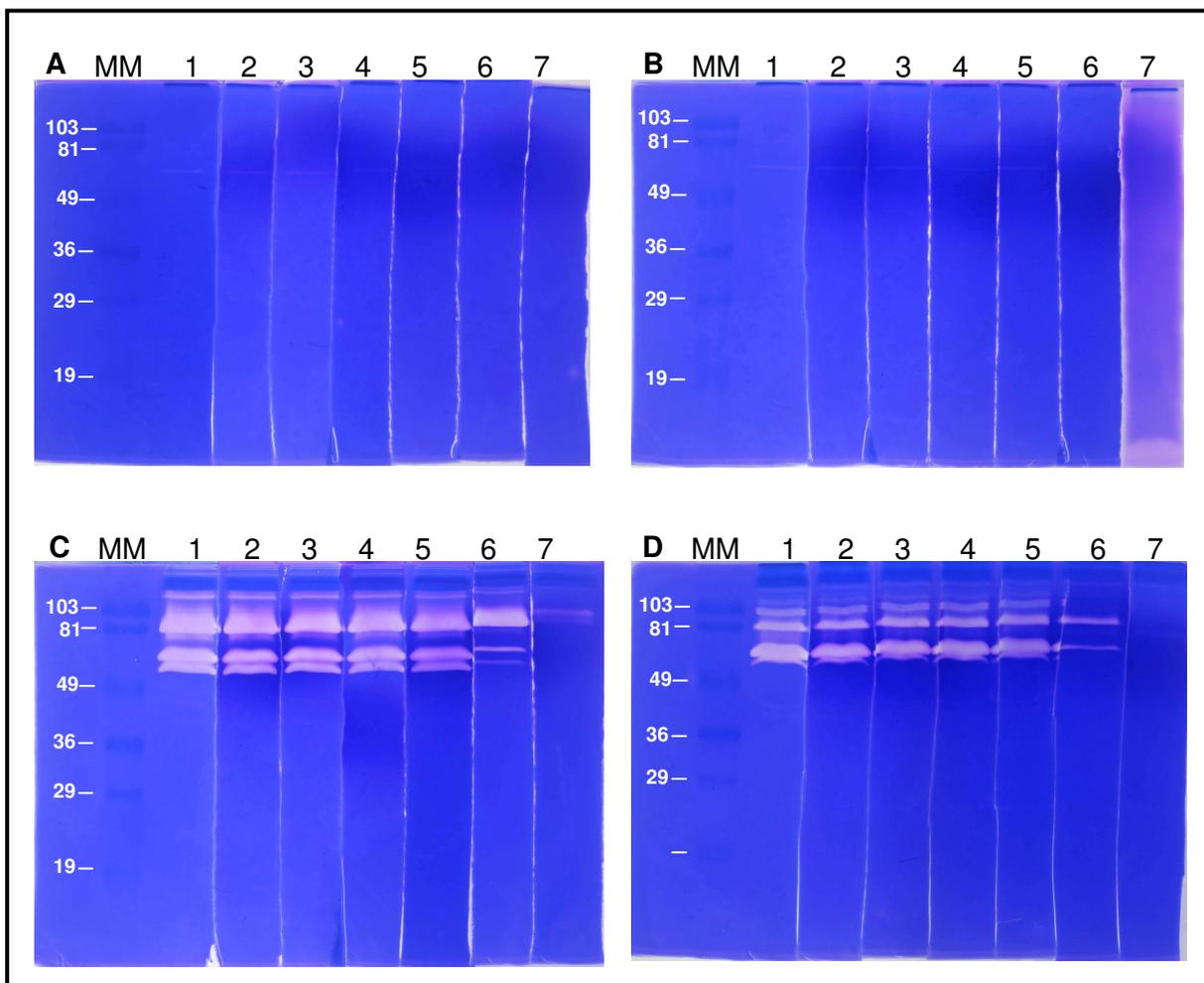


Figura 4.4: Zimografias de amostras de NAF coletadas de mamas doentes e sadias de duas pacientes. Cada painel refere-se a um gel único que foi dividido para a incubação com o tampão de hidrólise. Nos painéis A e C estão mostradas as amostras de NAF P29d e P34d, retiradas das mamas doentes de duas pacientes com câncer unilateral. As amostras de NAF das mamas sadias correspondentes estão mostradas nos painéis B (P29e) e D (P34e). Foram aplicados 20 μ g de proteínas de NAF em cada raia, incubadas alternadamente com: Raia 1, tampão de hidrólise sem inibidor; raia 2, tampão de hidrólise sem inibidor, pH 5,0 (controle específico para a pepstatina); raia 3, tampão de hidrólise contendo 1 μ M de pepstatina; raia 4, tampão de hidrólise contendo 10 μ M de E-64; raia 5, tampão de hidrólise contendo 100 μ M de APMSF; raia 6, tampão de hidrólise contendo 10 mM de EDTA; raia 7, tampão de hidrólise contendo 10 mM de ortofenantrolina. Os géis foram corados com azul de Coomassie R-250 e fotografados. MM, padrão de massa molecular pré-corado (*Low Range*, BioRad).

4.4. Otimização das condições de 2D-PAGE para as amostras de NAF

A análise proteômica do fluido mamário tem atraído atenção por ser um modo não invasivo de estudar o microambiente associado ao câncer de mama (Kuerer *et al.*, 2002). Vários estudos indicam que o NAF deve conter marcadores potenciais para o diagnóstico e prognóstico do câncer de mama, e as técnicas proteômicas representam as ferramentas mais modernas e sensíveis para se chegar à identificação dessas moléculas (Kuerer *et al.*, 2002; Pawlik *et al.*, 2005; Somiari *et al.*, 2005).

A 2D-PAGE é uma técnica com grande poder de resolução. Ela permite a separação de proteínas a partir de misturas complexas, gerando um mapa das cadeias peptídicas íntegras, onde se pode avaliar tanto os diferentes níveis de expressão proteica, quanto a presença de modificações pós-traducionais dando origem a diferentes isoformas. Entretanto, para se tirar o máximo proveito da técnica, para cada amostra em particular é preciso otimizar as condições de análise, incluindo os passos de extração, solubilização e fracionamento das proteínas (Görg *et al.*, 2004).

Para concentração das amostras de NAF e retirada de contaminantes (ex.: sais, polissacarídeos e ácidos nucleicos) que pudessem atrapalhar a eletroforese bidimensional, as amostras foram precipitadas com TCA e, após lavagens do *pellet* com acetona, imediatamente dissolvidas em solução aquosa contendo agente caotrópico, detergente e DTT. Esta solução é importante para garantir a completa desnaturação, desagregação, redução e solubilização das proteínas. A idéia é que todas as interações moleculares entre as proteínas do NAF sejam desfeitas, de forma que cada *spot* visualizado no 2D-PAGE represente uma única cadeia polipeptídica (Görg *et al.*, 2004). Para determinar as melhores condições de separação dos polipeptídeos na focalização isoeletrica do NAF, fizemos testes variando o intervalo de pH e o método de aplicação de amostra.

Inicialmente, foi realizada a separação das cadeias polipeptídicas de uma mesma amostra em fitas de 7 cm, nos intervalos de pH de 3-10 NL (Figura 4.5A) e 4-7 (Figura 4.5B). Como a maioria dos *spots* foi focalizada na faixa de pH entre 4 e 7, escolhemos este intervalo para todas as análises posteriores. Comparamos também os dois métodos de aplicação de amostra na focalização isoeletrica do NAF, o *cup loading* e a reidratação *in gel*. Pelo método de *cup loading* (Figura 4.6A), obtivemos 191 *spots* e melhor resolução dos polipeptídeos de

maior massa molecular e mais básicos. Entretanto, na região do anodo onde a amostra foi aplicada, verificamos um arraste vertical importante que atrapalha a visualização dos *spots*. Isso pode ser resultado de agregação ou precipitação das proteínas na região do *cup*, local onde a amostra é aplicada na fita de focalização (Görg *et al.*, 2004). Por reidratação *in gel* (Figura 4.6B), obtivemos um número maior de *spots* (217) com boa resolução e com menos arraste vertical, o que justificou a escolha deste último como método padrão de aplicação de amostras de NAF nos géis bidimensionais. Os géis mostrados nas Figuras 4.5 e 4.6 são representativos de pelo menos 5 replicatas (2 técnicas e 3 biológicas) reprodutíveis.

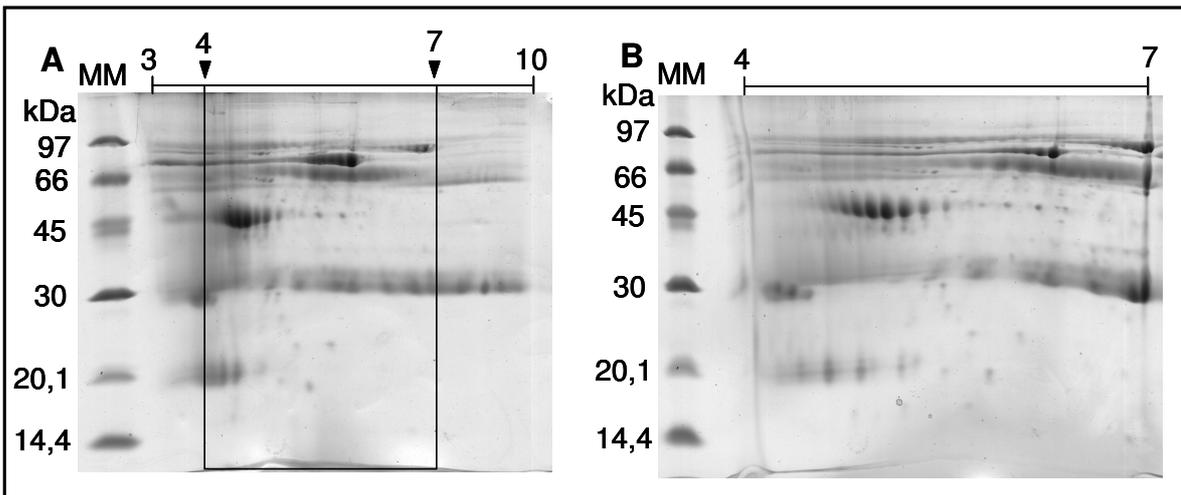


Figura 4.5: 2D-PAGE de uma amostra de NAF de indivíduo normal (P7d). Foram aplicados 150 µg de proteína pela técnica de reidratação *in gel*, em géis de 7 cm, com intervalos de pH de 3-10NL (A) e 4-7 (B). Os géis foram corados com azul de Coomassie G250 coloidal. MM, padrão de massa molecular.

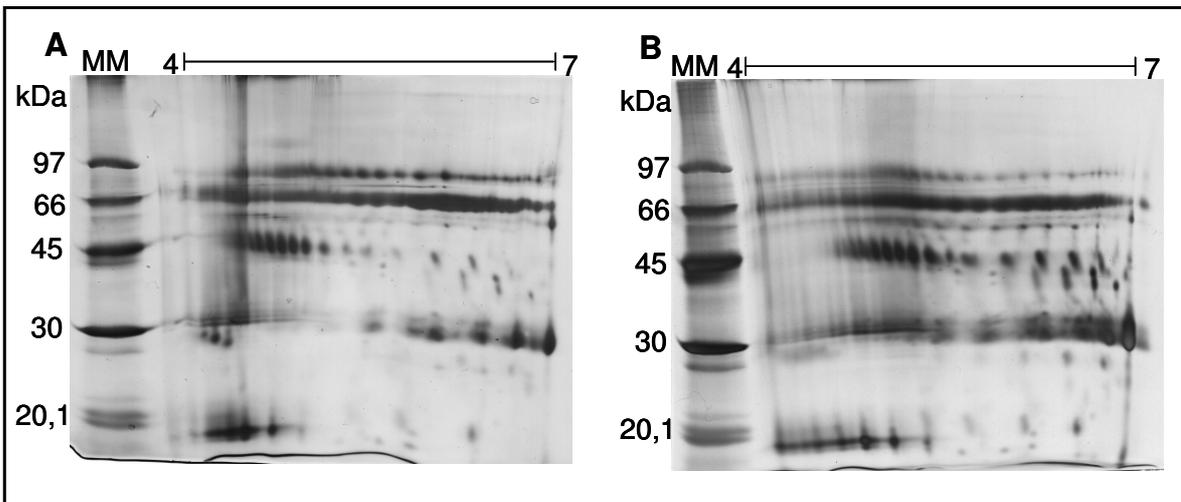


Figura 4.6: 2D-PAGE de uma amostra de NAF de indivíduo normal (P5d). Foram aplicados 150 µg de proteína em fitas de IEF de 7 cm pH 4-7, utilizando os métodos de aplicação de amostra de (A) *cup loading* e (B) reidratação *in gel*. Os géis foram revelados após impregnação com nitrato de prata. MM, padrão de massa molecular.

4.5. Caracterização do conteúdo proteico do NAF por 2D-PAGE e espectrometria de massas.

As análises proteômicas devem ser encaradas como uma abordagem geradora de hipóteses. As técnicas proteômicas permitem a caracterização do conteúdo proteico global de uma amostra, agregando informações moleculares valiosas a processos fisio-patológicos e/ou sugerindo possíveis candidatos a biomarcadores a serem validados em fase posterior do estudo, não necessariamente por proteômica (Somari *et al.*, 2005). Wulfkühle e colaboradores (2002) analisaram o tecido tumoral mamário de 6 pacientes por 2D-PAGE e espectrometria de massas (MS) e identificaram 57 proteínas diferencialmente expressas, sendo 10 posteriormente validadas como proteínas candidatas por imunohistoquímica. Essas proteínas estão envolvidas nas mais diversas funções celulares, tais como regulação do transporte intracelular, composição do citoesqueleto, apoptose e instabilidade genética (processamento de mRNA e manutenção telomérica) (Wulfkühle *et al.*, 2002), mostrando a importância dessas análises para se alcançar um melhor entendimento da doença.

Como a mama é um órgão pareado, há possibilidade de fazer análise comparativa na mesma paciente, em casos de câncer de mama unilateral, entre a amostra da mama doente e da mama sadia (que funciona como um controle interno) (Kuerer *et al.*, 2002). Portanto, para caracterização inicial da composição proteica das amostras brasileiras, o NAF das duas mamas de uma paciente com carcinoma ductal infiltrante unilateral (P2d e P2e) foi submetido à eletroforese bidimensional usando fitas de focalização isoeletrica de 18 cm, com intervalos de pH 4-7. Os mapas bidimensionais das duas amostras pareadas se mostraram equivalentes (Figura 4.7), com poucos *spots* muito abundantes distribuídos em fileiras (*spot trains*) sugestivas da presença de isoformas, ou seja, apresentando massas moleculares similares, mas pontos isoeletricos distintos. Esta característica é bastante comum em proteínas plasmáticas e se deve fundamentalmente à presença de combinações complexas de modificações pós-traducionais (especialmente glicosilações) nestas moléculas, que podem ser discriminadas pela eletroforese bidimensional (Anderson e Anderson, 2002). *Spots* distribuídos por todo o gel foram excisados e digeridos *in gel* com tripsina para identificação por espectrometria de massas (MALDI-TOF-TOF). Dos 54 *spots* excisados, 21 deram sequência por MS/MS, resultando na identificação de 7 proteínas diferentes. As proteínas detectadas foram albumina, diversas classes de imunoglobulina, zinco- α 2-

glicoproteína, transferrina e proteína induzida por prolactina. A lista completa das proteínas identificadas encontra-se na Tabela 4.3.

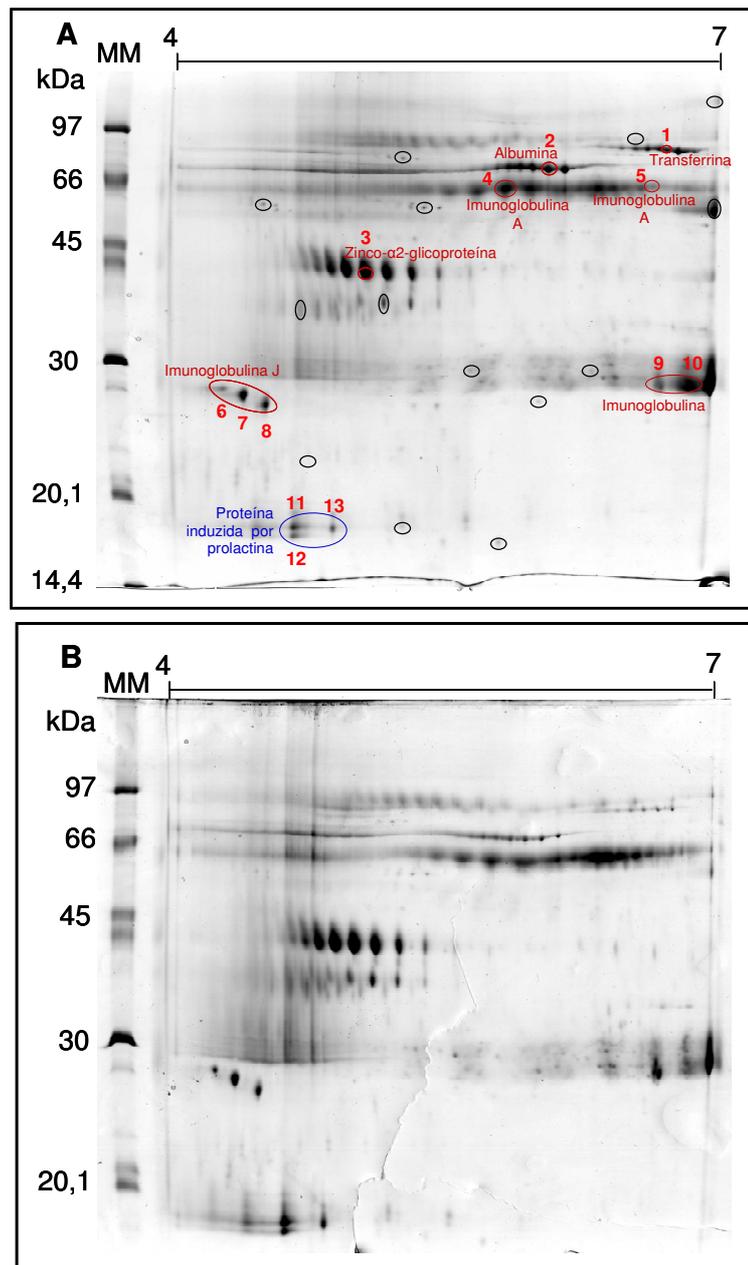


Figura 4.7: 2D-PAGE das amostras das mamas sadia (A) e doente (B) de uma paciente com carcinoma ductal infiltrante unilateral (P2d e P2e, respectivamente). Foram aplicados 400 μ g de proteína em cada gel através da técnica de reidratação *in gel*, usando fitas de IEF de 18 cm, no intervalo de pH 4-7. Os géis foram corados com azul de Coomassie G-250 coloidal. MM, padrão de massa molecular. Os *spots* proteicos identificados por MALDI-TOF/TOF estão indicados apenas no gel da mama sadia (a numeração em vermelho refere-se ao número do *spot* da Tabela 3), mas foram sequenciados nos dois géis. A proteína induzida por prolactina está assinalada em azul.

Tabela 4.3. Identificações por MALDI TOF/TOF dos spots de 2D-PAGE das duas mamas pareadas de uma paciente com câncer unilateral.

Nome da proteína	Código de acesso no NCBI	Número do spot (Posição na placa de MALDI)	Sequência	Mascot Ion score	Probabilidade do Scaffold	Precursor		Cobertura de sequência
						m/z	Carga	
Transferrina	gi 136191	1 (C4)	(K)EGYGYGTGAFR(C)	33,77	95%	1283,56	1	-7,67
			(R)KPVVEEYANC _{cam} HLAR(A)	24,08	95%	1586,77	1	-3,10
			(R)FDEFSEGC _{cam} APGSK(K)	44,08	95%	1577,64	1	-9,52
Albumina	gi 113576	2 (C5)	(K)KVPQVSTPLVEVSR(N)	22,58	95%	1639,94	1	-0,52
			(K)LVNEVTEFAK(T)	38,32	95%	1149,60	1	-11,70
			(K)SLHTLFGDK(L)	33,96	95%	1017,53	1	-2,31
Zinco- α -glicoproteína	gi 141596	3 (C11)	(K)AYLEEEC _{cam} PATLR(K)	57,48	95%	1451,73	1	29,20
			(R)AGEVQPELR(G)	34,45	95%	1127,61	1	32,77
			(K)HVEDVPAFGALGSLNDLQFFR(Y)	55,02	95%	2403,27	1	25,95
			(K)WEAEPVYVQR(A)	47,80	95%	1276,67	1	28,79
			(K)AREDFIMETLK(D)	40,11	95%	1352,73	1	28,65
			(R)YSLTYITGLSK(H)	61,79	95%	1408,77	1	24,17
			(R)QEPSQGTTFVAVTSILR(V)	57,72	95%	1835,95	1	1,77
			(R)DASGVFTWTPSSGK(S)	51,73	95%	1540,73	1	-0,16
			(K)TFC _{cam} TAAYPEK(T)	38,20	95%	1375,61	1	-5,39
			(R)QEPSQGTTFVAVTSILR(V)	46,76	95%	1835,96	1	3,52
			(R)DASGVFTWTPSSGK(S)	45,13	95%	1540,74	1	5,36
Imunoglobulina A	gi 113584	4 (C6)	(R)SSEDPNEDIVER(N)	28,24	95%	1389,67	1	41,70
			(R)FVYHLSDL _{Cam} K(K)	36,75	95%	1281,69	1	43,53
			(K) _{Cam} YTA _{cam} VPLVYGGETK(M)	42,56	95%	1656,93	1	59,32
Imunoglobulina J	gi 21489959	7 (C16)	(R)IIVPLNRR(E)	43,19	95%	938,62	1	44,80
			(R)FVYHLSDL _{Cam} K(K)	55,65	95%	1281,68	1	41,97
			(R)SSEDPNEDIVER(N)	63,72	95%	1389,67	1	41,70
Imunoglobulina cadeia leve	gi 149673891	8 (C17)	(K) _{Cam} YTA _{cam} VPLVYGGETK(M)	46,39	95%	1656,90	1	42,00
			(R)SSEDPNEDIVER(N)	50,85	95%	1389,70	1	62,71
			(R)FVYHLSDL _{Cam} K(K)	32,00	95%	1281,70	1	57,57
Proteína induzida por prolactina	gi 134170	9 (C21)	(K)SGTASV _{Cam} LLNNFYPR(E)	30,46	95%	1798,00	1	59,89
			(K)VYAcEVT _{cam} HQGLSSPVT(K)(S)	64,36	95%	1876,04	1	58,47
			(R)TVAAPSVFIFPPSDEQLK(S)	89,75	95%	1946,13	1	52,80
Proteína induzida por prolactina	gi 134170	10 (C22)	(K)VDNALQSGNSQESVTEQDSK(D)	73,80	95%	2136,10	1	58,92
			(K)DSTYLSLSLTLSK(A)	43,66	95%	1502,84	1	52,94
			(K)VYAc _{cam} EVTHQGLSSPVT(K)(S)	55,64	95%	1876,02	1	49,51
Proteína induzida por prolactina	gi 134170	11 (C24)	(K)VDNALQSGNSQESVTEQDSK(D)	60,51	95%	2136,09	1	57,09
			(R)TVAAPSVFIFPPSDEQLK(S)	60,95	95%	1946,13	1	54,34
			(K)SVRPNIDEVAVLAVGTQLK(E)	29,56	95%	2069,24	1	58,02
Proteína induzida por prolactina	gi 134170	12 (D1)	(R)FYTIELK(V)	63,20	95%	1026,63	1	37,07
			(R)TVQIAAVDVIR(E)	57,27	95%	1283,83	1	50,91
			(K)TYLISIPLOGAFNYK(Y)	42,86	95%	1815,07	1	54,57
Proteína induzida por prolactina	gi 134170	13 (D2)	(K)YTAC _{cam} LD _{cam} DDNPK(T)	69,98	95%	1356,62	1	44,17
			(R)ELGIC _{cam} PDDAAVPIK(N)	97,80	95%	1610,93	1	50,09
			(K)YTAC _{cam} LD _{cam} DDNPK(T)	54,05	95%	1356,48	1	-54,76
As modificações estão simbolizadas: carbamidometilação de cisteína - C _{cam} ; oxidação de metionina - Mox.								

A zinco- α 2-glicoproteína (ZAG) é uma proteína sintetizada em adipócitos, células epiteliais da próstata e do fígado e expressa em fluidos humanos como o plasma, sêmen, suor, saliva, urina, líquidos amniótico e cerebrospinal e no leite materno, citado por Hassan *et al.* (2008) (Hassan *et al.*, 2008). Ela está associada à mobilização de lipídeos e apresenta homologia com o fator de mobilização de lipídeos e expressão elevada na caquexia (extensiva redução de gordura corporal) associada ao câncer (Alexander *et al.*, 2004; Hassan *et al.*, 2008). A presença desta proteína em altas concentrações foi descrita em fluidos de mamas com doença benigna, chegando a níveis 30 a 50 vezes maiores do que as concentrações plasmáticas (Bundred *et al.*, 1987). Níveis aumentados também já foram detectados no soro de pacientes com câncer de mama ou com cistos mamários (Celis *et al.*, 2006). Pacientes com linfonodos positivos ou em estágios mais avançados da doença apresentaram níveis maiores de ZAG no soro quando comparados àquelas com linfonodo negativo ou estágios mais precoces (Bundred *et al.*, 1991). A função da ZAG em condições fisiológicas normais e no câncer ainda não é bem conhecida, no entanto esta proteína é considerada uma molécula marcadora da atividade apócrina da glândula mamária e pode ter valor prognóstico para o câncer de mama (Bundred *et al.*, 1991; Hassan *et al.*, 2008).

A proteína induzida por prolactina, também conhecida como GCDFP-15 (*gross cystic disease fluid protein*), é produzida principalmente em condições patológicas na mama e é expressa constitutivamente em várias glândulas exócrinas, tais como lacrimal, salivar e sudorípara (Mazoujian *et al.*, 1983; Alexander *et al.*, 2004). Sua secreção é induzida por prolactina e andrógeno e inibida por estrógeno (Alexander *et al.*, 2004). A GCDFP-15 é uma das principais proteínas identificadas no fluido acumulado nos macrocistos de mamas (*human breast gross cystic disease* – doença benigna), que apresentam secreção apócrina característica. Alguns tipos histológicos de carcinomas mamários revelam por imunohistoquímica a presença da GCDFP-15 (Mazoujian *et al.*, 1983). Contraditoriamente, um estudo com NAF de mulheres na pré-menopausa mostrou que os níveis de GCDFP-15 em amostras de câncer se apresentaram diminuídos quando comparados com os níveis nas amostras normais. A grande influência dos níveis hormonais sobre a expressão desta proteína poderia explicar a discrepância de resultados entre mulheres normais e pacientes em situações diferentes (Hall *et al.*, 1998; Alexander *et al.*, 2004). No presente trabalho, os *spots* de proteína induzida por prolactina (indicados em azul) parecem mais intensos na mama doente da paciente P2 (Figura 4.7B) do que na sua correspondente sadia (Figura 4.7A). Obviamente que não se

pode concluir nada utilizando uma única paciente. Comparações quantitativas utilizando mais pacientes também foram realizadas e serão discutidas mais adiante.

Além de identificar as duas proteínas descritas acima, possivelmente relacionadas a doenças de mama (benigna e câncer), também identificamos no NAF proteínas típicas do plasma humano, como a albumina, imunoglobulinas e transferrina. Em conjunto, elas correspondem à maior parte das proteínas do NAF detectadas no gel bidimensional pela coloração de azul de Coomassie G-250 coloidal. Para tentar enriquecer o NAF em seus componentes menos abundantes, resolvemos testar a depleção de algumas de suas proteínas mais abundantes. Nesta etapa de pré-fracionamento, utilizamos as abordagens de depleção por afinidade com anticorpo ou com proteína L, descritas a seguir.

4.6. Depleção de proteínas abundantes do NAF por afinidade com anticorpos

Proteínas muito abundantes prejudicam a separação e detecção de proteínas menos abundantes, uma vez que limitam a quantidade dessas proteínas carregadas no gel e as encobrem no perfil da eletroforese bidimensional (Görg *et al.*, 2004). A estratégia de remoção dessas proteínas tem sido amplamente utilizada em estudos proteômicos de amostras biológicas humanas com intervalo dinâmico muito grande, como o plasma (Varnum *et al.*, 2003; Albuquerque *et al.*, 2009). Em muitos casos, este pré-fracionamento possibilita o enriquecimento, visualização e identificação de proteínas menos abundantes em misturas complexas. A maior desvantagem é a perda de amostra, seja pela necessidade de maior manipulação e/ou pela possibilidade de ocorrência de depleções inespecíficas (Görg *et al.*, 2004; Omenn *et al.*, 2005).

Nesta dissertação, amostras de NAF de mamas saudáveis foram submetidas à cromatografia na coluna *Agilent* de múltipla afinidade originalmente desenvolvida para remoção das seis proteínas mais abundantes do plasma humano (albumina, IgG, IgA, transferrina, haptoglobina e antitripsina). A fração do NAF que interagiu com a coluna foi analisada por 2D-PAGE e todos os *spots* protéicos foram excisados, tripsinizados e identificados por espectrometria de massas (Figura 4.8 e Tabela 4.4). Esta análise mostrou que, além das proteínas alvo da coluna, a proteína induzida por prolactina também foi depletada, indicando que a coluna não é tão específica no caso de NAF. Por este motivo, o método de pré-purificação usando esta coluna de afinidade com anticorpos foi abandonada para o estudo do NAF. Em outros estudos com o NAF (Varnum *et al.*, 2003; Alexander *et al.*, 2004),

a remoção de albumina e imunoglobulinas utilizando kits comerciais também não foi efetiva e não melhorou sensivelmente a detecção de proteínas menos abundantes.

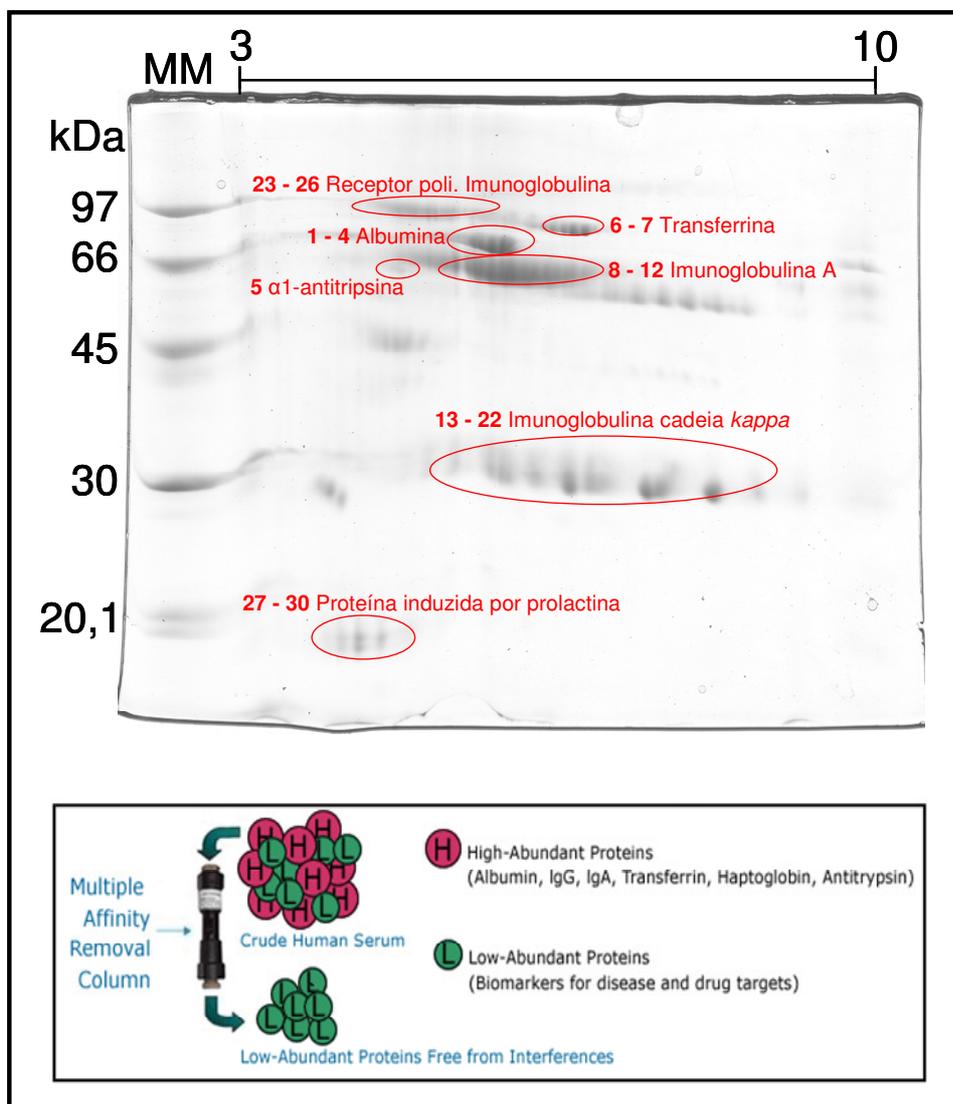


Figura 4.8: 2D-PAGE da fração de NAF da mama controle de indivíduo com câncer (P2d) que interagiu com a coluna de múltipla afinidade da Agilent. Foram aplicados 150 µg de proteína em uma fita de IEF de 7 cm, com intervalo de pH 3-10. A amostra foi aplicada por reidratação *in gel* e o gel foi corado com Coomassie G-250 coloidal. MM, padrão de peso molecular. Os spots identificadas por MALDI-TOF/TOF estão indicados em vermelho e os números em negrito e vermelho (1 – 30) estão relacionados ao número do spot na Tabela 4. O esquema abaixo da imagem do gel demonstra a depleção de proteínas séricas pela coluna. Fonte: Agilent Technologies.

Tabela 4.4: Identificações por MALDI TOF/TOF dos spots da 2D-PAGE da fração do NAF que interage com a coluna de múltipla afinidade da Agilent.

Nome da proteína	Código de acesso no NCBI	Número da banda (posição na placa do MALDI)	Sequência	Mascoat Ion score	Probabilidade do Scaffold	Precursor		Cobertura de sequência
						m/z	Erro (ppm)	
Albumina	gi 119626071	1 (A14)	(K)LVNEVTEFAK(T)	47,13	50%	1.149,67	1	50,5
			(K)FQNALLVR(Y)	31,44	95%	960,6196	1	58,79
			(R)FKDLGEEENFK(A)	51,7	95%	1.226,67	1	54,74
			(K)AVM _{ox} DDFAAFVEK(C)	35,45	95%	1.358,70	1	51,11
			(K)VHTEC _{cam} C _{cam} HGDLLC _{cam} ADDRDLAK(Y)	38,12	95%	2.585,28	1	64,31
			(K)LVNEVTEFAK(T)	45,63	43%	1.149,68	1	53,54
			(K)ADDKETC _{cam} FAEEGKK(L)	48,44	95%	1.627,80	1	47,11
			(R)FKDLGEEENFK(A)	57,65	95%	1.226,67	1	53,85
			(R)RPC _{cam} FSALEVDETYVPK(E)	101,3	95%	1.911,02	1	45,41
			(K)KVPQVSTPTLVEVSR(N)	47,91	51%	1.639,99	1	30,4
			(K)VFDEFKPLVEEPQNLK(Q)	85,7	95%	2.045,18	1	42,22
			(R)RHPDYSVLLLR(L)	26,37	94%	1.467,91	1	46,16
			(K)LYLEIAR(R)	43,91	95%	927,5439	1	53,86
			(K)VHTEC _{cam} C _{cam} HGDLLC _{cam} ADDRDLAK(Y)	46,64	95%	2.585,23	1	42,3
			(R)RPC _{cam} FSALEVDETYVPK(E)	36,84	95%	1.911,03	1	49,97
			(R)FKDLGEEENFK(A)	62,46	95%	1.226,68	1	62,08
			(K)AVM _{ox} DDFAAFVEK(C)	31,43	95%	1.358,70	1	53,46
α-1 anti-tripisina	gi 110350939	5 (A20)	(K)LYHSEAFVNEGDEEAKK(Q)	77,36	95%	2.186,14	1	44,4
			(R)KPVEEYANCC _{cam} HLAR(A)	58,88	95%	1.586,86	1	56,2
			(K)EGYGYTGAFR(C)	66,63	95%	1.283,65	1	61,35
			(R)KPVDEYKDC _{cam} HLAQVPSHTVVAR(S)	99,36	95%	2.549,42	1	51,38
			(K)W _{ox} C _{cam} ALSHHER(L)	50,48	95%	1.211,61	1	58,17
			(R)FDEFSEGC _{cam} APGSK(K)	34,42	95%	1.577,74	1	53,93
			(R)FDEFSEGC _{cam} APGSK(D)	60,75	95%	1.705,85	1	54,45
			(K)NLNEKDYELLC _{cam} LDGTR(K)	66,52	95%	1.953,04	1	51,81
			(K)ADRQYELLC _{cam} LDNTR(K)	50,69	95%	1.881,97	1	51,64
			(R)KPVEEYANCC _{cam} HLAR(A)	64,34	95%	1.586,86	1	52,8
			(R)KPVDEYKDC _{cam} HLAQVPSHTVVAR(S)	131,79	95%	2.549,43	1	53,77
			(K)EGYGYTGAFR(C)	52,71	95%	1.283,64	1	56,76
			(R)FDEFSEGC _{cam} APGSK(K)	40,84	95%	1.577,74	1	50,44
			(K)KASYLDC _{cam} IR(A)	32,48	95%	1.125,64	1	62
			(R)FDEFSEGC _{cam} APGSK(D)	67,38	95%	1.705,85	1	57,15
			(K)DSGFQMNQLR(G)	34,21	95%	1.195,62	1	55,08
			(K)NLNEKDYELLC _{cam} LDGTR(K)	134,11	95%	1.953,04	1	54,17
(-)EVLVESGGGLVQPGGSLR(L)	46,69	95%	1.882,09	1	47,05			
Transferrina	gi 110590599	11 (A22)						18%
Immunoglobulina	gi 121670547	11 (A22)						16%

Tabela 4.4: Continuação.

Nome da proteína	Código de acesso no NCBI	Número da banda (posição na placa do MALDI)	Sequência	Mascore (m score)	Probabilidade do Scaffold	Precursor		Cobertura de sequência
						m/z	Erro (ppm)	
Imunoglobulina A	gi 229537	8 (A18)	(R)WLQGSQELPR(E)	59,11	95%	1.213,69	1	50,22
			(K)TFTC _{cam} TAAYPEK(T)	39,13	95%	1.375,69	1	49,42
			(R)IQQPSQGITTFVAVTSILR(V)	89,25	95%	1.835,01	1	22,81
		9 (A19)	(R)WLQGSQELPR(E)	33,68	95%	1.213,68	1	35,89
		10 (A21)	(R)WLQGSQELPR(E)	45,09	95%	1.213,70	1	57,14
			(R)IQQPSQGITTFVAVTSILR(V)	86,56	95%	1.835,01	1	25,1
		11 (A22)	(R)WLQGSQELPR(E)	62,9	95%	1.213,71	1	63,57
			(R)IQQPSQGITTFVAVTSILR(V)	87,01	95%	1.835,00	1	17,74
		12 (A23)	(R)IQQPSQGITTFVAVTSILR(V)	92,58	95%	1.835,01	1	22,15
			(R)WLQGSQELPR(E)	49,79	95%	1.213,71	1	59,78
		13 (B23)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	50,71	63%	1.876,01	1	44,45
			(R)JTVAAPSVFIFPPSDEQLK(S)	59,31	95%	1.946,11	1	43,86
		14 (B24)	(K)VDNALQSGNSQESVTEQDSK(D)	38,35	95%	2.136,06	1	41,32
		15 (C1)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	75,16	95%	1.875,96	1	17,8
	(K)VDNALQSGNSQESVTEQDSK(D)	79,44	95%	2.136,00	1	15,29		
16 (C10)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	101,56	95%	1.875,98	1	28,99		
	(R)JTVAAPSVFIFPPSDEQLK(S)	58,88	95%	1.946,08	1	24,49		
	(K)VDNALQSGNSQESVTEQDSK(D)	114,09	95%	2.136,04	1	31,63		
17 (C11)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	70,49	95%	1.876,01	1	44,77		
	(R)JTVAAPSVFIFPPSDEQLK(S)	44,55	95%	1.946,10	1	37,59		
18 (C12)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	72,25	95%	1.875,97	1	20,68		
	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	94,94	95%	1.875,99	1	33,52		
19 (C3)	(R)JTVAAPSVFIFPPSDEQLK(S)	48,75	95%	1.946,09	1	30,5		
	(R)FSGSGGTDFTLISR(L)	59,95	95%	1.632,85	1	35,49		
20 (C4)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	78,82	95%	1.875,96	1	17,43		
21 (C5)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	95,36	95%	1.875,97	1	22,7		
	(K)VDNALQSGNSQESVTEQDSK(D)	112,92	95%	2.136,03	1	28,54		
22 (C8)	(K)IVYAC _{cam} EVTHQGLSSPVTK(S)	103,98	95%	1.875,98	1	27,61		

Imunoglobulina cadeia kappa

25%

Tabela 4.4.: Continuação.

Nome da proteína	Código de acesso no NCBI	Número da banda (posição na placa do MALDI)	Sequência	Mascof Ion score	Probabilidade do Scaffold	Precursor		Cobertura de sequência
						m/z	Erro (ppm)	
Receptor polimérico de imunoglobulina	gi B1377806	23 (A4)	(K)NADLQVLKPEPELVYEDLR(G)	97,79	95%	2.241,34	1	74,67
			(R)G S V T F H C _{cam} A L G P E V A N V A K (F)	85,58	95%	1.857,08	1	80,98
			(K)A F V N C _{cam} D E N S R (L)	63,23	95%	1.211,63	1	94,99
			(K)G V A G G S V A V L C _{cam} P Y N R (K)	38,81	95%	1.519,90	1	85,72
			(R)T V T I N C _{cam} P F K T E N A Q K (R)	52,29	95%	1.751,01	1	76,98
		(R)G G C _{cam} I T L I S S E G Y V S S K (Y)	40,45	95%	1.657,95	1	81,84	
		(K)A F V N C _{cam} D E N S R (L)	58,36	95%	1.211,61	1	79,22	
		(R)G S V T F H C _{cam} A L G P E V A N V A K (F)	64,41	95%	1.857,07	1	71,24	
		(R)T V T I N C _{cam} P F K T E N A Q K (R)	35,08	95%	1.751,01	1	73,89	
		(R)G G C _{cam} I T L I S S E G Y V S S K (Y)	72,83	95%	1.657,93	1	72,07	
		(K)NADLQVLKPEPELVYEDLR(G)	79,83	95%	2.241,33	1	67,44	
		(R)A D E G W Y W C _{cam} G V K (Q)	54,33	95%	1.370,69	1	76,38	
		(K)A F V N C _{cam} D E N S R (L)	63,59	95%	1.211,61	1	81,45	
		(R)G S V T F H C _{cam} A L G P E V A N V A K (F)	79,17	95%	1.857,06	1	67,47	
		(R)T V T I N C _{cam} P F K T E N A Q K (R)	28,09	81%	1.751,01	1	72,47	
(R)L S D A G Q Y L C _{cam} Q A G D D S N S N K K (N)	40,49	95%	2.171,10	1	61,53			
(K)G V A G G S V A V L C _{cam} P Y N R (K)	22,73	81%	1.519,89	1	76,64			
(R)T V T I N C _{cam} P F K T E N A Q K (R)	59,85	95%	1.751,01	1	74,52			
(R)G S V T F H C _{cam} A L G P E V A N V A K (F)	72,25	95%	1.857,06	1	68,65			
(R)L S D A G Q Y L C _{cam} Q A G D D S N S N K K (N)	95,47	95%	2.171,12	1	71,99			
(R)G G C I T L I S S E G Y V S S K (Y)	39,16	95%	1.657,91	1	59,77			
(K)Y T A C _{cam} L C D D N P K (T)	64,55	95%	1.356,59	1	25,6			
(R)E L G I C _{cam} P D D A A V I P I K N N R (N)	63,03	95%	1.610,90	1	30,66			
(R)E L G I C _{cam} P D D A A V I P I K N N R (F)	74,58	95%	1.995,08	1	25,61			
(K)Y T A C _{cam} L C D D N P K (T)	64,21	95%	1.356,60	1	29,87			
(K)Y T A C _{cam} L C D D N P K (T)	47,33	95%	1.356,61	1	41,08			
(R)E L G I C _{cam} P D D A A V I P I K N N R (F)	46,23	95%	1.995,10	1	31,42			
(K)Y T A C _{cam} L C D D N P K (T)	67,98	95%	1.356,61	1	40,85			
(K)T Y L I S S I P L Q G A F N Y K (Y)	36,91	95%	1.814,98	1	4,052			

As modificações estão simbolizadas: carbamidometilação de cisteína - Ccam; oxidação de metionina - Mox; oxidação de triptofano - Wox.

4.7. Depleção de proteínas abundantes do NAF por afinidade com proteína L

Os resultados anteriores mostraram claramente que as imunoglobulinas correspondem a uma fração importante das proteínas mais abundantes do NAF. Varnum e colaboradores estimaram que elas podem representar cerca de 30% do conteúdo proteico total deste fluido (Varnum *et al.*, 2003). Portanto, decidimos testar a depleção somente das imunoglobulinas do NAF usando proteína L agarose, proteína bacteriana recombinante que tem afinidade pelas cadeias leves do tipo kappa de todas as classes de imunoglobulina humana, incluindo IgG, IgM, IgA, IgE e IgD. A resina foi empacotada em colunas do tipo *micro-spin* e os procedimentos de aplicação de amostra, lavagem da coluna e eluição da fração ligada foram feitos sob centrifugação suave. Todas as frações cromatográficas foram quantificadas por leitura de absorvância a 280 nm e submetidas à análise por SDS-PAGE (Figura 4.9). Somente as bandas reveladas na fração que interagiu com a proteína L foram tripsinizadas e identificadas por espectrometria de massas. Todas as identificações indicaram a presença apenas de imunoglobulina, sugerindo que a depleção do NAF utilizando proteína L recombinante ocorreu de modo específico (Tabela 4.5).

Alíquotas depletadas e sem depleção de uma mesma amostra de NAF (P5d) foram submetidas à eletroforese bidimensional a fim de comparar o ganho em termos de número e resolução de *spots* (Figura 4.10). No gel referente à fração depletada (Figura 4.10A) foram detectados 112 *spots* e no gel da amostra bruta (Figura 4.10B), 169 *spots*. Como esperado, observamos que as regiões do gel onde se localizam as cadeias pesada e leve das imunoglobulinas (indicadas em vermelho na Figura 4.10) apresentam-se menos intensamente coradas no gel depletado. No entanto, não houve melhora significativa no perfil bidimensional como um todo. Muito pelo contrário, a diminuição do número de *spots* visualizados foi bastante expressiva na amostra depletada. Após quatro experimentos com amostras diferentes de NAF (P5d, P5e, P10d e P11d), a perda de proteína total no processo de depleção estimada pela leitura espectrofotométrica foi de $36 \pm 3\%$. Tomando como 100% o total de proteína recuperado, observamos que a quantidade de imunoglobulina (proteínas que interagiram com a proteína L) variou entre 5 e 18% nas diferentes amostras. Considerando a perda significativa de material, sem incremento da qualidade do perfil eletroforético de NAF, decidimos abandonar a idéia de depleção das imunoglobulinas, mantendo o processamento da amostra o mais simples possível. Para tentar identificar as proteínas menos abundantes do NAF, passamos a testar a detecção das proteínas do gel

bidimensional utilizando fluorescência, metodologia mais sensível e com maior intervalo dinâmico linear do que a revelação pelo Coomassie coloidal.

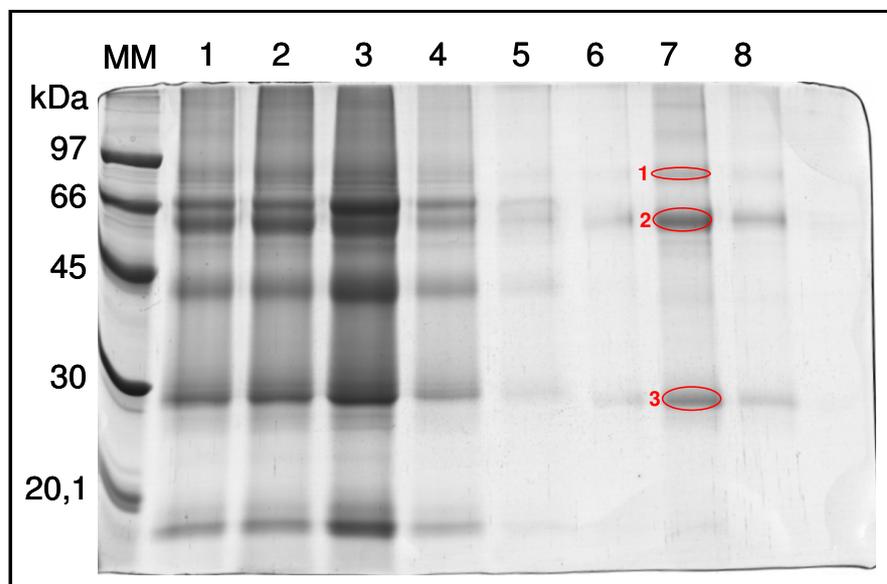


Figura 4.9: SDS-PAGE das frações de NAF de uma amostra controle (P5e) eluídas da coluna de proteína L agarose. Raias 1 e 2, NAF bruto sem depleção; raias 3-5, frações da lavagem da coluna (*flowthrough*, proteínas não-ligadas); raias 6-8, frações eluídas da coluna (proteínas que interagiram com a proteína L). As bandas assinaladas em vermelho e numeradas de 1 a 3 correspondem àquelas identificadas por espectrometria de massas (MALDI-TOF/TOF) como imunoglobulinas. Foram aplicados 20 μ L de cada fração cromatográfica. Os géis foram corados com azul de Coomassie G-250 coloidal. MM, padrão de massa molecular.

Tabela 4.5: Identificações por MALDI TOF/TOF das bandas de eletroforese unidimensional da fração do NAF que interage com a Proteína L.

Nome da proteína	Código de acesso no NCBI	Número da banda (posição na placa do MALDI)	Sequência	Mascot / <i>ion score</i>	Probabilidade do Scaffold	Precursor			Cobertura de sequência
						m/z	Carga	Erro (ppm)	
Receptor polimérico de imunoglobulina	gi 219109287	1 (D3)	(K)DGFSFVITGLR(K)	67,65	95%	1250,70	1	17,55	6%
			(K)QGHFYGETAAVYVAVEER(K)	72,50	95%	2026,01	1	21,20	
			(K)NADLQVLKPEPELVYEDLR(G)	68,06	95%	2241,22	1	20,41	
Imunoglobulina A	gi 2135473	2 (D4)	(R)WLGSGQELPR(E)	82,40	95%	1213,65	1	11,91	11%
			(R)QEPSQGTTFVAVTSILR(V)	103,19	95%	1835,96	1	3,41	
			(K)SAVQGGPPER(D)	45,15	52%	940,49	1	4,28	
Imunoglobulina cadeia kappa	gi 10835792	3 (D5)	(K)VVAC ^{cam} EVTHQGLSSPVTK(S)	94,75	95%	1875,97	1	21,26	17%
			(R)TVAAPS ^{cam} VIFPPSDEQLK(S)	73,99	95%	1946,07	1	20,48	
			(K)VDNALQSGNSQESVTEQDSK(D)	101,88	95%	2136,02	1	23,39	

A modificação está simbolizada: carbamidometilação de cisteína - C^{cam}.

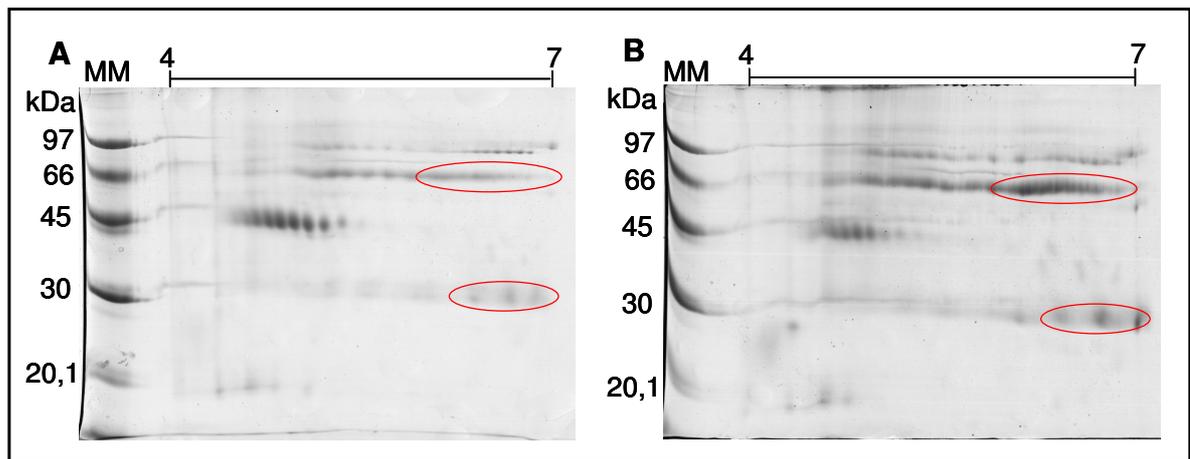


Figura 4.10: 2D-PAGE da mesma amostra de NAF de um indivíduo controle (P5d) após a depleção das imunoglobulinas por proteína L (A) e sem depleção (amostra bruta) (B). As cadeias pesada e leve das imunoglobulinas estão indicadas em vermelho. Foram aplicados 150 µg de proteína em cada gel, em fitas de IEF de 7 cm, com intervalos de pH 4-7. Os géis foram revelados por impregnação com nitrato de prata. MM, padrão de massa molecular.

4.8 – Análise de amostras pareadas de NAF por DIGE

Na eletroforese de fluorescência diferencial, dois grupos de amostras (ex.: controle saudável e câncer) são marcados com fluoróforos espectralmente independentes (Cy3 e Cy5), mas estruturalmente similares (Unlu *et al.*, 1997). Em seguida, as amostras são misturadas e submetidas ao mesmo gel de eletroforese bidimensional. Um padrão interno marcado com um terceiro fluoróforo (Cy2) também é aplicado no mesmo gel (Alban *et al.*, 2003). Trata-se de um *pool* preparado a partir de partes iguais de cada uma das amostras do estudo, contendo, portanto, o universo das proteínas de todos os indivíduos. Ele é utilizado para normalizar as quantidades de amostra nos diferentes géis, além de facilitar a sobreposição das imagens dos géis sendo comparados. Esta abordagem reduz o número de géis por experimento, diminui a variabilidade experimental e torna mais confiável o pareamento entre os géis e a quantificação dos *spots* proteicos, simplificando a comparação entre amostras (Marouga *et al.*, 2005). O DIGE é uma técnica de revelação com grande intervalo dinâmico linear (4 ordens de grandeza) e muito sensível, com limite mínimo de detecção de cerca de 1 ng de proteína por *spot* (Görg *et al.*, 2004).

Nesta etapa do estudo, as 14 amostras de NAF (sem qualquer depleção) das mamas controle e doente de sete pacientes com câncer de mama unilateral foram analisadas por DIGE. Elas foram marcadas alternadamente com Cy3 ou Cy5 (Tabela 4.6), para evitar a introdução de viés em caso de marcação preferencial das amostras por um dos fluoróforos. O padrão interno, constituído pela mistura da mesma quantidade de proteína de cada uma das 14 amostras de NAF, foi marcado com Cy2. As duas amostras marcadas com Cy3 (50 µg) e Cy5 (50 µg) de cada paciente e uma amostra do padrão interno (50 µg) foram analisadas conjuntamente no mesmo gel bidimensional.

As imagens obtidas através do scanner de fluorescência *Typhoon Trio* foram analisadas no módulo DIA do *software DeCyder*, que faz a co-deteção e a normalização dos volumes dos *spots* das duas amostras de cada gel em relação ao seu padrão interno. As imagens do *pool* interno se mostraram muito semelhantes em todos os géis, indicando que a metodologia de análise foi reprodutiva (Figura 4.11). De modo interessante, as duas mamas de cada uma das pacientes mostraram perfis de expressão proteica semelhantes, independentemente da presença da doença (Figuras 4.12 e 4.13). Por outro lado, diferenças significativas foram detectadas entre os padrões de imagem bidimensional das várias pacientes, indicando

variabilidade individual importante. Considerando o padrão de distribuição de todos os *spots* em cada gel, observamos vários perfis proteicos diferentes. Entretanto, para simplificar a análise, resolvemos classificar os perfis dos géis em dois grandes grupos: o perfil A foi visualizado em todas as imagens do padrão interno (Figura 4.11) e nas imagens de ambas as mamas de quatro pacientes. Ele é equivalente às imagens dos demais géis bidimensionais feitos nas etapas anteriores deste estudo, apresentando dois grupos típicos de *spots* ácidos ($pI \approx 4,5$), com massas moleculares de 29,5 e 18 kDa (Figura 4.12, grupos de *spots* # 1-2, respectivamente); O perfil B foi observado nas amostras de NAF das outras três pacientes e é caracterizado pela ausência dos dois grupos típicos de *spots* do perfil A (Figura 4.13). As heterogeneidades observadas dentro de cada grupo não mostraram correlação aparente com as características demográficas das pacientes (Tabela 4.1, vide item 4.1).

Como as imagens das mamas de algumas pacientes (P34, P35 e P36, Figura 4.13) diferiram bastante do perfil A típico do padrão interno, não foi possível fazer a co-deteção dos *spots* nas três imagens de cada um dos três géis destas amostras. Esta etapa é fundamental para possibilitar a normalização dos volumes dos *spots* e o pareamento posterior das imagens nos diferentes géis. Desta forma, ainda que qualitativamente tenhamos detectado diferenças importantes, infelizmente não foi possível utilizar a técnica de DIGE para comparar quantitativamente a expressão diferencial de proteína entre as 14 amostras do estudo.

Tabela 4.6: Amostras pareadas de mamas controle e mamas com câncer utilizadas na 2D-DIGE.

Código da amostra	Grupo	Concentração de proteína ($\mu\text{g}/\mu\text{L}$)	Marcação
P29e	Controle	141,1	Cy3
P30e	Controle	79,3	Cy5
P31d	Controle	70,1	Cy3
P33e	Controle	40,3	Cy5
P34e	Controle	246,4	Cy3
P35d	Controle	142,1	Cy5
P36d	Controle	133,8	Cy3
P29d	Câncer	141,9	Cy5
P30d	Câncer	107,7	Cy3
P31e	Câncer	69,3	Cy5
P33d	Câncer	63,4	Cy3
P34d	Câncer	52,9	Cy5
P35e	Câncer	255,6	Cy3
P36e	Câncer	82,5	Cy5

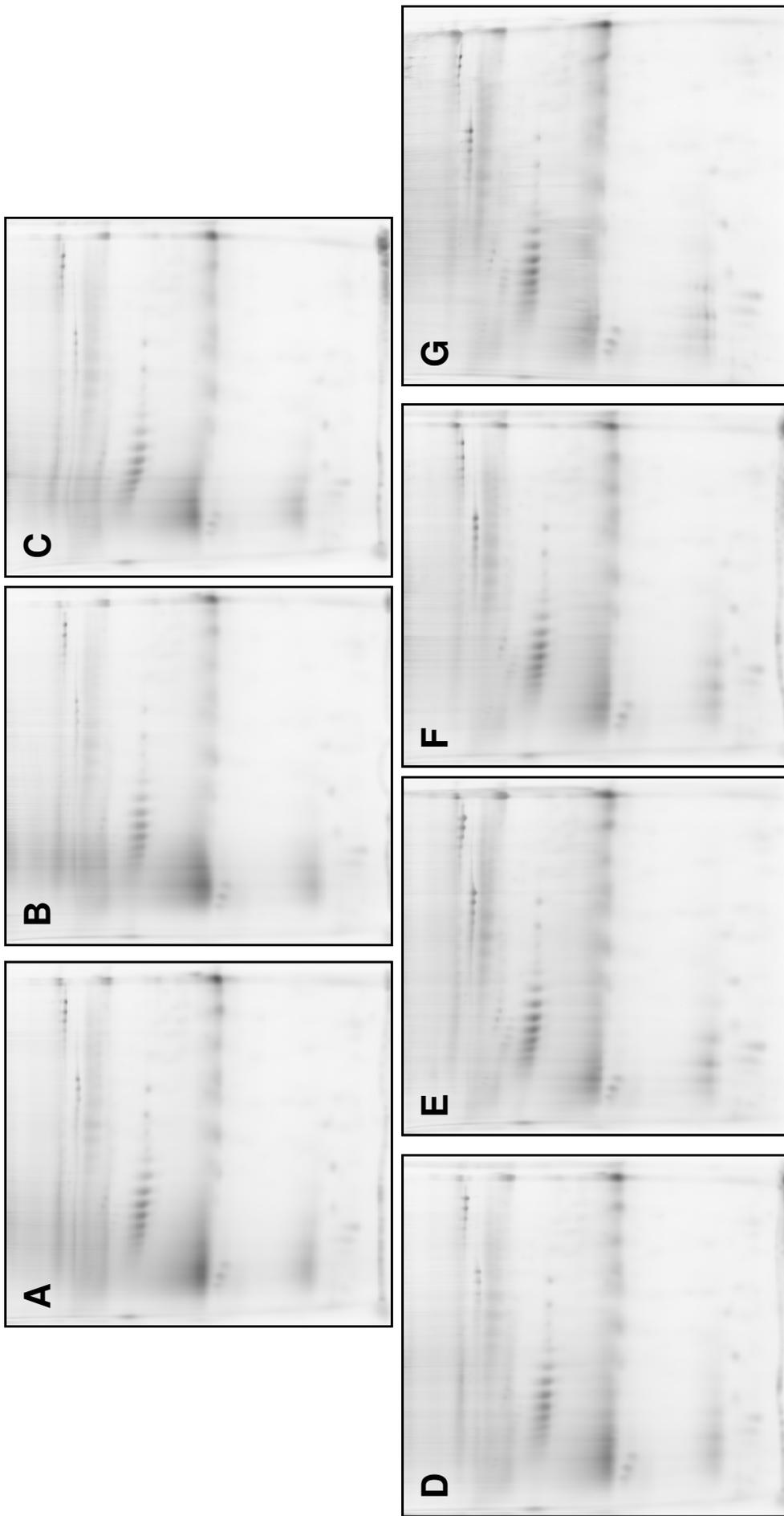


Figura 4.11: Imagens da amostra de padrão interno marcado com Cy2 dos sete géis usados no experimento de DIGE. Foram aplicados 50 μg do mesmo *pool* em cada gel, utilizando reidratação *in gel* e fitas de focalização isoeletrica de 18 cm, com intervalo de pH 4-7. As imagens dos géis referentes às amostras das seguintes pacientes estão mostradas: P29 (A), P30 (B), P31 (C), P33 (D), P34 (E), P35 (F), P36 (G).

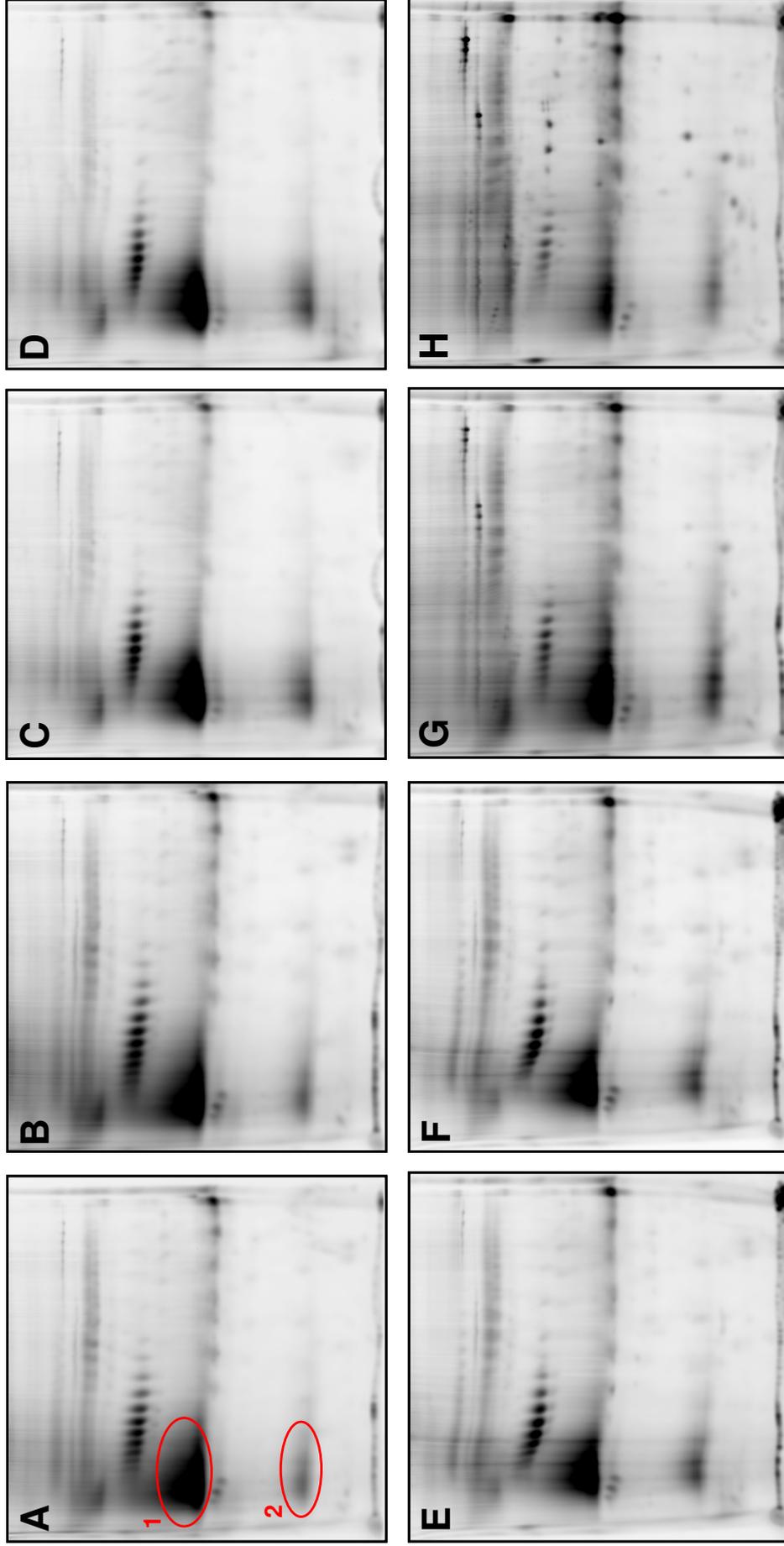


Figura 4.12: Imagens de DIGE das amostras de NAF classificadas no perfil A. Paciente P 29: (A) Mama controle marcada com Cy3 e (B) Mama com câncer marcada com Cy5. Paciente P 30: (C) Mama controle com Cy5 e (D) Mama com câncer com Cy3. Paciente P 31: (E) Mama controle com Cy3 e (F) Mama com câncer com Cy5. Paciente P 33: (G) Mama controle marcada com Cy5 e (H) Mama com câncer marcada com Cy3. Foram aplicados 50 µg de cada amostra (controle e câncer) por reidratação *in gel* em fitas de focalização isoeletrica de 18 cm, com intervalo de pH 4-7. Os grupos de spots 1-2 mostrados no painel A em vermelho são característicos do perfil A de NAF.

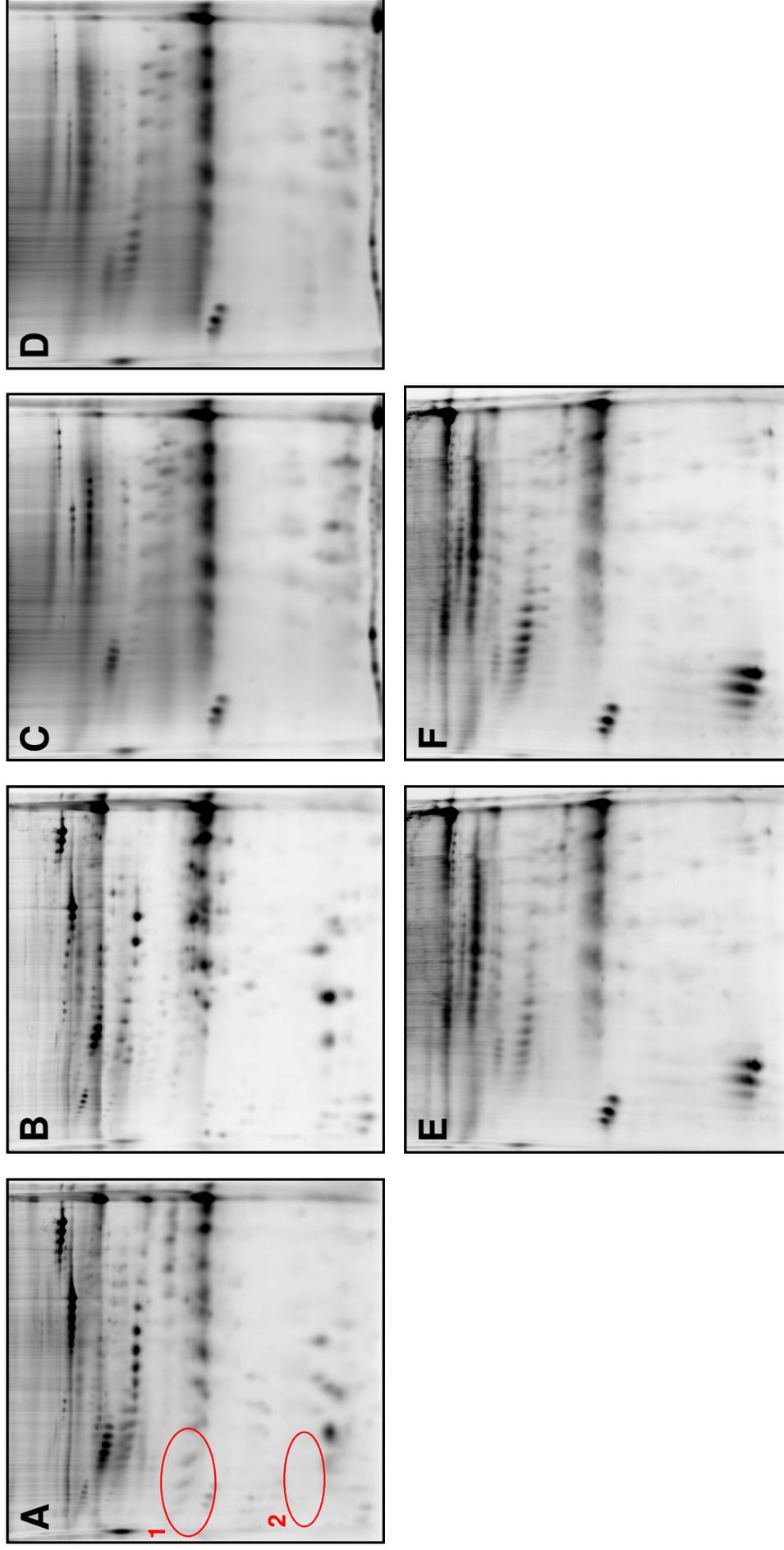


Figura 4.13: Imagens de DIGE das amostras de NAF classificadas no perfil B. Paciente P34: (A) Mama controle com Cy3 e (B) Mama com câncer com Cy5. Paciente P35: (C) Mama controle marcada com Cy5 e (D) Mama com câncer marcada com Cy3. Paciente P36: (E) Mama controle com Cy3 e (F) Mama com câncer com Cy5. Foram aplicados 50 μg de cada amostra (controle e câncer) por reidratação *in gel* em fitas de focalização isoeletrica de 18 cm, com intervalo de pH 4-7. Os grupos de spots 1-2 mostrados no perfil A em vermelho (Figura 4.12) estão ausentes no perfil B de NAF, como indicado em vermelho.

Na literatura, há poucas análises de NAF de doadoras saudáveis ou de pacientes com câncer de mama por 2D-PAGE. Nenhuma delas usou DIGE, a maioria utilizou um número amostral bastante reduzido ($n < 3$) e somente um trabalho analisou qualitativamente 20 amostras por gel bidimensional, mas não fez qualquer análise estatística das imagens dos géis. Alguns sugeriram que há variação na expressão protéica do NAF entre mulheres com doença benigna ou câncer, ainda que o perfil entre as mamas de cada mulher se mantenha constante (Klein *et al.*, 2001; Kuerer *et al.*, 2002; Alexander *et al.*, 2004). Kuerer *et al.* (2002) avaliaram as mamas sadias e doentes de três pacientes e encontraram diferenças qualitativas substanciais nos padrões de expressão entre as mamas de cada paciente, além de diferenças entre pacientes. Nossos resultados de DIGE mostraram que há variação no perfil de expressão proteica entre as pacientes, ainda que as duas mamas de uma mesma paciente tenham apresentado perfis comparáveis na maioria dos casos. Por outro lado, nossos zimogramas sugeriram que mama doente pode apresentar níveis de atividade proteolítica aumentado em relação ao encontrado na mama normal da mesma paciente (Figura 4.4). Como não identificamos qualquer metaloprotease nos *spots* principais do NAF, esta(s) enzima(s) deve(m) representar proteína(s) minoritária(s) da secreção. Portanto, considerando os componentes menos abundantes do NAF, é bastante provável que existam diferenças entre as duas mamas de uma mesma paciente, mas que não foram detectadas pela metodologia utilizada.

Estudos com culturas de células cancerosas de mama sugeriram que hormônios podem influenciar a expressão de proteínas associadas à atividade apócrina da glândula mamária, citado por Huang *et al.* (2007) (Huang *et al.*, 2007). Chatterton *et al.* (2004) mediram a concentração de estradiol no NAF de mulheres sadias pré-menopausa em pontos fixos do ciclo menstrual ao longo de 15 meses. Eles encontraram alta estabilidade na concentração de estradiol ao longo do tempo, indicando que fatores ambientais têm pouco efeito imediato nos níveis de estradiol no NAF. Petrakis *et al.* (1987) verificaram que não houve correlação significativa entre os níveis de estrógenos no NAF e o ciclo menstrual, o *status* de pré- e pós-menopausa ou a presença de doença (benigna ou câncer) (Petrakis *et al.*, 1987). Porém, a alta concentração de esteróides no NAF (colesterol e seus produtos de oxidação) apresentou correlação com doença proliferativa benigna da mama (Ernster *et al.*, 1987). Algumas proteínas reguladas por hormônios sexuais detectáveis no NAF, tais como a pS2 e apolipoproteína D, apresentaram variações em função das diferenças na concentração de

estrógeno circulante no soro por terapias hormonais (goserelina e tamoxifeno) ou pela reposição hormonal (Harding *et al.*, 2000). Outro estudo avaliou as variações do perfil proteômico do NAF ao longo do ciclo menstrual em mulheres pré-menopausa saudáveis. O perfil das proteínas do NAF não variou substancialmente com o ciclo menstrual, mas houve grandes diferenças entre as pacientes (Noble *et al.*, 2007). Portanto, os níveis de hormônios no NAF parecem estáveis ao longo do ciclo menstrual e as variações hormonais sistêmicas parecem ter pouca influência no perfil proteico do NAF.

Utilizando 2D-DIGE, definimos dois perfis bidimensionais bem característicos de amostras de NAF de pacientes com câncer de mama unilateral. As amostras de mulheres sem a doença (P5, P7 e P11) ou com tumor benigno (P4 e P10) e de cinco das oito pacientes (P2, P29, P30, P31 e P33) apresentam o mesmo padrão. Este perfil A foi considerado o mais clássico do NAF. A classificação do NAF em diferentes perfis eletroforéticos havia sido mostrada anteriormente por SDS-PAGE (Sanchez *et al.*, 1992). Os autores relataram que as pacientes com câncer de mama que apresentaram o padrão de bandas de eletroforese do Tipo II (vide item 4.2) mostravam melhor prognóstico quando comparado com aquelas do Tipo I (Vizoso *et al.*, 2001). Na presente dissertação, mostramos claramente que a complexidade do NAF é maior do que o suposto por SDS-PAGE. É possível que análises proteômicas com número maior de pacientes revele ainda a existência de vários outros perfis de expressão proteica diferentes.

Para tentar compreender melhor a origem da variabilidade individual das amostras de NAF, os géis de DIGE foram corados por Coomassie coloidal e os grupos de *spots* característicos de cada perfil foram excisados, tripsinizados, dessalinizados e submetidos à análise por MALDI-TOF/TOF. Infelizmente, não conseguimos qualquer identificação, possivelmente em função da quantidade limitada de material. Novas tentativas serão feitas utilizando mais material.

4.9. Análise do conteúdo de glicoproteínas do NAF

Os diferentes *spots* característicos dos perfis A e B podem corresponder a proteínas diferentes ou representar diferentes isoformas das mesmas proteínas. Por exemplo, alguns estudos sugerem que muitas diferenças entre amostras sadias e com câncer ocorrem devido à glicosilação, uma modificação pós-traducional que pode influenciar as interações célula-célula em favor da metástase, facilitando o escape das células cancerosas à vigilância do sistema imunológico (Dwek *et al.*, 2001). Alguns tumores benignos de mama expressam

tipos de glicosilação que parecem contribuir para a proliferação e transformação celulares (Gallegos *et al.*, 2010). No entanto, apesar de promissoras, as análises proteômicas em tecido mamário sofrem com a heterogeneidade dos vários tipos celulares que compõem essas amostras. O fluido aspirado de mama, no entanto, contém proteínas secretadas pelo epitélio mamário nas suas formas totalmente processadas, permitindo a análise das mudanças dos níveis de expressão de proteínas e também dos padrões de modificação pós-traducional associados à doença (Alexander *et al.*, 2004).

Para analisar o conteúdo de glicoproteínas do NAF, utilizamos quatro amostras de duas pacientes (duas de mamas controle pareadas com as mamas com câncer) que representam os perfis A ou B. As amostras foram separadas por 2D-PAGE, os géis foram revelados com *Pro-Q Emerald 488 (Glycoprotein Gel and Blot Stain Kit, Invitrogen, EUA)* e as imagens digitalizadas no scanner *Typhoon Trio*. O Pro-Q Emerald é um corante fluorescente que se liga especificamente a carboidratos oxidados por periodato. Posteriormente, os mesmos géis foram revelados com Coomassie G-250 coloidal para detecção de proteínas totais e as imagens digitalizados no *ImageScanner III*. As duas imagens obtidas de cada gel (detecções por *Pro-Q Emerald* e por Coomassie G-250 coloidal) foram sobrepostas e comparadas utilizando o programa *Image Master 7.0*. A sobreposição das duas imagens de 2D-PAGE de cada uma das pacientes mostrou que a glicosilação deve contribuir de modo importante para a heterogeneidade do NAF. Na paciente com perfil A (Figura 4.14), foram detectados 97 *spots* proteicos em ambas as mamas, sendo 21 glicosilados ($\approx 22\%$). Na paciente com perfil B (Figura 4.15), foram detectados cerca de 143 *spots* nas duas mamas, sendo 73 glicosilados ($\approx 51\%$). Os perfis de glicoproteínas, assim como os perfis de DIGE, foram semelhantes entre as mamas da mesma paciente, independente da presença de doença ou não, mas variaram entre as pacientes.

Apenas dois estudos de NAF determinaram por imunensaio a concentração de proteínas glicosiladas associadas à adesão, ao crescimento e à migração celulares, tais como P-caderina, antígenos TF (Thomsen-Friedenreich) e Tn (Kumar *et al.*, 2005; Mannello *et al.*, 2008). Os resultados indicaram altas concentrações destas proteínas no NAF de mamas com câncer e quantidades praticamente indetectáveis no fluido de mamas saudáveis. Portanto, é possível que a simples medida do conteúdo de determinadas glicoproteínas do NAF venha a ser útil para o prognóstico ou para o diagnóstico não invasivo do câncer de mama.

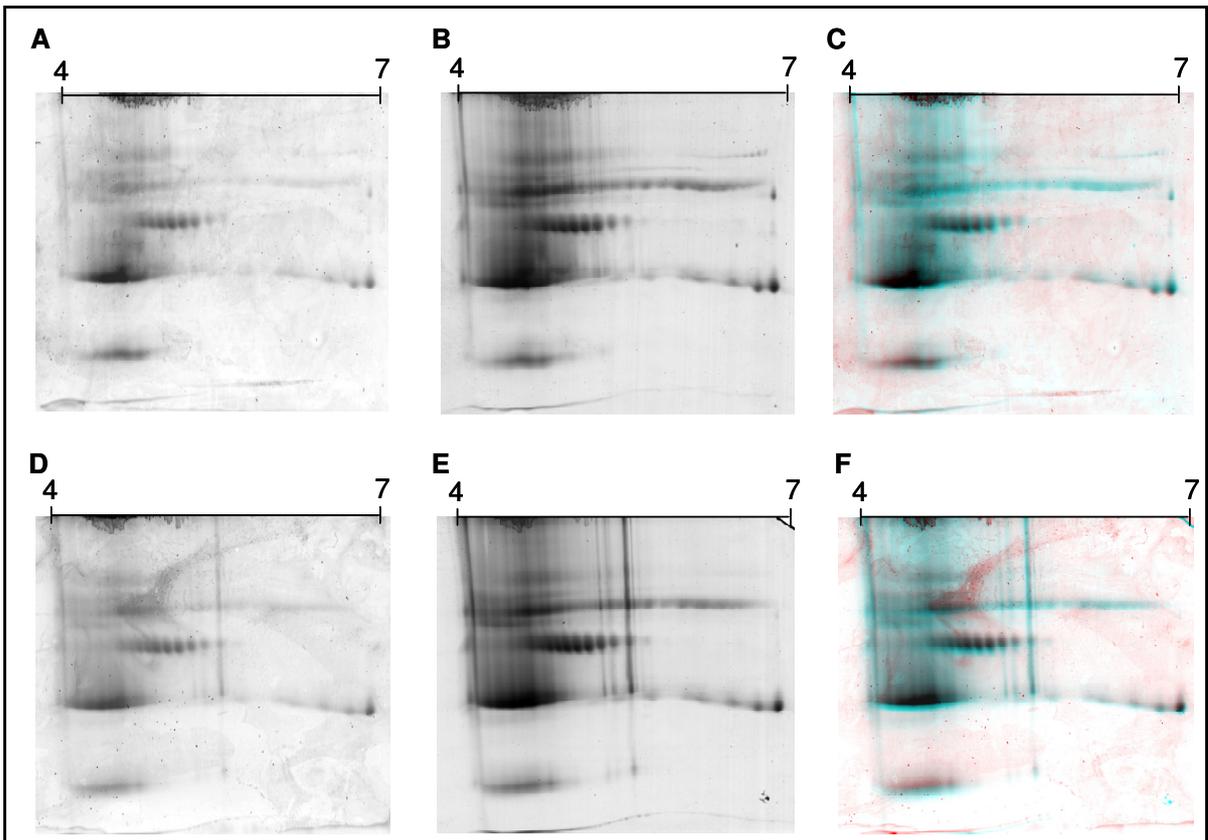


Figura 4.14: 2D-PAGE de NAF das mamas da paciente P29 que apresenta o perfil A. Gel da mama sadia (P29e) corado com *Pro-Q Emerald* (A) e com Coomassie G-250 (B). A sobreposição das duas imagens (*Pro-Q Emerald* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel C. Gel da mama com câncer (P29d) corado com *Pro-Q Emerald* (D) e com Coomassie G-250 (E). A sobreposição das duas imagens (*Pro-Q Emerald* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel F. Os spots comuns às duas imagens estão mostrados em preto nos painéis C e F. Cada imagem contém 150 µg de proteína aplicados por reidratação *in gel* em fitas de IEF de 7 cm, com intervalo de pH 4-7.

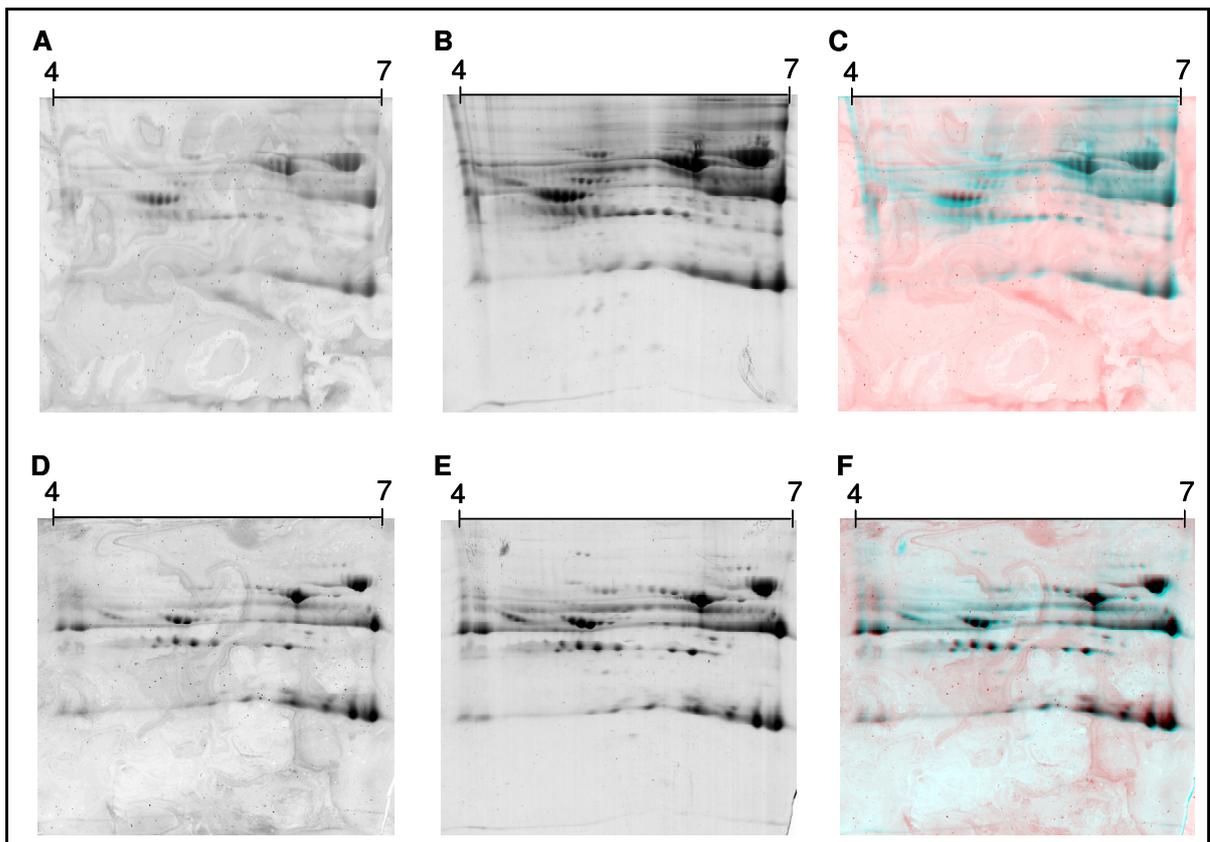


Figura 4.15: 2D-PAGE de NAF das mamas da paciente P34 que apresenta o perfil B. Gel da mama sadia (P34e) corado com *Pro-Q Emerald* (A) e com Coomassie G-250 (B). A sobreposição das duas imagens (*Pro-Q Emerald* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel C. Gel da mama com câncer (P34d) corado com *Pro-Q Emerald* (D) e com Coomassie G-250 (E). A sobreposição das duas imagens (*Pro-Q Emerald* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel F. Os spots comuns às duas imagens estão mostrados em preto nos painéis C e F. Cada imagem contém 150 µg de proteína aplicados por reidratação *in gel* em fitas de IEF de 7 cm, com intervalo de pH 4-7.

4.10. Análise do conteúdo de fosfoproteínas do NAF

Várias proteínas de sinalização celular, quando se encontram no estado fosforilado, participam de vias de transdução de sinal fundamentais em processos de crescimento celular e diferenciação (Alberts *et al.*, 1994). Alguns receptores celulares relacionados ao câncer de mama, como o receptor tipo 2 do fator de crescimento epidérmico humano (HER-2) e o receptor de estrogênio (ER), utilizam a fosforilação para a sinalização de processos de proliferação celular. O HER-2 é uma proteína transmembrana tirosino-quinase que fosforila moléculas receptoras citoplasmáticas para ativação de vias promotoras do crescimento; sua superexpressão está associada a um pior prognóstico no câncer de mama (Roy *et al.*, 2009). Através do *crossstalk* entre o estrógeno e alguns fatores de crescimento no microambiente tumoral da mama as vias de ativação do HER-2 também ativam o ER por fosforilação. Como resultado, vários genes promotores do crescimento celular e, conseqüentemente, do crescimento do tumor, são ativados por fatores de transcrição (Yamaguchi, 2007).

Para analisarmos o conteúdo de fosfoproteínas no NAF, utilizamos as mesmas amostras dos perfis A e B das análises glicoproteicas. Desta vez, os géis bidimensionais foram revelados com *Pro-Q Diamond (Phosphoprotein Gel Stain, Invitrogen, USA)*. Trata-se de um corante fluorescente que se liga a grupamentos fosfato em resíduos de tirosina, serina ou treonina. Após digitalização das imagens no scanner *Typhoon Trio*, os mesmos géis foram revelados com Coomassie G-250 coloidal para detecção das proteínas totais. As duas imagens obtidas nas diferentes revelações foram sobrepostas e comparadas como descrito acima (item 4.9). A revelação com *Pro-Q Diamond* mostrou que a amostra da paciente com o perfil A não apresentou *spots* detectáveis de fosfoproteínas (Figura 4.16). No caso do perfil B, foram detectados cerca de 38 *spots* entre 63 e 66 kDa, na região mais ácida do gel (Figura 4.17). Utilizamos como referência de massa molecular a banda da ovalbumina do padrão de massa molecular (vide item 3.4) que é uma fosfoproteína de 45 kDa revelada pelo *Pro-Q Diamond*.

As pacientes estudadas apresentaram nenhuma (P29) ou poucas (P34) proteínas fosforiladas no NAF de suas duas mamas. É importante lembrar que as amostras de NAF utilizadas neste estudo representam apenas as proteínas secretadas no sistema de ductos da mama, sem contaminação de células esfoliadas, que foram separadas por centrifugação no momento da coleta. Sabe-se que a maioria das proteínas fosforiláveis constituem proteínas celulares que utilizam este mecanismo para o controle de suas atividades de sinalização celular,

sinalização transmembrana, amplificação intracelular de sinais, controle do ciclo celular, entre outras funções exercidas no citoplasma e no núcleo (Alberts *et al.*, 1994). A maioria dos estudos que relaciona fosfoproteínas com o câncer de mama se restringe a proteínas transmembranares ou de cascatas intracelulares (Yamaguchi, 2007; Caiazza *et al.*, 2010). Segundo Alexander *et al* (2004), as proteínas secretadas no NAF seriam totalmente processadas e glicosiladas não sofrendo influência de modificações transitórias, como a fosforilação, que afetaria apenas as proteínas dentro das células. No entanto, já foram encontradas no NAF de mulheres americanas e quenianas sem câncer algumas caseínas, fosfoproteínas predominantemente encontradas no leite materno (Varnum *et al.*, 2003). Interessantemente, uma das pacientes do nosso estudo apresentou *spots* fosforilados no NAF das duas mamas com massa molecular diferente das caseínas (≈ 23 kDa). Para complementar nossos resultados, pretendemos ainda identificar estes *spots* por espectrometria de massas.

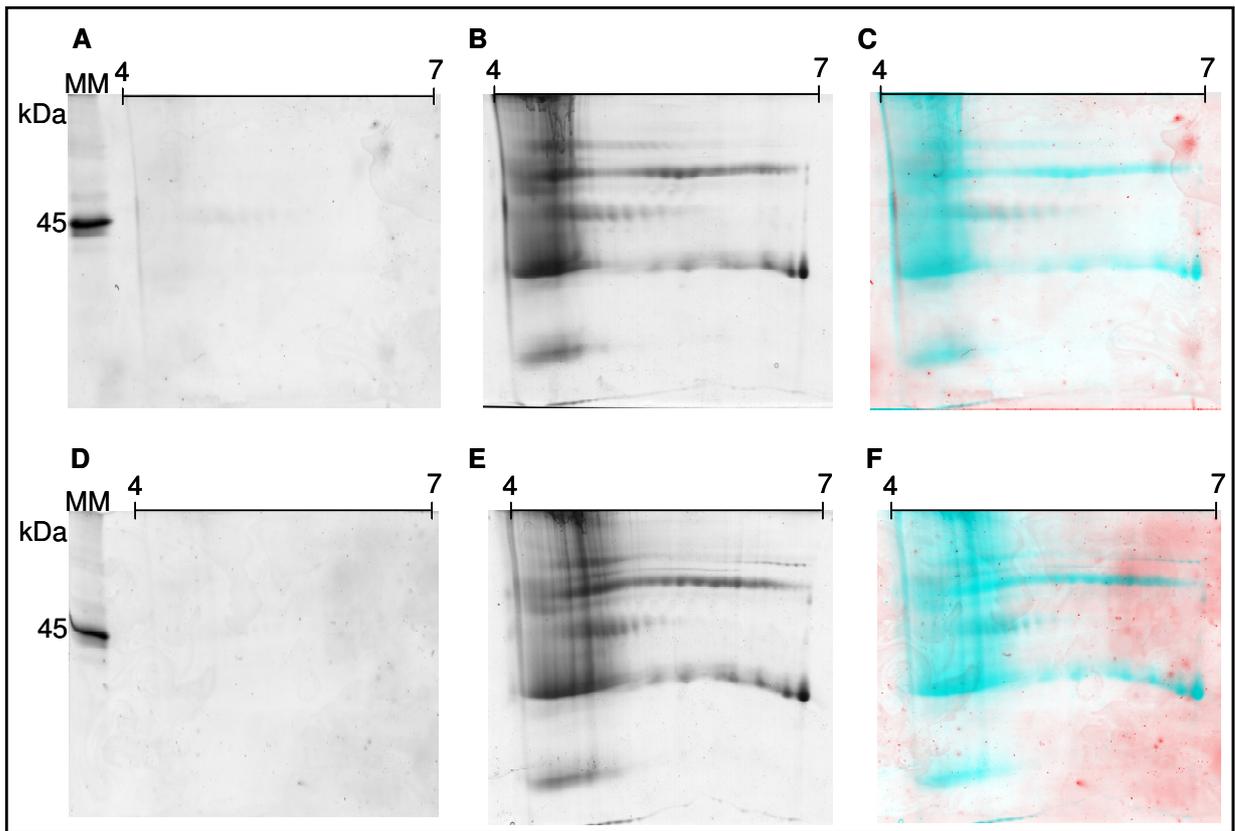


Figura 4.16: 2D-PAGE de NAF das mamas da paciente P29 que apresenta o perfil A. Gel da mama sadia corado com *Pro-Q Diamond* (A) e com Coomassie G-250 (B). A sobreposição das duas imagens (*Pro-Q Diamond* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel C. Gel da mama com câncer corado com *Pro-Q Diamond* (D) e com Coomassie G-250 (E). A sobreposição das duas imagens (*Pro-Q Diamond* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel F. Os *spots* comuns às duas imagens (AxB e Dx E) não foram visualizados. Cada imagem contém 150 μg de proteína aplicados por reidratação *in gel* em fitas de IEF de 7 cm, com intervalo de pH 4-7.

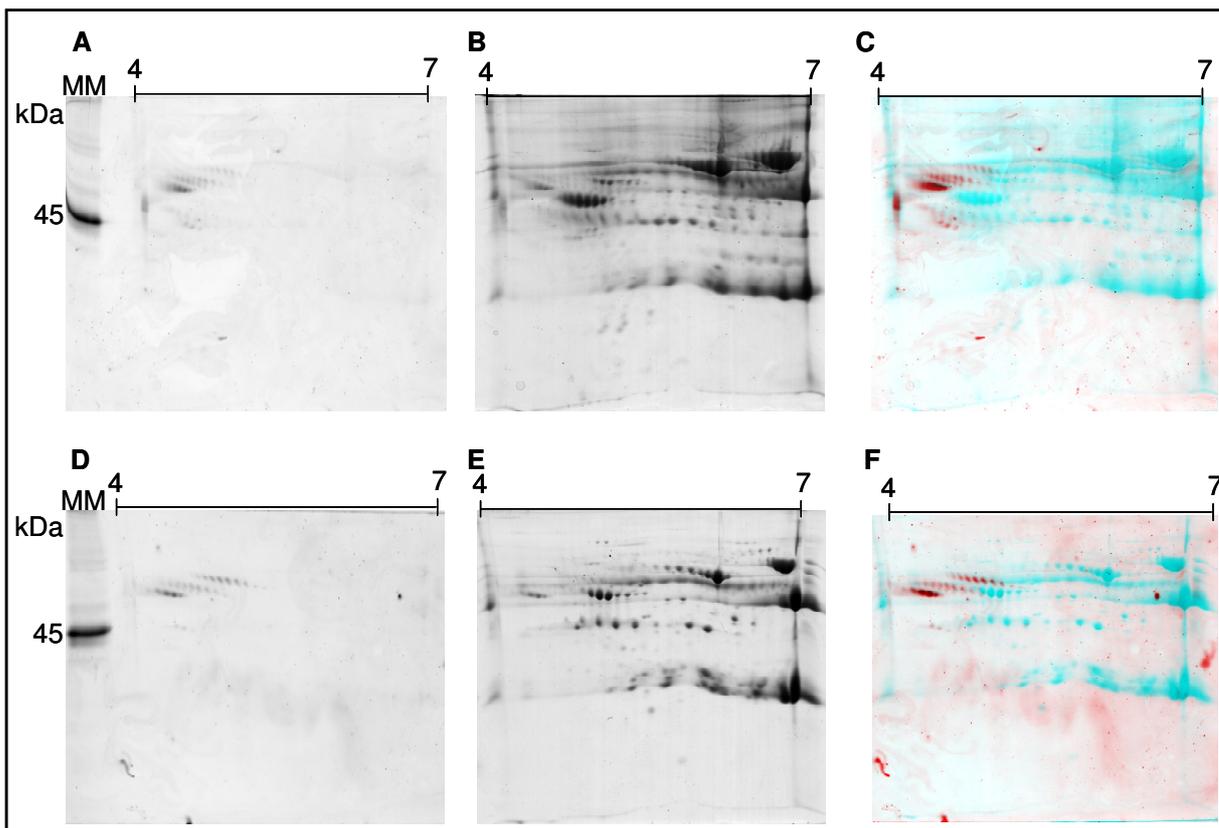


Figura 4.17: 2D-PAGE de NAF das mamas da paciente P34 que apresenta o perfil B. Gel da mama sadia corado com *Pro-Q Diamond* (A) e com Coomassie G-250 (B). A sobreposição das duas imagens (*Pro-Q Diamond* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel C. Gel da mama com câncer corado com *Pro-Q Diamond* (D) e com Coomassie G-250 (E). A sobreposição das duas imagens (*Pro-Q Diamond* em vermelho e Coomassie G-250 coloidal em azul) está mostrada no painel F. Os *spots* comuns às duas imagens estão mostrados em preto nos painéis C e F. Cada imagem contém 150 μg de proteína aplicados por reidratação *in gel* em fitas de IEF de 7 cm, com intervalo de pH 4-7.

4.11. Perspectivas futuras: caracterização do NAF por LC-MS/MS

O tamanho amostral utilizado nos caros e trabalhosos estudos proteômicos comparativos é normalmente bastante reduzido. Isto pode prejudicar seriamente a qualidade das conclusões obtidas, mesmo considerando que se trata normalmente apenas de uma fase exploratória inicial (*discovery phase*) (Zolg, 2006). Por exemplo, dos poucos trabalhos de 2D-PAGE do NAF já publicados, a maioria comparou em média cerca de três indivíduos (Klein *et al.*, 2001; Kuerer *et al.*, 2002). No nosso trabalho, as comparações foram feitas com sete amostras em cada grupo, já que tivemos muita dificuldade para obter quantidades suficientes de amostras pareadas, como já discutido no item 4.1. Ainda assim, o estudo foi importante para aumentar nosso conhecimento qualitativo acerca da composição proteica das amostras de NAF.

Salvo engano, até então não existiam trabalhos publicados que descrevessem o estudo sistemático do NAF utilizando pacientes brasileiras com câncer de mama unilateral. Esta amostra biológica é muito valiosa, pois contém proteínas diretamente relacionadas com o microambiente de desenvolvimento e progressão do tumor. Nosso trabalho mostrou claramente que existem diferenças individuais importantes que não devem ser menosprezadas na definição do desenho experimental da pesquisa. Idealmente, as mulheres a serem incluídas em um próximo estudo deverão constituir um grupo maior e mais homogêneo, por exemplo, no que diz respeito à idade, estado de pré ou pós-menopausa e uso de hormônios exógenos (contraceptivos orais e terapias de reposição hormonal). A variabilidade proteômica do NAF das pacientes estudadas se refletiu não só na concentração e conteúdo de proteínas totais, assim como também na sua composição de metaloproteases, glicoproteínas e fosfoproteínas. Os resultados de 2D-DIGE mostraram ainda que o NAF é uma amostra com grande intervalo dinâmico, ou seja, constituído por algumas poucas proteínas muito abundantes e muitos outros componentes minoritários, ambos ricos em modificações pós-traducionais, principalmente glicosilação. O estudo do NAF por técnicas proteômicas quantitativas mais modernas é fundamental para conseguirmos detectar e identificar candidatos a biomarcadores nesta secreção. Estas moléculas são de grande importância para a saúde pública, pois teoricamente poderiam auxiliar no diagnóstico precoce da doença maligna da mama, assim como na definição do

prognóstico e na escolha do tratamento mais adequado para cada paciente (Cavalli *et al.*, 2004).

Portanto, para conseguirmos identificar as proteínas menos abundantes do NAF (e seus possíveis candidatos a biomarcadores) (Hunt *et al.*, 2005; Zolg, 2006), precisaremos: a) utilizar métodos de separação de proteínas capazes de reduzir a complexidade das amostras e seu intervalo dinâmico; b) fazer uso de técnicas de enriquecimento de proteínas/peptídeos modificados (glicosilados e fosforilados) que permitam a caracterização de modificações pós-traducionais importantes; c) empregar técnicas de identificação por espectrometria de massas mais sensíveis, acuradas e com amplo intervalo dinâmico linear; d) escolher uma abordagem que permita análises quantitativas sem que haja necessidade de sobreposição de padrões similares de imagens, como na eletroforese bidimensional.

Neste contexto, a metodologia proteômica que nos parece mais adequada atualmente é a análise de peptídeos por nanocromatografia líquida acoplada à espectrometria de massas de alta resolução (Han *et al.*, 2008). Uma das abordagens mais populares envolve a tripsinização em solução de todas as proteínas de uma amostra complexa, seguida da separação dos peptídeos por cromatografia bidimensional (geralmente troca catiônica forte seguida de fase reversa) e identificação por nESI-MS/MS (*shot gun proteomics*). O Laboratório de Toxinologia em breve estará apto a trabalhar com esta tecnologia de ponta, já que acaba de adquirir um nESI LTQ-Orbitrap XL (Thermo Scientific, EUA) com nanocromatografia acoplada, que deverá ser instalado nos próximos meses. Portanto, como um experimento piloto de “prova de conceito”, amostras de NAF tripsinizadas foram enviadas para análise em um equipamento deste mesmo tipo instalado no laboratório do *Protein Research Group*, Departamento de Bioquímica e Biologia Molecular, Universidade do Sul da Dinamarca, Odense, Dinamarca.

Em cada uma das quatro análises por LC-MS/MS, foram aplicados menos de 0,5 µg de digesto trípico do NAF de ambas as mamas das pacientes P29 (perfil A por 2D-PAGE) ou P34 (perfil B). Os peptídeos foram separados por cromatografia de fase reversa, utilizando um gradiente de 0 a 34% de solvente B em 100 min, seguido de análise *on line* no nESI LTQ-Orbitrap. Vale lembrar que, para identificação por MALDI-TOF/TOF, utilizamos 300 µg de proteína em cada 2D-PAGE, quantidade pelo menos 600 vezes maior do que a utilizada neste experimento piloto. A possibilidade de reduzir tão drasticamente a quantidade de proteína

utilizada por ensaio é muito importante no caso do NAF, pois poderia reduzir o volume mínimo necessário de amostra de 15 µL para até 1 µL por coleta.

Ainda que separando uma pequena quantidade de peptídeos em uma única dimensão, conseguimos expandir o número de proteínas identificadas no NAF. Albumina, imunoglobulina A e cadeia kappa de imunoglobulina foram comuns às quatro amostras. Zinco- α 2-glicoproteína, receptor de imunoglobulina, apolipoproteína D e proteína induzida por prolactina foram encontradas nas duas mamas da paciente P29 (perfil A). A lactoferrina e a clusterina foram encontradas apenas nas mamas com câncer e sadia desta mesma paciente, respectivamente. Em ambas as mamas da paciente P34 (perfil B) foram encontradas imunoglobulina G e haptoglobina; na mama doente foi identificada lactalbumina e na mama sadia, transferrina. A Tabela 4.7 mostra todos os detalhes das identificações. Chama atenção a qualidade excepcional da maior parte dos dados: a) proteínas identificadas por diversos peptídeos e com boa cobertura de sequência; b) erro de massa do precursor menor do que 2 ppm; c) ótimos valores de XCorr e delta cnc calculados pelo Sequest; d) probabilidade de acerto maior do que 95% calculada pelo Scaffold.

Utilizando cromatografia bidimensional para fracionar o NAF antes da análise no nESI-LTQ Orbitrap, esperamos ser capazes de aumentar exponencialmente o número de proteínas identificadas. Neste sentido, pretendemos utilizar na primeira dimensão uma coluna de troca catiônica forte eluída com gradiente linear crescente de sal. Na segunda dimensão, as diferentes frações cromatográficas eluídas na primeira etapa serão submetidas individualmente à cromatografia de fase reversa *on line* com o nESI LTQ-Orbitrap (Fang *et al.*, 2010). Para medir as diferenças de expressão proteica nas amostras de NAF, pretendemos marcar quimicamente os peptídeos trípticos por dimetilação com isótopos estáveis, um método de quantificação simples, rápido, barato e eficiente (Boersema *et al.*, 2009).

Tabela 4.7: Identificações do NAF de quatro pacientes por LC-MS/MS

Nome da proteína	Código de acesso no NCBI	Amostra	Sequencia	SEQUEST XCorr	SEQUEST deltaCn	Probabilidade do Scaffold	Precursor		Cobertura de sequencia
							m/z	Carga	
Albumina	gij113576	P29d	(K)AEFAEVSK(L)	2,85	0,34	95%	440,72	2	-0,38
			(K)DDNPMLPR(L)	2,57	0,35	95%	470,73	2	-0,24
			(K)FQNALVLR(Y)	2,67	0,16	95%	480,79	2	-0,32
			(K)TYETLTK(G)	1,78	0,21	95%	492,75	2	0,22
			(K)LVAAASQAALGL(-)	2,70	0,45	95%	507,30	2	-0,23
			(K)SLHTLFGDK(L)	2,35	0,22	95%	509,27	2	0,42
			(K)LDELRLDEGK(A)	2,54	0,20	95%	537,78	2	1,34
			(K)CcamCcamTESLVNR(R)	2,66	0,46	95%	569,75	2	-1,02
			(K)AAFTFCcamCcamQAADK(A)	3,23	0,42	95%	686,29	2	1,51
			(K)YICcamENQDSSSK(L)	3,13	0,29	95%	722,33	2	-0,10
			(K)TCcamVADESAENCcamDK(S)	4,49	0,74	95%	749,80	2	2,70
			(K)KVPQVSTPTLVEVSR(N)	3,36	0,53	95%	820,47	2	-0,96
			(K)QNCcamELFEQLGEYK(F)	3,89	0,39	95%	829,38	2	2,28
			(K)AEFAEVSK(L)	2,65	0,21	95%	440,72	2	-0,10
			(K)DDNPMLPR(L)	2,09	0,33	95%	470,73	2	-0,24
(K)FQNALVLR(Y)	2,91	0,21	95%	480,79	2	-0,38			
(K)TYETLTK(G)	1,94	0,18	95%	492,75	2	0,96			
(K)LVAAASQAALGL(-)	2,61	0,46	95%	507,30	2	-1,38			
(K)CcamCcamTESLVNR(R)	2,48	0,54	95%	569,75	2	-0,49			
(K)LVNEVTEFAK(I)	3,19	0,49	95%	575,31	2	-1,45			
(K)AAFTFCcamCcamQAADK(A)	2,64	0,38	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,80	0,38	95%	722,33	2	1,43			
(K)TCcamVADESAENCcamDK(S)	4,45	0,72	95%	749,79	2	1,64			
(K)QNCcamELFEQLGEYK(F)	2,58	0,49	95%	829,38	2	2,20			
(R)FKDLGEEFK(A)	2,11	0,14	95%	409,54	2	-0,51			
(K)AEFAEVSK(L)	2,92	0,12	95%	440,73	2	0,94			
(K)LYLIEIAR(R)	2,50	0,10	95%	464,25	2	-0,97			
(K)DDNPMLPR(L)	2,27	0,43	95%	470,73	2	-0,83			
(K)FQNALVLR(Y)	3,01	0,22	95%	480,79	2	-0,63			
(K)TYETLTK(G)	2,19	0,37	95%	492,75	2	-0,77			
(K)QTALVELVK(H)	2,68	0,24	95%	500,81	2	-0,09			
(K)LVAAASQAALGL(-)	3,67	0,50	95%	507,30	2	-0,17			
(K)SLHTLFGDK(L)	2,26	0,14	95%	509,27	2	0,60			
(K)LDELRLDEGK(A)	2,32	0,19	95%	537,78	2	0,21			
(K)KQALVELVK(H)	2,39	0,31	95%	564,85	2	0,43			
(K)CcamCcamTESLVNR(R)	2,97	0,37	95%	569,75	2	0,16			
(K)LVNEVTEFAK(I)	3,73	0,41	95%	575,31	2	0,04			
(K)AAFTFCcamCcamQAADK(A)	3,67	0,57	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,95	0,43	95%	722,33	2	2,36			
(K)TCcamVADESAENCcamDK(S)	4,89	0,76	95%	749,80	2	3,03			
(K)KVPQVSTPTLVEVSR(N)	3,90	0,49	95%	756,43	2	2,70			
(K)KVPQVSTPTLVEVSR(N)	3,75	0,49	95%	820,48	2	2,69			
(K)QNCcamELFEQLGEYK(F)	3,66	0,58	95%	829,38	2	1,25			
Albumina	gij113576	P29e	(K)AEFAEVSK(L)	2,85	0,34	95%	440,72	2	-0,38
			(K)DDNPMLPR(L)	2,57	0,35	95%	470,73	2	-0,24
			(K)FQNALVLR(Y)	2,67	0,16	95%	480,79	2	-0,32
			(K)TYETLTK(G)	1,78	0,21	95%	492,75	2	0,22
			(K)LVAAASQAALGL(-)	2,70	0,45	95%	507,30	2	-0,23
			(K)SLHTLFGDK(L)	2,35	0,22	95%	509,27	2	0,42
			(K)LDELRLDEGK(A)	2,54	0,20	95%	537,78	2	1,34
			(K)CcamCcamTESLVNR(R)	2,66	0,46	95%	569,75	2	-1,02
			(K)AAFTFCcamCcamQAADK(A)	3,23	0,42	95%	686,29	2	1,51
			(K)YICcamENQDSSSK(L)	3,13	0,29	95%	722,33	2	-0,10
			(K)TCcamVADESAENCcamDK(S)	4,49	0,74	95%	749,80	2	2,70
			(K)KVPQVSTPTLVEVSR(N)	3,36	0,53	95%	820,47	2	-0,96
			(K)QNCcamELFEQLGEYK(F)	3,89	0,39	95%	829,38	2	2,28
			(K)AEFAEVSK(L)	2,65	0,21	95%	440,72	2	-0,10
			(K)DDNPMLPR(L)	2,09	0,33	95%	470,73	2	-0,24
(K)FQNALVLR(Y)	2,91	0,21	95%	480,79	2	-0,38			
(K)TYETLTK(G)	1,94	0,18	95%	492,75	2	0,96			
(K)LVAAASQAALGL(-)	2,61	0,46	95%	507,30	2	-1,38			
(K)CcamCcamTESLVNR(R)	2,48	0,54	95%	569,75	2	-0,49			
(K)LVNEVTEFAK(I)	3,19	0,49	95%	575,31	2	-1,45			
(K)AAFTFCcamCcamQAADK(A)	2,64	0,38	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,80	0,38	95%	722,33	2	1,43			
(K)TCcamVADESAENCcamDK(S)	4,45	0,72	95%	749,79	2	1,64			
(K)QNCcamELFEQLGEYK(F)	2,58	0,49	95%	829,38	2	2,20			
(R)FKDLGEEFK(A)	2,11	0,14	95%	409,54	2	-0,51			
(K)AEFAEVSK(L)	2,92	0,12	95%	440,73	2	0,94			
(K)LYLIEIAR(R)	2,50	0,10	95%	464,25	2	-0,97			
(K)DDNPMLPR(L)	2,27	0,43	95%	470,73	2	-0,83			
(K)FQNALVLR(Y)	3,01	0,22	95%	480,79	2	-0,63			
(K)TYETLTK(G)	2,19	0,37	95%	492,75	2	-0,77			
(K)QTALVELVK(H)	2,68	0,24	95%	500,81	2	-0,09			
(K)LVAAASQAALGL(-)	3,67	0,50	95%	507,30	2	-0,17			
(K)SLHTLFGDK(L)	2,26	0,14	95%	509,27	2	0,60			
(K)LDELRLDEGK(A)	2,32	0,19	95%	537,78	2	0,21			
(K)KQALVELVK(H)	2,39	0,31	95%	564,85	2	0,43			
(K)CcamCcamTESLVNR(R)	2,97	0,37	95%	569,75	2	0,16			
(K)LVNEVTEFAK(I)	3,73	0,41	95%	575,31	2	0,04			
(K)AAFTFCcamCcamQAADK(A)	3,67	0,57	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,95	0,43	95%	722,33	2	2,36			
(K)TCcamVADESAENCcamDK(S)	4,89	0,76	95%	749,80	2	3,03			
(K)KVPQVSTPTLVEVSR(N)	3,90	0,49	95%	756,43	2	2,70			
(K)KVPQVSTPTLVEVSR(N)	3,75	0,49	95%	820,48	2	2,69			
(K)QNCcamELFEQLGEYK(F)	3,66	0,58	95%	829,38	2	1,25			
Albumina	gij113576	P34d	(K)AEFAEVSK(L)	2,85	0,34	95%	440,72	2	-0,38
			(K)DDNPMLPR(L)	2,57	0,35	95%	470,73	2	-0,24
			(K)FQNALVLR(Y)	2,67	0,16	95%	480,79	2	-0,32
			(K)TYETLTK(G)	1,78	0,21	95%	492,75	2	0,22
			(K)LVAAASQAALGL(-)	2,70	0,45	95%	507,30	2	-0,23
			(K)SLHTLFGDK(L)	2,35	0,22	95%	509,27	2	0,42
			(K)LDELRLDEGK(A)	2,54	0,20	95%	537,78	2	1,34
			(K)CcamCcamTESLVNR(R)	2,66	0,46	95%	569,75	2	-1,02
			(K)AAFTFCcamCcamQAADK(A)	3,23	0,42	95%	686,29	2	1,51
			(K)YICcamENQDSSSK(L)	3,13	0,29	95%	722,33	2	-0,10
			(K)TCcamVADESAENCcamDK(S)	4,49	0,74	95%	749,80	2	2,70
			(K)KVPQVSTPTLVEVSR(N)	3,36	0,53	95%	820,47	2	-0,96
			(K)QNCcamELFEQLGEYK(F)	3,89	0,39	95%	829,38	2	2,28
			(K)AEFAEVSK(L)	2,65	0,21	95%	440,72	2	-0,10
			(K)DDNPMLPR(L)	2,09	0,33	95%	470,73	2	-0,24
(K)FQNALVLR(Y)	2,91	0,21	95%	480,79	2	-0,38			
(K)TYETLTK(G)	1,94	0,18	95%	492,75	2	0,96			
(K)LVAAASQAALGL(-)	2,61	0,46	95%	507,30	2	-1,38			
(K)CcamCcamTESLVNR(R)	2,48	0,54	95%	569,75	2	-0,49			
(K)LVNEVTEFAK(I)	3,19	0,49	95%	575,31	2	-1,45			
(K)AAFTFCcamCcamQAADK(A)	2,64	0,38	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,80	0,38	95%	722,33	2	1,43			
(K)TCcamVADESAENCcamDK(S)	4,45	0,72	95%	749,79	2	1,64			
(K)QNCcamELFEQLGEYK(F)	2,58	0,49	95%	829,38	2	2,20			
(R)FKDLGEEFK(A)	2,11	0,14	95%	409,54	2	-0,51			
(K)AEFAEVSK(L)	2,92	0,12	95%	440,73	2	0,94			
(K)LYLIEIAR(R)	2,50	0,10	95%	464,25	2	-0,97			
(K)DDNPMLPR(L)	2,27	0,43	95%	470,73	2	-0,83			
(K)FQNALVLR(Y)	3,01	0,22	95%	480,79	2	-0,63			
(K)TYETLTK(G)	2,19	0,37	95%	492,75	2	-0,77			
(K)QTALVELVK(H)	2,68	0,24	95%	500,81	2	-0,09			
(K)LVAAASQAALGL(-)	3,67	0,50	95%	507,30	2	-0,17			
(K)SLHTLFGDK(L)	2,26	0,14	95%	509,27	2	0,60			
(K)LDELRLDEGK(A)	2,32	0,19	95%	537,78	2	0,21			
(K)KQALVELVK(H)	2,39	0,31	95%	564,85	2	0,43			
(K)CcamCcamTESLVNR(R)	2,97	0,37	95%	569,75	2	0,16			
(K)LVNEVTEFAK(I)	3,73	0,41	95%	575,31	2	0,04			
(K)AAFTFCcamCcamQAADK(A)	3,67	0,57	95%	686,29	2	0,80			
(K)YICcamENQDSSSK(L)	3,95	0,43	95%	722,33	2	2,36			
(K)TCcamVADESAENCcamDK(S)	4,89	0,76	95%	749,80	2	3,03			
(K)KVPQVSTPTLVEVSR(N)	3,90	0,49	95%	756,43	2	2,70			
(K)KVPQVSTPTLVEVSR(N)	3,75	0,49	95%	820,48	2	2,69			
(K)QNCcamELFEQLGEYK(F)	3,66	0,58	95%	829,38	2	1,25			

Tabela 4.7. Continuação

Nome da proteína	Código de acesso no NCBI	Amostra	Sequencia	SEQUEST XCorr	SEQUEST deltaCn	Probabilidade do Scaffold	Precursor		Cobertura de sequencia	
							Carga	Erro(ppm)		
							m/z			
Albumina	gi 113576	P34e	(R)FKDLGEENFK(A)	2,60	0,22	95%	409,54	2	-1,85	
			(K)AEFAEVSK(L)	2,70	0,26	95%	440,73	2	2,95	
			(K)LYEJAR(R)	2,72	0,09	95%	464,25	2	-1,04	
			(K)DDNPMLPR(L)	2,54	0,34	95%	470,73	2	-0,24	
			(K)FQNALLVR(Y)	3,20	0,22	95%	480,79	2	-0,70	
			(K)TYETLLEK(C)	2,64	0,39	95%	492,75	2	0,90	
			(K)QTALVELVK(H)	2,97	0,27	95%	500,81	2	-0,15	
			(K)LVAASQAALGL(-)	3,47	0,54	95%	507,30	2	0,19	
			(K)SLHTLFGDK(L)	2,36	0,25	95%	509,27	2	-0,60	
			(K)LDELRLDEGK(A)	2,25	0,24	95%	537,78	2	3,05	
			(K)ADDKETC _{cam} FAEEGK(L)	2,74	0,33	95%	543,25	2	1,85	
			(K)KVPQVSTPTLVEVSR(N)	2,06	0,32	95%	547,32	2	0,63	
			(K)KQTALVELVK(H)	2,57	0,30	95%	564,85	2	-0,22	32%
			(K) _{Ccam} C _{cam} TESLVNR(R)	3,04	0,47	95%	569,75	2	1,01	
			(K)KLVAAASQAALGL(-)	3,06	0,31	95%	571,35	2	1,96	
			(K)LVNEVTEFAK(T)	3,16	0,44	95%	575,31	2	0,04	
			(K)AAAFTEC _{cam} C _{cam} QAADK(A)	3,72	0,53	95%	686,29	2	2,31	
			(K)YIC _{cam} ENQDISSK(L)	3,94	0,35	95%	722,33	2	0,83	
			(K)TC _{cam} VADESAENcDK(S)	4,54	0,75	95%	749,79	2	1,40	
			(K)VPQVSTPTLVEVSR(N)	3,91	0,30	95%	756,42	2	-1,66	
(K)LKEC _{cam} EKPLLEK(S)	3,57	0,30	95%	773,90	2	0,72				
(K)ADDKETC _{cam} FAEEGK(L)	3,32	0,42	95%	814,37	2	1,89				
(K)KVPQVSTPTLVEVSR(N)	4,02	0,62	95%	820,47	2	1,28				
(K)QNC _{cam} ELFEQLGEYK(F)	3,66	0,59	95%	829,38	2	-0,96				
P29d			(K)TPLTATLSK(S)	2,82	0,26	95%	466,28	2	-0,17	
			(K)SAVQGGPPER(D)	2,91	0,39	95%	470,75	2	-0,09	
			(R)WLQGSQELPR(E)	3,36	0,25	95%	607,32	2	0,82	20%
			(K)TFTC _{cam} TAAYPEsk(T)	3,15	0,45	95%	688,32	2	1,65	
			(R)DASGVFTWTPSSGK(S)	3,88	0,55	95%	770,87	2	0,36	
			(R)QEPSQGITTFVTSILR(V)	3,36	0,60	95%	918,48	2	2,26	
Imunoglobulina A	gi 113584	P29e	(K)SAVQGGPPER(D)	2,74	0,34	95%	470,75	2	-1,13	
			(R)WLQGSQELPR(E)	3,44	0,22	95%	607,32	2	2,13	
			(K)TFTC _{cam} TAAYPEsk(T)	3,13	0,40	95%	688,32	2	1,73	18%
			(R)DASGVFTWTPSSGK(S)	3,27	0,57	95%	770,87	2	-0,27	
P34d			(R)QEPSQGITTFVTSILR(V)	3,09	0,55	95%	918,48	2	2,13	
			(K)SAVQGGPPER(D)	2,91	0,39	95%	470,75	2	-0,02	3%
P34e			(K)SAVQGGPPER(D)	2,74	0,35	95%	470,75	2	-0,87	
			(R)WLQGSQELPR(E)	3,38	0,26	95%	607,32	2	0,32	5%

Tabela 4.7: Continuação

Nome da proteína	Código de acesso no NCBI	Amostra	Sequencia	SEQUEST XCorr	SEQUEST deltaCn	Probabilidade do Scaffold	Precursor		Cobertura de sequencia		
							m/z	Erro(ppm)			
Imunoglobulina cadeia kappa	P29d	gij125145	(K)SFNRGECcam(-)	2,68	0,52	95%	435,18	2	0,37	35%	
			(K)DSTYSLSSTLTLSK(A)	2,99	0,26	95%	751,88	2	-1,38		
			(-)TVAAPSVFIIPPSDEQLK(S)	3,89	0,52	95%	973,52	2	1,09		
			(K)SFNRGECcam(-)	2,27	0,45	95%	435,18	2	-0,05		
			(K)VVDNALQSGNSQESVTEQDSK(D)	4,76	0,62	95%	1068,49	2	2,13		
			(K)SFNRGECcam(-)	2,47	0,59	95%	435,18	2	-0,61		
	P34d	gij125145	(-)TVAAPSVFIIPPSDEQLK(S)	3,25	0,37	95%	973,52	2	1,41	24%	
			(K)SFNRGECcam(-)	2,52	0,52	95%	435,18	2	-0,05		
			(K)VVDNALQSGNSQESVTEQDSK(D)	4,91	0,68	95%	1068,49	2	2,70		
	Receptor de imunoglobulina	P29d	gij150421625	(K)RAPAFEGRI(I)	2,93	0,18	95%	452,24	2	0,10	4%
				(K)AFVNCcamDENSRL(L)	3,18	0,43	95%	606,26	2	2,06	
				(R)ADEGWYWCcamGVK(Q)	2,65	0,43	95%	685,80	2	2,48	
(K)AFVNCcamDENSRL(L)				3,00	0,42	95%	606,26	2	0,55		
Proteína induzida por prolactina	P29d	gij134170	(R)FYTIKVE(-)	3,27	0,30	95%	627,85	2	0,85	25%	
			(K)YTAcamLcDDNPK(T)	3,52	0,54	95%	678,78	2	2,03		
			(R)ELGICcamPDDAAVPIK(N)	4,31	0,44	95%	805,93	2	3,54		
			(K)YTAcamLcDDNPK(T)	3,62	0,63	95%	678,78	2	-1,30		
			(R)ELGICcamPDDAAVPIK(N)	2,98	0,40	95%	805,93	2	1,42		
	P29e	gij114034	(K)VLNQLR(A)	3,07	0,11	95%	436,25	2	-0,45	18%	
			(K)NILTSNNIDVK(K)	3,60	0,20	95%	615,84	2	-0,10		
			(K)MTVTDQVNCcamPK(L)	3,38	0,39	95%	646,80	2	0,77		
			(K)MoxTVTDQVNCcamPK(L)	3,57	0,48	95%	654,80	2	0,64		
			(K)NILTSNNIDVKK(M)	3,54	0,33	95%	679,89	2	1,23		
Apolipoproteína D	P29e	gij114034	(R)NPNLPPETVDSLK(N)	3,21	0,32	95%	712,38	2	0,48	30%	
			(K)CcamPNPPVQENFDVNK(Y)	4,08	0,50	95%	829,39	2	1,99		
			(R)WYEIEK(I)	2,29	0,14	95%	434,22	2	-0,51		
			(K)NILTSNNIDVK(K)	3,96	0,28	95%	615,84	2	0,79		
			(K)MTVTDQVNCcamPK(L)	3,82	0,34	95%	646,80	2	0,11		
P29e	gij114034	gij114034	(K)MoxTVTDQVNCcamPK(L)	3,93	0,50	95%	654,80	2	2,41	33%	
			(K)NILTSNNIDVKK(M)	3,91	0,36	95%	679,89	2	1,23		
			(R)NPNLPPETVDSLK(N)	3,39	0,45	95%	712,38	2	-0,29		
			(K)CcamPNPPVQENFDVNK(Y)	3,80	0,51	95%	829,39	2	0,74		

Tabela 4.7: Continuação

Nome da proteína	Código de acesso no NCBI	Amostra	Sequencia	SEQUEST XCorr	SEQUEST deltaCn	Probabilidade do Scaffold	Precursor		Cobertura de sequencia	
							m/z	Carga		Erro(ppm)
Zinco-α2-glicoproteína	gi 141596	P29d	(K)YYDGGK(D)	2,20	0,39	95%	404,68	2	0,32	14%
			(R)AGEVQPELR(G)	3,55	0,37	95%	564,29	2	2,44	
			(K)WEAEPVYQR(A)	2,91	0,22	95%	638,82	2	0,78	
			(R)YSLTYITGLSK(H)	4,08	0,62	95%	704,87	2	3,06	
			(K)AYLEEECamPATLR(K)	4,30	0,56	95%	726,35	2	1,74	
			(R)AGEVQPELR(G)	3,70	0,36	95%	564,29	2	0,06	
			(K)WEAEPVYQR(A)	2,70	0,13	95%	638,82	2	1,55	
			(R)YSLTYITGLSK(H)	4,20	0,60	95%	704,87	2	1,58	
			(K)AYLEEECamPATLR(K)	4,40	0,57	95%	726,35	2	0,48	
			(R)DGAGDVAFIR(E)	3,40	0,25	95%	510,76	2	-0,38	
Lactoferrina	gi 85700158	P29d	(R)VVWCcamAVGEQELR(K)	3,65	0,15	95%	723,37	2	0,90	5%
			(K)NLLFNDNTECamLAR(L)	3,57	0,38	95%	790,38	2	0,38	
Clusterina	gi 116533	P29e	(R)ELDESILQVAER(L)	3,80	0,35	95%	644,82	2	-0,38	6%
			(R)EILSVDCcamSTNPFSGAK(L)	4,02	0,59	95%	881,92	2	1,20	
Haptoglobina	gi 123508	P34d	(K)DYAEVGR(V)	2,95	0,32	95%	405,19	2	0,96	8%
			(K)SCcamVAEYGVYK(V)	3,13	0,32	95%	673,33	2	-2,62	
			(R)TEGDGVYTLNNEK(Q)	3,41	0,49	95%	720,34	2	1,52	
			(K)DIAPTLTYVGG(K)	2,90	0,34	95%	645,87	2	2,51	
Imunoglobulina G	gi 121039	P34e	(K)SCcamVAEYGVYK(V)	3,77	0,38	95%	673,33	2	-1,98	9%
			(R)TEGDGVYTLNNEK(Q)	2,81	0,28	95%	720,34	2	1,43	
			(K)NGVSLTCcamLVK(G)	2,83	0,32	95%	581,32	2	1,83	
			(K)STSGGTAALGCcamLVK(D)	3,73	0,40	95%	661,34	2	0,65	
Lactalbumina	gi 125946	P34d	(K)NGVSLTCcamLVK(G)	3,09	0,29	95%	581,32	2	0,36	12%
			(K)STSGGTAALGCcamLVK(D)	3,08	0,23	95%	661,34	2	-1,01	
			(K)TPPVLDSDGSFFLYSK(L)	2,42	0,45	95%	937,47	2	2,12	
			(K)AGVETITPSK(Q)	3,45	0,33	95%	495,76	2	0,08	
Transferrina	gi 136191	P34e	(K)DGAGDVAFVK(H)	3,53	0,38	95%	489,75	2	-0,19	4%
			(K)YLGEEYK(A)	2,40	0,16	95%	500,75	2	0,50	
			(K)HQTVPQNTGGK(N)	2,50	0,44	95%	583,80	2	0,20	
			(K)EGYGYTGAFRIC	2,44	0,37	95%	642,29	2	0,62	

As modificações estão simbolizadas: Ccam: metilação de cisteína - Ccam: oxidação de metionina - Mox.

5 – CONCLUSÕES

- A concentração proteica do NAF foi alta e bastante variável (média $115 \pm 60 \mu\text{g}/\mu\text{L}$), sem diferenças estatisticamente significativas entre os grupos tumor e controle.
- As amostras de mulheres controle e com doença benigna mostraram perfis eletroforéticos semelhantes, comparáveis também com outros estudos publicados na literatura utilizando amostras de mulheres sem câncer.
- A análise zimográfica do NAF mostrou correspondência entre os perfis de bandas proteolíticas das duas mamas de uma mesma paciente, independente da presença ou ausência da doença. Entre as diferentes pacientes, os perfis se mostraram bastante distintos. A atividade proteolítica do NAF se mostrou sensível a quelantes de cátions metálicos.
- Na eletroforese bidimensional, a aplicação da amostra por reidratação *in gel* em fitas de focalização isoeletrica no intervalo de pH 4-7 resultou na melhor resolução dos *spots* do NAF.
- Os perfis de 2D-PAGE das duas mamas de uma mesma paciente se mostraram reprodutíveis e permitiram a identificação de diversas proteínas características do NAF.
- A depleção das proteínas mais abundantes do NAF utilizando diferentes abordagens se mostrou pouco eficiente, não sendo recomendada nas análises futuras desta amostra biológica.
- A análise do NAF por eletroforese de fluorescência diferencial (DIGE) revelou diferenças individuais importantes, embora as duas mamas de uma mesma paciente tenham apresentado perfis bastante similares. Os diferentes perfis observados entre as pacientes foram classificados em dois grandes grupos (perfil A e perfil B). Devido à variação individual encontrada nos perfis eletroforéticos das amostras de algumas pacientes não foi possível, através da análise por DIGE, realizar a análise de expressão diferencial entre as mamas saudáveis e doentes.

- O NAF se mostrou rico em glicoproteínas, mas pouco abundante em proteínas fosforiladas. O mesmo padrão de variação individual descrito acima foi observado na análise destes subproteomas.
- A utilização de técnicas proteômicas mais modernas (ex.: nLC-ESI-LTQ Orbitrap) será necessária para identificação de proteínas menos abundantes no NAF, incluindo potenciais biomarcadores da doença e realizar a comparação das diferenças expressão entre mamas sadias e com câncer de mama.

6 - REFERÊNCIAS BIBLIOGRÁFICAS

Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature*. 2003 Mar 13;422(6928):198-207.

Akerstrom B, Bjorck L. Protein L: an immunoglobulin light chain-binding bacterial protein. Characterization of binding and physicochemical properties. *J Biol Chem*. 1989 Nov 25;264(33):19740-6.

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I. A novel experimental design for comparative two-dimensional gel analysis: two-dimensional difference gel electrophoresis incorporating a pooled internal standard. *Proteomics*. 2003 Jan;3(1):36-44.

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD. *The molecular biology of the cell*. 3rd ed. New York: Garland Publishing; 1994.

Albuquerque LM, Trugilho MR, Chapeaurouge A, Jurgilas PB, Bozza PT, Bozza FA, Perales J, Neves-Ferreira AG. Two-dimensional difference gel electrophoresis (DiGE) analysis of plasmas from dengue fever patients. *J Proteome Res*. 2009 Dec;8(12):5431-41.

Alexander H, Stegner AL, Wagner-Mann C, Du Bois GC, Alexander S, Sauter ER. Proteomic analysis to identify breast cancer biomarkers in nipple aspirate fluid. *Clin Cancer Res*. 2004 Nov 15;10(22):7500-10.

Ali S, Coombes RC. Endocrine-responsive breast cancer and strategies for combating resistance. *Nat Rev Cancer*. 2002 Feb;2(2): 101-12.

Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002 Nov;1(11):845-67.

Beynon RJ, Bond, J.S. *Proteolytic Enzymes*. Rickwood DaH, B.D., editor. New York: IRL press at Oxford University Press; 1996.

Boersema PJ, Raijmakers R, Lemeer S, Mohammed S, Heck AJ. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat Protoc*. 2009;4(4):484-94.

Bray GA. The underlying basis for obesity: relationship to cancer. *J Nutr*. 2002 Nov;132(11 Suppl):3451S-5S.

Bundred NJ, Miller WR, Walker RA. An immunohistochemical study of the tissue distribution of the breast cyst fluid protein, zinc alpha 2 glycoprotein. *Histopathology*. 1987 Jun;11(6):603-10.

Bundred NJ, Scott WN, Davies SJ, Miller WR, Mansel RE. Zinc alpha-2 glycoprotein levels in serum and breast fluids: a potential marker of apocrine activity. *Eur J Cancer*. 1991;27(5):549-52.

Caiazza F, Harvey BJ, Thomas W. Cytosolic Phospholipase A2 Activation Correlates with HER2 Overexpression and Mediates Estrogen-Dependent Breast Cancer Cell Growth. *Mol Endocrinol*. Baltimore, Md. 2010 Mar 8.

Cavalli F, Hansen H, Kaye SB. Textbook of medical oncology: Breast cancer. 3^a ed. Londres: Martin Dunitz; 2004.

Celis JE, Gromov P, Moreira JM, Cabezon T, Friis E, Vejborg IM, Proess G, Rank F, Gromova I. Apocrine cysts of the breast: biomarkers, origin, enlargement, and relation with cancer phenotype. *Mol Cell Proteomics*. 2006 Mar;5(3):462-83.

Chatterton RT, Jr., Geiger AS, Khan SA, Helenowski IB, Jovanovic BD, Gann PH. Variation in estradiol, estradiol precursors, and estrogen-related products in nipple aspirate fluid from normal premenopausal women. *Cancer Epidemiol Biomarkers Prev*. 2004 Jun;13(6):928-35.

Choe LH, Werner BG, Lee KH. Two-dimensional protein electrophoresis: from molecular pathway discovery to biomarker discovery in neurological disorders. *NeuroRx*. 2006 Jul;3(3):327-35.

Contran R, Kumar, V, Collins, T. Robbins: Patologia estrutural e funcional. . 6^a ed ed. Rio de Janeiro: Editora Guanabara Koogan; 2000.

Coombes KR, Fritsche HA, Jr., Clarke C, Chen JN, Baggerly KA, Morris JS, Xiao LC, Hung MC, Kuerer HM. Quality control and peak finding for proteomics data collected from nipple aspirate fluid by surface-enhanced laser desorption and ionization. *Clin Chem*. 2003 Oct;49(10):1615-23.

Craig R, Beavis RC. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun Mass Spectrom*. 2003;17(20):2310-6.

Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteomics*. 2004 Apr;3(4):367-78.

Djuric Z, Visscher DW, Heilbrun LK, Chen G, Atkins M, Covington CY. Influence of lactation history on breast nipple aspirate fluid yields and fluid composition. *Breast J*. 2005 Mar-Apr;11(2):92-9.

Dwek MV, Ross HA, Leatham AJ. Proteome and glycosylation mapping identifies post-translational modifications associated with aggressive breast cancer. *Proteomics*. 2001 Jun;1(6):756-62.

Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*. 1991 Nov;19(5):403-10.

Ernster VL, Wrensch MR, Petrakis NL, King EB, Miike R, Murai J, Goodson WH, 3rd, Siiteri PK. Benign and malignant breast disease: initial study results of serum and breast fluid analyses of endogenous estrogens. *J Natl Cancer Inst.* 1987 Nov;79(5):949-60.

Fang Y, Robinson DP, Foster LJ. Quantitative Analysis of Proteome Coverage and Recovery Rates for Upstream Fractionation Methods in Proteomics. *J Proteome Res.* 2010 Feb 4.

Fazekas de St Groth S, Webster RG, Datyner A. Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochim Biophys Acta.* 1963 May 14;71:377-91.

Franco J. *Mastologia: Formação do especialista.* São Paulo: Editora Atheneu; 1997.

Gallegos B, Perez-Campos E, Martinez R, Leyva P, Martinez M, Hernandez R, Pina S, Hernandez C, Zenteno E, Hernandez P. O-glycosylation expression in fibroadenoma. *Prep Biochem Biotechnol.* 2010;40(1):1-12.

Gesase AP, Satoh Y. Apocrine secretory mechanism: recent findings and unresolved problems. *Histol Histopathol.* 2003 Apr;18(2):597-608.

Gharbi S, Gaffney P, Yang A, Zvelebil MJ, Cramer R, Waterfield MD, Timms JF. Evaluation of two-dimensional differential gel electrophoresis for proteomic expression analysis of a model breast cancer cell system. *Mol Cell Proteomics.* 2002 Feb;1(2):91-8.

Good DM, Thongboonkerd V, Novak J, Bascands JL, Schanstra JP, Coon JJ, Dominiczak A, Mischak H. Body fluid proteomics for biomarker discovery: lessons from the past hold the key to success in the future. *J Proteome Res.* 2007 Dec;6(12):4549-55.

Gordon NG. *Cecil Tratado de Medicina Interna.* 21ª ed. Rio de Janeiro: Editora Guanabara; 2001.

Görg A, Weiss W, Dunn MJ. Current two-dimensional electrophoresis technology for proteomics. *Proteomics.* 2004 Dec;4(12):3665-85.

Görg A, Klaus A, Lück C, Weiland F, Weiss W. *Two-Dimensional Electrophoresis with Immobilized pH Gradients for Proteome Analysis: A LABORATORY MANUAL.* 2007 [updated 2007; cited]; Available from: <http://www.wzw.tum.de/proteomik/>.

Graham DR, Elliott ST, Van Eyk JE. Broad-based proteomic strategies: a practical guide to proteomics and functional screening. *J Physiol.* 2005 Feb 15;563(Pt 1):1-9.

Griffiths A, Miller JH, Suzuki DT, Lewontin RC, Gelbart WM. *Introdução a genética.* Sétima ed. Rio de Janeiro: Editora Guanabara Koogan; 2002.

Gross JH. *Mass Spectrometry: A textbook.* Berlin: Springer-Verlag; 2004.

Hall RE, Clements JA, Birrell SN, Tilley WD. Prostate-specific antigen and gross cystic disease fluid protein-15 are co-expressed in androgen receptor-positive breast tumours. *Br J Cancer*. 1998 Aug;78(3):360-5.

Han X, Aslanian A, Yates JR, 3rd. Mass spectrometry for proteomics. *Curr Opin Chem Biol*. 2008 Oct;12(5):483-90.

Hanash S. Disease proteomics. *Nature*. 2003 Mar 13;422(6928):226-32.

Harding C, Osundeko O, Tetlow L, Faragher EB, Howell A, Bundred NJ. Hormonally-regulated proteins in breast secretions are markers of target organ sensitivity. *Br J Cancer*. 2000 Jan;82(2):354-60.

Hassan MI, Waheed A, Yadav S, Singh TP, Ahmad F. Zinc alpha 2-glycoprotein: a multidisciplinary protein. *Mol Cancer Res*. 2008 Jun;6(6):892-906.

Heukeshoven J, Dernick R. Simplified method for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining. *Electrophoresis*. 1985;6:130-12.

Heussen C, Dowdle EB. Electrophoretic analysis of plasminogen activators in polyacrylamide gels containing sodium dodecyl sulfate and copolymerized substrates. *Anal Biochem*. 1980;102:196-202.

Higgins SA, Matloff ET, Rimm DL, Dziura J, Haffty BG, King BL. Patterns of reduced nipple aspirate fluid production and ductal lavage cellularity in women at high risk for breast cancer. *Breast Cancer Res*. 2005;7(6):R1017-22.

Hjerno K. Protein identification by peptide mass fingerprinting. In: Matthiesen R, editor. *Mass Spectrometry Data Analysis in Proteomics*. Totowa, New Jersey: Humana Press; 2007. p. 61-75.

Hondermarck H. Breast cancer: when proteomics challenges biological complexity. *Mol Cell Proteomics*. 2003 May;2(5):281-91.

Hondermarck H, Tastet C, El Yazidi-Belkoura I, Toillon RA, Le Bourhis X. Proteomics of breast cancer: the quest for markers and therapeutic targets. *J Proteome Res*. 2008 Apr;7(4):1403-11.

Huang Y, Nagamani M, Anderson KE, Kurosky A, Haag AM, Grady JJ, Lu LJ. A strong association between body fat mass and protein profiles in nipple aspirate fluid of healthy premenopausal non-lactating women. *Breast Cancer Res Treat*. 2007 Jul;104(1):57-66.

Hunt SM, Thomas MR, Sebastian LT, Pedersen SK, Harcourt RL, Sloane AJ, Wilkins MR. Optimal replication and the importance of experimental design for gel-based quantitative proteomics. *J Proteome Res*. 2005 May-Jun;4(3):809-19.

INCA. Controle do câncer de mama: documento de consenso. Rio de Janeiro; 2004 [updated 2004; cited 25 march]; Available from: <http://www.inca.gov.br/publicacoes/ConsensoIntegra.pdf>.

INCA. Estimativa 2010: Incidência de câncer no Brasil. Rio de Janeiro; 2009 [updated 2009; cited 25 march 2010]; Available from: <http://www.inca.gov.br/estimativa/2010/estimativa20091201.pdf>.

Keller A, Nesvizhskii AI, Kolker E, Aebersold R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem*. 2002 Oct 15;74(20):5383-92.

Klein P, Glaser E, Grogan L, Keane M, Lipkowitz S, Soballe P, Brooks L, Jenkins J, Steinberg SM, DeMarini DM, Kirsch I. Biomarker assays in nipple aspirate fluid. *Breast J*. 2001 Nov-Dec;7(6):378-87.

Kuerer HM, Goldknopf IL, Fritsche H, Krishnamurthy S, Sheta EA, Hunt KK. Identification of distinct protein expression patterns in bilateral matched pair breast ductal fluid specimens from women with unilateral invasive breast carcinoma. High-throughput biomarker discovery. *Cancer*. 2002 Dec 1;95(11):2276-82.

Kuerer HM, Coombes KR, Chen JN, Xiao L, Clarke C, Fritsche H, Krishnamurthy S, Marcy S, Hung MC, Hunt KK. Association between ductal fluid proteomic expression profiles and the presence of lymph node metastases in women with breast cancer. *Surgery*. 2004 Nov;136(5):1061-9.

Kumar SR, Sauter ER, Quinn TP, Deutscher SL. Thomsen-Friedenreich and Tn antigens in nipple fluid: carbohydrate biomarkers for breast cancer detection. *Clin Cancer Res*. 2005 Oct 1;11(19 Pt 1):6868-71.

La Rocca G, Pucci-Minafra I, Marrazzo A, Taormina P, Minafra S. Zymographic detection and clinical correlations of MMP-2 and MMP-9 in breast cancer sera. *Br J Cancer*. 2004 Apr 5;90(7):1414-21.

Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. 1970;227(259):680-5.

LaRocca G, Pucci-Minafra I, Marrazzo A, Taormina P, Minafra S. Zymographic detection and clinical correlations of MMP-2 and MMP-9 in breast cancer sera. *Br J Cancer*. 2004 Apr 5;90(7):1414-21.

Lefkovits I. Functional and structural proteomics: a critical appraisal. *J Chromatogr*. 2003 Apr 5;787(1):1-10.

Lescuyer P, Hochstrasser D, Rabilloud T. How shall we use the proteomics toolbox for biomarker discovery? *J Proteome Res*. 2007 Sep;6(9):3371-6.

Li CI, Uribe DJ, Daling JR. Clinical characteristics of different histologic types of breast cancer. *Br J Cancer*. 2005 Oct 31;93(9):1046-52.

Liotta LA, Kohn EC. The microenvironment of the tumour-host interface. *Nature*. 2001 May 17;411(6835):375-9.

Lohaus C, Nolte A, Bluggel M, Scheer C, Klose J, Gobom J, Schuler A, Wiebringhaus T, Meyer HE, Marcus K. Multidimensional chromatography: a powerful tool for the analysis of membrane proteins in mouse brain. *J Proteome Res*. 2007 Jan;6(1):105-13.

Malatesta M, Mannello F, Bianchi G, Sebastiani M, Gazzanelli G. Biochemical and ultrastructural features of human milk and nipple aspirate fluids. *J Clin Lab Anal*. 2000;14(6):330-5.

Mannello F, Sebastiani M. Zymographic analyses and measurement of matrix metalloproteinase-2 and -9 in nipple aspirate fluids. *Clin Chem*. 2003 Sep;49(9):1546-50.

Mannello F, Tonti GA, Medda V, Pederzoli A, Sauter ER. Increased shedding of soluble fragments of P-cadherin in nipple aspirate fluids from women with breast cancer. *Cancer Sci*. 2008 Nov;99(11):2160-9.

Marouga R, David S, Hawkins E. The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem*. 2005 Jun;382(3):669-78.

Mazoujian G, Pinkus GS, Davis S, Haagensen DE, Jr. Immunohistochemistry of a gross cystic disease fluid protein (GCDFP-15) of the breast. A marker of apocrine epithelium and breast carcinomas with apocrine features. *Am J Pathol*. 1983 Feb;110(2):105-12.

Mitchell G, Sibley PE, Wilson AP, Sauter E, A'Hern R, Eeles RA. Prostate-specific antigen in nipple aspiration fluid: menstrual cycle variability and correlation with serum prostate-specific antigen. *Tumour Biol*. 2002 Sep-Oct;23(5):287-97.

Mitulovic G, Mechtler K. HPLC techniques for proteomics analysis--a short overview of latest developments. *Brief Funct Genomic Proteomics*. 2006 Dec;5(4):249-60.

Mueller MM, Fusenig NE. Friends or foes - bipolar effects of the tumour stroma in cancer. *Nat Rev Cancer*. 2004 Nov;4(11):839-49.

Nesvizhskii AI, Keller A, Kolker E, Aebersold R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem*. 2003 Sep 1;75(17):4646-58.

Noble J, Dua RS, Locke I, Eeles R, Gui GP, Isacke CM. Proteomic analysis of nipple aspirate fluid throughout the menstrual cycle in healthy pre-menopausal women. *Breast Cancer Res Treat*. 2007 Aug;104(2):191-6.

Nystrom L, Andersson I, Bjurstam N, Frisell J, Nordenskjold B, Rutqvist LE. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet*. 2002 Mar 16;359(9310):909-19.

O'Farrell PH. High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. 1975 May 25;250(10):4007-21.

Omenn GS, States DJ, Adamski M, Blackwell TW, Menon R, Hermjakob H, Apweiler R, Haab BB, Simpson RJ, Eddes JS, Kapp EA, Moritz RL, Chan DW, Rai AJ, Admon A, Aebersold R, Eng J, Hancock WS, Hefta SA, Meyer H, Paik YK, Yoo JS, Ping P, Pounds J, Adkins J, Qian X, Wang R, Wasinger V, Wu CY, Zhao X, Zeng R, Archakov A, Tsugita A, Beer I, Pandey A, Pisano M, Andrews P, Tammen H, Speicher DW, Hanash SM. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics*. 2005 Aug;5(13):3226-45.

Parker SH, Jobe WE, Dennis MA, Stavros AT, Johnson KK, Yakes WF, Truell JE, Price JG, Kortz AB, Clark DG. US-guided automated large-core breast biopsy. *Radiology*. 1993 May;187(2):507-11.

Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. *CA Cancer J Clin*. 2005 Mar-Apr;55(2):74-108.

Pawlik TM, Fritsche H, Coombes KR, Xiao L, Krishnamurthy S, Hunt KK, Pusztai L, Chen JN, Clarke CH, Arun B, Hung MC, Kuerer HM. Significant differences in nipple aspirate fluid protein expression between healthy women and those with breast cancer demonstrated by time-of-flight mass spectrometry. *Breast Cancer Res Treat*. 2005 Jan;89(2):149-57.

Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*. 1999 Dec;20(18):3551-67.

Petrakis NL, Wrensch MR, Ernster VL, Miike R, Murai J, Simberg N, Siiteri PK. Influence of pregnancy and lactation on serum and breast fluid estrogen levels: implications for breast cancer risk. *Int J Cancer*. 1987 Nov 15;40(5):587-91.

Petrakis NL, Lee RE, Miike R, Dupuy ME, Morris M. Coloration of breast fluid related to concentration of cholesterol, cholesterol epoxides, estrogen, and lipid peroxides. *Am J Clin Pathol*. 1988 Jan;89(1):117-20.

Petrakis NL. ASPO Distinguished Achievement Award Lecture. Studies on the epidemiology and natural history of benign breast disease and breast cancer using nipple aspirate fluid. *Cancer Epidemiol Biomarkers Prev*. 1993 Jan-Feb;2(1):3-10.

Petrakis NL, Lowenstein JM, Wiencke JK, Lee MM, Wrensch MR, King EB, Hilton JF, Miike R. Gross cystic disease fluid protein in nipple aspirates of breast fluid of Asian and non-Asian women. *Cancer Epidemiol Biomarkers Prev*. 1993 Nov-Dec;2(6):573-9.

Qian WJ, Jacobs JM, Liu T, Camp DG, 2nd, Smith RD. Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics*. 2006 Oct;5(10):1727-44.

Rabilloud T, Charmont S. Detection of proteins on two-dimensional electrophoresis gels. In: Rabilloud T, editor. *Proteome Research: Two-Dimensional Gel Electrophoresis and Identification Methods*. New York: Springer; 2000. p. 107-26.

Ranuncolo SM, Armanasco E, Cresta C, Bal De Kier Joffe E, Puricelli L. Plasma MMP-9 (92 kDa-MMP) activity is useful in the follow-up and in the assessment of prognosis in breast cancer patients. *Int J Cancer*. 2003 Sep 20;106(5):745-51.

Roy R, Yang J, Moses MA. Matrix metalloproteinases as novel biomarkers and potential therapeutic targets in human cancer. *J Clin Oncol*. 2009 Nov 1;27(31):5287-97.

Sanchez LM, Vizoso F, Diez-Itza I, Lopez-Otin C. Identification of the major protein components in breast secretions from women with benign and malignant breast diseases. *Cancer Res*. 1992 Jan 1;52(1):95-100.

Sauter ER, Ross E, Daly M, Klein-Szanto A, Engstrom PF, Sorling A, Malick J, Ehya H. Nipple aspirate fluid: a promising non-invasive method to identify cellular markers of breast cancer risk. *Br J Cancer*. 1997;76(4):494-501.

Sauter ER, Tichansky DS, Chervoneva I, Diamandis EP. Circulating testosterone and prostate-specific antigen in nipple aspirate fluid and tissue are associated with breast cancer. *Environ Health Perspect*. 2002 Mar;110(3):241-6.

Sauter ER, Klein G, Wagner-Mann C, Diamandis EP. Prostate-specific antigen expression in nipple aspirate fluid is associated with advanced breast cancer. *Cancer Detect Prev*. 2004;28(1):27-31.

Sauter ERS, S.; Hewett, J.E.; Speckman, P.; Du Bois, G. Proteomic analysis of nipple aspirate fluid using SELDI-TOF-MS. *Int J Cancer*. 2005;114:791-6.

Schnitt SJ, Guidi AJ. *Pathology of invasive breast cancer*. 3rd ed. Harris J, Lippman, ME, Morrow, M, Osborne, CK, editor. Philadelphia: Lippincott, Williams and Wilkins; 2004

Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem*. 1996 Mar 1;68(5):850-8.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. Measurement of protein using bicinchoninic acid. *Anal Biochem*. 1985;150(1):76-85.

Somiari RI, Somiari S, Russell S, Shriver CD. Proteomics of breast carcinoma. *J Chromatogr*. 2005 Feb 5;815(1-2):215-25.

Steen H, Mann M. The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol.* 2004 Sep;5(9):699-711.

Stone KL, Williams KR. Enzymatic digestion of proteins and HPLC peptide isolation. In: Matsudaira P, editor. *A Practical Guide to Protein and Peptide Purification for Microsequencing.* San Diego: Academic Press; 1993. p. 43-69.

Thingholm TE, Jorgensen TJ, Jensen ON, Larsen MR. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat Protoc.* 2006;1(4):1929-35.

Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis.* 1997 Oct;18(11):2071-7.

Varnum SM, Covington CC, Woodbury RL, Petritis K, Kangas LJ, Abdullah MS, Pounds JG, Smith RD, Zangar RC. Proteomic characterization of nipple aspirate fluid: identification of potential biomarkers of breast cancer. *Breast Cancer Res Treat.* 2003 Jul;80(1):87-97.

Vizoso F, Sanchez LM, Diez-Itza I, Luz Lamelas M, Lopez-Otin C. Factors affecting protein composition of breast secretions from nonlactating women. *Breast Cancer Res Treat.* 1992;23(3):251-8.

Vizoso F, Plaza E, Vazquez J, Serra C, Lamelas ML, Gonzalez LO, Merino AM, Mendez J. Lysozyme expression by breast carcinomas, correlation with clinicopathologic parameters, and prognostic significance. *Ann Surg Oncol.* 2001 Sep;8(8):667-74.

Vizoso FJ, Gonzalez LO, Corte MD, Rodriguez JC, Vazquez J, Lamelas ML, Junquera S, Merino AM, Garcia-Muniz JL. Study of matrix metalloproteinases and their inhibitors in breast cancer. *Br J Cancer.* 2007 Mar 26;96(6):903-11.

Wasinger VC, Cordwell SJ, Cerpa-Poljak A, Yan JX, Gooley AA, Wilkins MR, Duncan MW, Harris R, Williams KL, Humphrey-Smith I. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis.* 1995;16:1090-4.

WHO. World Cancer Report, 2008. . Lyon: International Agency for Research on Cancer.; 2009 [updated 2009; cited 25 march]; Available from: http://www.iarc.fr/en/publications/pdfs-online/wcr/2008/wcr_.pdf.

Woessner JF, Nagase H. *Matrix metalloproteinases and TIMPs.* Oxford: Oxford University Press; 2000.

Wulfkuhle JD, Sgroi DC, Krutzsch H, McLean K, McGarvey K, Knowlton M, Chen S, Shu H, Sahin A, Kurek R, Wallwiener D, Merino MJ, Petricoin EF, 3rd, Zhao Y, Steeg PS. Proteomics of human breast ductal carcinoma in situ. *Cancer Res.* 2002 Nov 15;62(22):6740-9.

Yamaguchi Y. Microenvironmental regulation of estrogen signals in breast cancer. *Breast Cancer.* 2007;14(2):175-81.

Yates JR, 3rd. Mass spectral analysis in proteomics. *Annu Rev Biophys Biomol Struct.* 2004;33:297-316.

Zhang W, Chait BT. ProFound: an expert system for protein identification using mass spectrometric peptide mapping information. *Anal Chem.* 2000 Jun 1;72(11):2482-9.

Zolg W. The proteomic search for diagnostic biomarkers: lost in translation? *Mol Cell Proteomics.* 2006 Oct;5(10):1720-6.

7 – ANEXOS

7.1. Anexo I

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Título do Projeto: *Análise proteômica do fluido ductal mamário e estudo de mutação no gene TP53 em células presentes neste fluido em casos de câncer de mama.*

Investigador: Dante Pagnoncelli

Local da Pesquisa: Instituto Fernandes Figueira - FIOCRUZ

Endereço e telefone: Av. Rui Barbosa, 716 - Flamengo - Rio de Janeiro - RJ

Tel: (21) 2553-0052 - **Fax:** (21) 2553-6730

Você está sendo convidada a participar de uma pesquisa, coordenada por um profissional de saúde agora denominado pesquisador. Para poder participar, é necessário que você leia este documento com atenção. Ele pode conter palavras que você não entende. Por favor, peça aos responsáveis pelo estudo para explicar qualquer palavra ou procedimento que você não entenda claramente.

O propósito deste documento é dar a você as informações sobre a pesquisa e, se assinado, dará a sua permissão para participar no estudo. O documento descreve o objetivo, procedimentos, benefícios e eventuais riscos ou desconfortos caso queira participar. Você só deve participar do estudo se você quiser. Você pode se recusar a participar ou se retirar deste estudo a qualquer momento.

O Objetivo deste estudo é analisar a composição de proteínas do fluido mamário e pesquisar mutações (alterações) em um gene chamado TP53 nas células nele presentes. Aproveitaremos também este fluido para fazer uma pesquisa de alterações nas células que vem neste fluido da mesma maneira que se faz na coleta de material para “preventivo” de câncer de colo de útero.

Você será entrevistada e examinada para que possamos classificar seu caso conforme suas características próprias e de sua família e assim proceder a uma análise comparativa entre os grupos. Um primeiro grupo é daquelas mulheres com casos de câncer de mama comprovado por biópsia ou cirurgia, o segundo grupo é o de mulheres que possuem história de câncer na família que possam ser classificadas em um grupo de alto risco e que estejam em atendimento no Ambulatório de Alto Risco do IFF do Serviço de Mastologia, e o terceiro grupo é aquele sem história ou suspeita de câncer de mama e com mamografia normal.

A coleta do material a ser examinado, o aspirado de fluido de mama, é feito por meio de coleta não invasiva, sem dor e sem desconforto e facilmente realizável. Ele consiste na aplicação de uma compressa morna por 5 minutos, seguida por massagem suave da mama pela própria paciente durante mais 5 minutos, e após a limpeza do mamilo com álcool, é colocada uma bomba de aspiração (do tipo que se usa para extração de leite materno) e, se houver saída de pequenas gotas, estas serão colhidas e imediatamente armazenadas para posterior análise.

Não são conhecidos danos relacionados ao procedimento de coleta do fluido de mama.

As amostras coletadas serão encaminhadas aos laboratórios responsáveis pelos estudos:

1. Laboratório de Biologia Molecular de Tumores do Departamento de Biologia Celular e Genética da UERJ, sob a responsabilidade da Profa. Cláudia Vitória de Moura Gallo, para estudo de mutações (alterações) no gene TP53

2. Laboratório de Toxinologia do Departamento de Fisiologia e Farmacodinâmica da FIOCRUZ, sob a responsabilidade do Prof. Jonas Perales para estudo das proteínas que estão contidas neste fluido;

O armazenamento e a guarda do material ficarão a cargo do Laboratório de Biologia Molecular de Tumores do Departamento de Biologia Celular e Genética da UERJ, sob a responsabilidade da Profa. Cláudia Vitória de Moura Gallo.

O estudo das proteínas no fluido de mama e a pesquisa mutação do gene TP53 nas células recolhidas, embora seja uma nova área de pesquisa com promessas para o futuro, no momento não vão revelar nenhum resultado que possa afetar sua escolha de tratamento e ou diagnóstico ou prognóstico para câncer de mama.

O único teste que pode ser útil para a voluntária é a citopatologia oncológica, que é um exame semelhante ao preventivo de colo de útero (Papanicolau). Embora raramente possa ser encontrado algum resultado que tenha valor prático, caso achemos algum indício neste exame que possa estar associado ao câncer de mama, nos obrigamos a comunicar imediatamente o resultado somente à você ou a quem for de sua indicação para receber esta informação assim como para seu médico no IFF, para que as providências para o esclarecimento da dúvida levantada sejam tomadas.

É também importante lembrar que um resultado negativo nesta citologia não afasta e nem diminui o risco de câncer de mama, não servindo como substituto de nenhum dos métodos reconhecidos e indicados de rastreamento para o câncer de mama, que você deve continuar a ser submetida.

Sua decisão em participar deste estudo é voluntária. Você pode decidir não participar no estudo. Uma vez que você decidiu participar do estudo, você pode retirar seu consentimento e participação a qualquer momento. Se você decidir não continuar no estudo e retirar sua participação, você não será punida ou perderá qualquer benefício ao qual você tem direito.

Não haverá nenhum custo a você relacionado aos procedimentos previstos no estudo.

Sua participação é voluntária, portanto você não será pago por sua participação neste estudo.

O Investigador responsável pelo estudo e equipe irão coletar informações sobre você. Em todos esses registros um código substituirá seu nome. Todos os dados coletados serão mantidos de forma confidencial. Os dados coletados serão usados para a avaliação do estudo, membros das Autoridades de Saúde ou do Comitê de Ética, podem revisar os dados fornecidos. Os dados também podem ser usados em publicações científicas sobre o assunto pesquisado. Porém, sua identidade não será revelada em qualquer circunstância.

Você tem direito de acesso aos seus dados. Você pode discutir esta questão mais adiante com seu médico do estudo.

Se você ou seu(s) parente(s) tiver(em) alguma dúvida com relação ao estudo, direitos do paciente, ou no caso de danos relacionados ao estudo, você deve contatar o Investigador do estudo ou sua equipe (Dr. Dante Pagnoncelli – tel: 2553-0052 – ramal da ginecologia). Se você tiver dúvidas sobre seus direitos como um paciente de pesquisa, você pode contatar Comitê de Ética em Pesquisa do Instituto Fernandes Figueira (telefone: 2554-1730). Este comitê é composto de um grupo de indivíduos com conhecimento científicos e não científicos que realizam a revisão ética inicial e continuada do estudo de pesquisa para mantê-lo seguro e proteger seus direitos.

Declaração de consentimento do paciente:

Eu li e discuti com o investigador responsável pelo presente estudo os detalhes descritos neste documento. Entendo que eu sou livre para aceitar ou recusar, e que eu posso interromper minha participação a qualquer momento sem dar uma razão, sem nenhum prejuízo no caso de estar em tratamento médico. Eu concordo que os dados coletados para o estudo sejam usados para o propósito acima descrito.

Estou sendo informada que minha participação não afetará minha escolha de tratamento e ou diagnóstico e prognóstico do câncer de mama.

Eu entendi a informação apresentada neste termo de consentimento. Eu tive a oportunidade para fazer perguntas e todas as minhas perguntas foram respondidas.

Eu concordo em participar do projeto: Análise proteômica do fluido ductal mamário e estudo de mutação em gene p53 em células colhidas deste fluido em casos de câncer de mama

Eu receberei uma cópia assinada e datada deste Documento de Consentimento Informado.

1. Dou minha permissão aos pesquisadores deste projeto para me constatarem por telefone ou através de mensagem pelo correio se informações adicionais forem necessárias, conforme informações que constam em meu prontuário no IFF. () SIM () NÃO

2. Eu gostaria de ser informada sobre o resultado da citologia oncológica que pode fornecer algum resultado para câncer de mama. () SIM () NÃO

3. Eu gostaria que a pessoa abaixo relacionada fosse informada sobre o resultado da citologia oncológica que pode fornecer algum resultado para câncer de mama. () SIM () NÃO

Nome da voluntária: _____

Telefones: _____ Prontuário: _____

_____ Identidade: _____

ASSINATURA DA VOLUNTÁRIA

Nome da pessoa referenciada: _____

Endereço: _____

Tel: _____ Condição do(a) referenciado (a): _____

ASSINATURA DA PESSOA REFERENCIADA

Eu expliquei a Sra _____ a proposta desta pesquisa e os procedimentos necessários à coleta das amostras e os riscos e benefícios deste estudo.

ASSINATURA DO PESQUISADOR

DATA: _____

7.2. Anexo II

Parecer Consubstanciado de Projeto de Pesquisa

Título do Projeto: Análise proteômica do fluido ductal mamário e estudo de mutação em células deste fluido em casos de câncer de mama

Pesquisador Responsável Dante Pagnoncelli

Data da Versão 20/07/2006

Cadastro FR101935

Data do Parecer 05/10/2006

Grupo Área Temática I.1 Genética Humana

Objetivos do Projeto

Realizar análise comparativa da composição de proteínas do fluido ductal mamário e da presença de mutações no gene TP53 em material colhido de três grupos de mulheres: aquelas com câncer de mama unilateral, mulheres sem câncer de mama, porém portadoras de mutação nos genes BRCA1 e/ou BRCA2 e mulheres sem história familiar de câncer de mama e com mamografia normal.

Sumário do Projeto

Este projeto é relativo à investigação do câncer de mama através de novas técnicas moleculares, como a análise proteômica de fluido ductal mamário, como forma de avaliar a utilização desta nova tecnologia como ferramenta de detecção precoce do tumor e de acompanhamento da progressão do câncer de mama.

Itens Metodológicos e Éticos	Situação
Título	Adequado
Autores	Adequados
Local de Origem na Instituição	Adequado
Projeto elaborado por patrocinador	Não
Aprovação no país de origem	Não necessita
Local de Realização	Própria instituição
Outras instituições envolvidas	Sim
Condições para realização	Adequadas

Comentários sobre os itens de Identificação

Os itens relativos à identificação do projeto foram corretamente apresentados.

Introdução	Adequada
------------	----------

Comentários sobre a Introdução

O Projeto está bem formulado e em sua introdução o pesquisador responsável apresenta, de forma clara e justificada, a relevância de sua proposta para a área.

Objetivos	Adequados
-----------	-----------

Comentários sobre os Objetivos

O projeto apresentado tem por objetivo testar a aplicação de novas metodologias moleculares, como a análise proteômica de fluido ductal mamário, como ferramenta de detecção precoce e acompanhamento da progressão do câncer de mama. A justificativa para a realização deste projeto é atual e está embasada em dados da literatura.

Pacientes e Métodos	
Delineamento	Adequado
Tamanho de amostra	Total 200 Local
Cálculo do tamanho da amostra	Não informado
Participantes pertencentes a grupos especiais	Não
Seleção equitativa dos indivíduos participantes	Ausente
Crêrios de inclusão e exclusão	Adequados
Relação risco- benefício	Adequada
Uso de placebo	Não utiliza
Período de suspensão de uso de drogas (wash out)	Não utiliza
Monitoramento da segurança e dados	Adequado



Coordenador do Comitê de Ética em Pesquisa com Seres Humanos
 INSTITUTO FERNANDES FIGUEIRA - IFF/FIOCRUZ

APROVADO Página 1-2

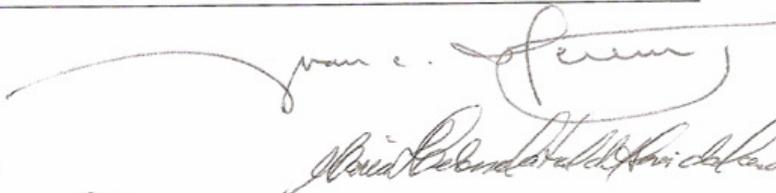
Válido Até 30/12/2008

Comitê de Ética em Pesquisa com Seres Humanos
 INSTITUTO FERNANDES FIGUEIRA - IFF/FIOCRUZ

Fonte de financiamento externa	Agência de fomento
Comentários sobre o Cronograma e o Orçamento	
Ver comentários gerais.	
Referências Bibliográficas	Adequadas
Comentários sobre as Referências Bibliográficas	

Recomendação
 Aprovar

Comentários Gerais sobre o Projeto
Anexar um documento tanto da equipe técnica da análise proteômica (laboratório de Toxicologia - Dept. de Fisiologia e Farmacodinâmica do IOC - FIOCRUZ) como da equipe técnica da análise do gene TP53 (Laboratório de Biologia Molecular de Tumores do Dept. de Biologia Celular e Genética da UERJ) com o compromisso junto ao pesquisador principal da disponibilidade de verbas para realização e término da pesquisa, apesar de explícito tais compromissos no Material e Métodos do projeto apresentado ao CEP/IFF-FIOCRUZ.



APROVADO

Válido Até 30/12/2008

Comitê de Ética em Pesquisa com Seres Humanos
 INSTITUTO FERNANDES FIGUEIRA - IFF/FIOCRUZ
 Telefone: 2552-8491 / 2554-1700 r. 1730

Coordenador do Comitê de Ética em Pesquisa com Seres Humanos
 INSTITUTO FERNANDES FIGUEIRA - IFF/FIOCRUZ

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)