

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE BIOCÊNCIAS

CURSO DE BIOTECNOLOGIA

DEBORAH DA CRUZ SCHAFHAUSER

AVALIAÇÃO DO USO DE PEPTÍDEOS URINÁRIOS COMO BIOMARCADORES PARA
O DIAGNÓSTICO DO CÂNCER DE PRÓSTATA

Porto Alegre

2019

DEBORAH DA CRUZ SCHAFHAUSER

AVALIAÇÃO DO USO DE PEPTÍDEOS URINÁRIOS COMO BIOMARCADORES PARA
O DIAGNÓSTICO DO CÂNCER DE PRÓSTATA

Trabalho de Conclusão de curso apresentado como
requisito parcial para obtenção do título de Bacharel em
Biotecnologia com ênfase em Biotecnologia Molecular na
Universidade Federal do Rio Grande do Sul.

Orientador(a): Prof^ª. Dr^ª. Karina Mariante Monteiro

Porto Alegre

2019

DEBORAH DA CRUZ SCHAFHAUSER

AVALIAÇÃO DO USO DE PEPTÍDEOS URINÁRIOS COMO BIOMARCADORES PARA
O DIAGNÓSTICO DO CÂNCER DE PRÓSTATA

Trabalho de Conclusão de Curso, apresentado
como requisito parcial para obtenção de grau de
Bacharel em Biotecnologia Molecular, pela
Universidade Federal do Rio Grande do Sul.

Data da Avaliação: __/__/__

Banca Examinadora

Profa. Dra. Karina Mariante Monteiro

Orientadora – Universidade Federal do Rio Grande do Sul

Profa. Dra. Fernanda Bueno Morrone

Banca – Pontifca Universidade Católica do Rio Grande do Sul

Prof. Dr. Charley Christian Staats

Banca – Universidade Federal do Rio Grande do Sul

Este trabalho foi realizado no Laboratório de Genômica Estrutural e Funcional do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul, com apoio financeiro da Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), do Programa Pesquisa para o SUS (PPSUS), do Departamento de Ciência e Tecnologia da Secretaria de Ciência, Tecnologia e Insumos Estratégicos do Ministério da Saúde (Decit/SCTIE/MS), da Secretaria da Saúde do Rio Grande do Sul (SES/RS) e do Programa de Iniciação Científica BIC/UFRGS.

Agradecimentos

Este trabalho finaliza, mas não define, os anos dedicados à graduação. Todas as frustrações me trouxeram aprendizados, assim como todas as risadas me mostraram que a vida também pode ser leve. Agradeço primeiramente aos meus pais, Jair e Maria, que nunca tiveram a chance de cursar o ensino superior, mas dedicaram tempo e trabalho para que eu pudesse ter esta e tantas oportunidades. Sempre com muito amor, se fizeram presentes nos momentos de maior dificuldade e é por causa deles que me tornei quem sou. Agradeço também àqueles que passaram pela minha vida e deixaram sua marca. Aos amigos que emprestaram seus ombros, ouvidos e abraços, me ensinaram muito sobre diversidade, força e gentileza.

Agradeço imensamente à Professora Karina, que me aceitou como orientada e me guiou pelos experimentos, procurando sempre soluções alternativas, sem deixar que as dificuldades me abalassem. A todo laboratório 206, 204 e 210 que me acolheu com muito carinho, principalmente à Cris que me acompanhou a cada passo. Agradeço à Gabriela Paludo, por me auxiliar com tanta paciência e dedicação.

Por fim, agradeço a Deus, por nunca me deixar cair.

Resumo

O Câncer de Próstata (CaP) é uma doença de alto impacto socioeconômico no Brasil, sendo o segundo tipo de câncer mais incidente na população masculina, impactando a qualidade de vida, assim como os investimentos em saúde pública. O diagnóstico é confirmado a partir da biópsia do tecido prostático, e seu encaminhamento é relativo ao exame de toque retal suspeito e aos valores de PSA aumentados (do inglês: *prostate-specific antigen*), contudo, poucos homens realizam o exame de toque e os valores de PSA podem estar relacionados a questões benignas da próstata, se mostrando pouco específico para câncer. Portanto, a utilização de outro componente biológico como biomarcador mais específico para a condição de adenocarcinoma prostático geraria grandes benefícios. Este trabalho buscou na urina de pacientes com suspeita de diagnóstico para adenocarcinoma de próstata e hiperplasia benigna da próstata (HPB), peptídeos circulantes que estivessem em diferente abundância comparado ao grupo controle.

As amostras coletadas tiveram as proteínas depletadas e foram submetidas a cromatografia líquida acoplada a espectrometria de massas. Os resultados apresentaram ao menos um peptídeo diferencialmente abundante entre cada grupo de estudo, sendo um deles exclusivo do grupo controle. Por fim, quatro peptídeos demonstraram potencial para biomarcador para diagnóstico de patologia da próstata. Curvas ROC (*Receiver Operating Characteristic Curve*) foram executadas para analisar a sensibilidade e especificidade destes peptídeos para diferenciar os grupos.

O peptídeo que apresentou maior área sob a curva (0,697), estava diferencialmente abundante na comparação entre os grupos câncer e controle com dados normalizados a partir da mediana. Identificamos um peptídeo que exclusivamente diferenciava hiperplasia benigna de câncer, que é a maior problemática associada ao PSA, atualmente. Mais análises precisam ser feitas utilizando um grupo amostral maior e uma técnica mais sensível para validar a força preditiva dos peptídeos.

Palavras-chave: câncer de próstata, biomarcador, peptidômica, diagnóstico.

Abstract

Prostate Cancer is a disease with high social economic impact in Brazil, for being the second most incident type of cancer at the male population, impacting life quality, as the public investment in wealth. Its diagnostic is confirmed with prostatic tissue biopsy, and it is given by a suspect digital rectal exam and by the raise of PSA values (prostate-specific antigen), however, few man perform the digital rectal exam and PSA values may be related to benign prostate conditions, being low specific to cancer. Therefore, using another biological component as a more specific biomarker to prostate adenocarcinoma would generate great benefits. This work looked for circulating peptides at different abundances between prostate cancer, benign prostate hyperplasia and control groups, using patients with suspect diagnose urine as sample source.

All collected samples had their proteins depleted and were subjected to liquid chromatography coupled to mass spectrometry. Results presented at least one differentially abundant peptide between each study group, one of them being unique to the control group. Finally, four peptides demonstrated potential for biomarker for diagnosis of prostate pathology. Receiver Operating Characteristic Curves were performed to analyze the sensitivity and specificity of these peptides to differentiate the groups.

The peptide with the largest area under the curve (0.697) was differentially abundant when comparing cancer and control groups with normalized data from the median. We identified a peptide that uniquely differentiated benign hyperplasia from cancer, which is the biggest problem associated with PSA today. Further analysis needs to be done using a larger sample group and a more sensitive technique to validate the predictive strength of peptides.

Key-words: prostate cancer, biomarker, peptidomics, diagnostic.

Sumário

1. Introdução.....	8
1.1. Câncer de Próstata.....	8
1.2. Diagnóstico e Tratamento de CaP.....	9
1.3. Biomarcadores.....	12
1.4. Peptídeos Biomarcadores na Urina.....	13
1.5. Validação de Biomarcadores Clínicos.....	14
2. Justificativa.....	15
3. Objetivos.....	16
3.1. Objetivo Geral.....	16
3.2. Objetivos Específicos.....	16
4. Metodologia.....	16
4.1. Obtenção das Amostras.....	16
4.2. Definição dos Grupos.....	16
4.3. Isolamento dos Peptídeos Urinários.....	18
4.4. Análise do Peptidoma Urinário.....	18
4.5. Processamento e análise dos dados de MS.....	18
4.6. Análises Estatísticas.....	19
5. Resultados e Discussão.....	19
6. Conclusão e Perspectiva.....	32
7. Referências.....	32
Anexo 1	38

1. Introdução

1.1. Câncer de próstata

Atualmente, o câncer é a segunda principal causa de mortes no mundo, assim como a segunda principal causa de morbidade e mortes prematuras (índices *YLL-years of life*, ou anos de vida, e *DALY-disability-adjusted life year*, ou anos de vida ajustados por incapacidade), segundo a Organização Mundial de Saúde (OMS). Esses números implicam em uma cobrança de reação nas políticas públicas para garantir acesso a saúde e proteção aos pacientes com câncer, garantindo o diagnóstico precoce e o tratamento adequado. Para tanto, o Plano de Ação Global da OMS para controle e prevenção do câncer e outras doenças não transmissíveis se concentra no objetivo da redução de mortalidade prematura (GLOBAL BURDEN OF DISEASE CANCER COLLABORATION, 2019).

O Câncer de Próstata (CaP) é o fator causador de 358 mil mortes no mundo (THE GLOBAL CANCER OBSERVATORY, 2018) e no Brasil é o segundo tipo de câncer mais incidente entre os homens, abaixo apenas do câncer de pele não melanoma, sendo estimados 68.220 novos casos da doença para cada ano do biênio 2018-2019 (INCA, 2017). Segundo dados do DATASUS, apenas em 2012, foram gastos aproximadamente R\$ 150 milhões com o tratamento do CaP, sendo R\$ 127 milhões com hormonioterapia de primeira e segunda linha, R\$ 20 milhões com quimioterapia e R\$ 1,5 milhões com orquiectomias. Portanto, o câncer de próstata possui grande importância em termos de saúde pública, desde seu diagnóstico até o tratamento em si (DINI; KOFF, 2006).

Até o início deste século, diversos estudos buscaram compreender o plano de fundo genético associado a determinadas doenças utilizando o estudo filogenético familiar com associação a condições de dominância seguindo o modelo Mendeliano. Algumas associações foram obtidas para o câncer, contudo, o uso de amostragem familiar nestes casos não são suficientes para compreender as heranças recessivas, traços de associação gênica e as diferentes vias oncológicas associadas ao fenótipo (LANDER; SCHORK, 1994; LICHTENSTEIN et al., 2000). Portanto, pouco se conhecia até então sobre a relação destas vias e o câncer de próstata, dificultando o desenvolvimento de ferramentas mais tecnológicas para seu diagnóstico. Hábitos como o consumo de cigarros e bebida alcoólica são comumente associados ao aumento do risco do desenvolvimento de determinados tipos de câncer. Para o câncer de próstata a associação com tabaco e o aumento da taxa de risco foi negativo em estudos de uma variável no Reino Unido (JACOB et al., 2018). Por outro lado, o CaP é altamente associado a hereditariedade: em 2003 foi encontrada uma taxa de risco partindo de 1,8 vezes

maior em desenvolver câncer de próstata para homens com histórico familiar de CaP, em qualquer grau de parentesco, comparado a homens sem o mesmo histórico na família (BRUNER et al., 2003). Também foi possível traçar a etiologia genética mais associada com os casos de câncer de próstata, sendo a descendência Afro-Americana a de maior incidência, seguida da Europeia e Asiática.

Com as análises de associação genômica ampla (ou GWAS, do inglês: *genome-wide association study*), o estudo das populações com maior incidência de CaP abriu espaço para a análise de variantes associadas a doença. Nesses estudos foram encontrados mais de 160 *loci* comuns a suscetibilidade ao câncer de próstata, o que sugere associação poligênica. Esses dados permitiram o cálculo do escore de risco poligênico do CaP, mas também revelam a necessidade da meta-análise e reanálise dos dados, a fim de melhor compreender as vias biológicas envolvidas, diferenciar as relações de causalidade entre fatores, e para encontrar fatores clinicamente significativos para compreensão, diagnóstico e tratamento do CaP (FARASHI et al., 2019).

1.2. Diagnóstico e tratamento de CaP

Diferentes morbidades podem acometer a glândula da próstata: inflamação ou prostatite, hiperplasia benigna e câncer. Apesar de serem doenças diferentes, estudos indicam que inflamações recorrentes da próstata, seja por mudanças hormonais, infecções bacterianas ou virais, respostas imunes ou refluxo urinário, causam danos ao tecido da glândula. Os processos para tratar os danos causados pelas infecções crônicas podem gerar mudanças estruturais no tecido e o aumento da próstata, o que pode levar ao desenvolvimento da hiperplasia benigna e/ou do carcinoma (VERZE; CAI; LORENZETTI, 2016). Para o diagnóstico em tempo hábil para o tratamento, é recomendação do sistema de saúde que se mantenha uma rotina de exames periódicos de toque retal para avaliar hiperplasia da glândula prostática em homens a partir dos 45 anos de idade. Por questões socioculturais, poucos brasileiros realizam o exame preventivo (MOREIRA, 2012). Em 2015, de acordo com pesquisa do Datafolha em parceria com a Sociedade Brasileira de Urologia, por exemplo, apenas 32% da população masculina realizam o exame de toque retal no Brasil.

O antígeno prostático específico, ou PSA (do inglês: *prostate-specific antigen*), é secretado pela próstata em condições normais. Ele pode ser detectado no sangue e sua detecção em níveis elevados é associada a condições anormais da próstata a partir dos anos

1990, quando o teste passou a ser difundido. Um aumento no diagnóstico de câncer de próstata a partir desta época é atribuído a detecção precoce de casos (TABAYOYONG; ABOUASSALY, 2015). Contudo, o PSA tem limitações importantes como a falta de especificidade. Hiperplasia Prostática Benigna (HPB), prostatites e até mesmo infecções do trato urinário, podem causar uma elevação do PSA sérico (LILJA; ULMERT; VICKERS, 2008). Essa baixa especificidade leva a um excesso de falsos positivos, aumentando os custos em saúde e diminuindo a qualidade de vida do paciente (HEIJNSDIJK et al., 2015). Por isso, hoje a dosagem do PSA sérico é complementar ao exame de toque retal e auxilia no diagnóstico precoce da doença. Pacientes com níveis séricos de PSA aumentados ou exame de toque retal suspeito são então submetidos a biópsia por agulha para a confirmação histopatológica da doença (LITWIN; TAN, 2017).

Em 2012, a agência americana *Food and Drug Administration* (FDA) registrou o teste PCA3 como ferramenta de auxílio no encaminhamento de biópsias de próstata para detecção de CaP, este teste detecta a presença de um lncRNA em amostras de urina (SANTOS, 2015). O gene codificador do lncRNA PCA3 (gene *PCA3*, do inglês, *Prostate Cancer Antigen 3*) é comprovadamente superexpresso em casos de câncer de próstata e o lncRNA sintetizado por ele pode ser encontrado em maior quantidade na urina dos pacientes com adenocarcinoma. É recomendado que o teste seja efetuado após o estímulo da próstata pelo toque, no recolhimento da primeira amostra de urina após o exame. Para os casos em que seja provável a necessidade de realizar o procedimento de biópsia uma segunda vez, o reconhecimento de baixos níveis de PCA3 (> 20) associados a baixos níveis de PSA sérico (> 4 ng/ml), poderiam dispensar o procedimento para 8% da população em estudo, contudo, ainda há riscos de subdiagnósticos (WEI et al., 2014).

A primeira técnica descrita de diagnóstico de câncer de próstata foi a biópsia transperineal aberta, por Young e Davis em 1926. Esta apresentava menores taxas de infecção que a biópsia usando a rota transretal, mas era necessário que o tumor estivesse presente em 95% do tecido para que a detecção fosse possível, além de resultar em altas taxas de incontinência urinária, disfunção erétil e hospitalizações (SIVARAMAN et al., 2016). Atualmente a técnica considerada padrão ouro na biópsia da próstata é a transperineal ou transretal ecoguiada, sendo a transperineal menos utilizada. Ela se constitui pelo uso de uma agulha que retira de 6 a 12 amostras da glândula e esse número depende do volume prostático, se caso o volume esteja em torno de 40 ml, por exemplo, se recomenda a retirada de no mínimo 8 amostras (LOPES et al., 2015). A ultrassonografia comumente é usada para acompanhamento

e guia do processo de biópsia, contudo seu baixo contraste dificulta a diferenciação entre tecido normal e cancerígeno, o que pode levar a biópsia de regiões de diagnóstico incerto e uma possível repetição do procedimento, levando ao máximo desconforto do paciente. Uma alternativa promissora é a Ressonância Magnética Multiparamétrica, por produzir melhor visualização de imagem, reduzindo o número de regiões de amostragem necessárias, e permitindo um diagnóstico mais acurado. Contudo, essa técnica ainda possui um custo muito elevado para ampla utilização (BRIZMOHUN APPAYYA et al., 2018; BROWN et al., 2015).

A próstata passa por várias etapas de modificação para enfim desenvolver o nível de malignidade do tecido, sendo o primeiro deles classificado como neoplasia intraepitelial de próstata (PIN), que pode ser de baixo ou alto grau. O PIN significa que foram encontradas células que parecem ser câncer, mas que elas não aparentam estar em outras regiões da próstata. Outro possível diagnóstico é a proliferação atípica de pequenos ácinos (ASAP), que significa que as células aparentam como CaP, mas estão em pouca quantidade. Por fim, o diagnóstico da biópsia pode também relatar atrofia inflamatória atípica, quando as células simplesmente se apresentam menores que o normal (“Biópsia para Diagnóstico do Câncer de Próstata - Instituto Oncoguia”, [s.d.]). Donald Gleason, em 1974, desenvolveu uma grade para avaliação do padrão histológico do adenocarcinoma de próstata, a qual constitui dois níveis de grade, predominante e secundário, ambos em escala de 1 a 5 (LITWIN; TAN, 2017; SHEN; ABATE-SHEN, 2010) (Tabela 1). Por exemplo, uma biópsia que apresente escala Gleason 7 (3 + 4) significa que o tecido maligno mais abundante tem malignidade nível 3, mas a parte menos abundante das amostras biopsiadas apresenta células de malignidade nível 4.

Tabela 1: Sistema de Graduação Patológica definido pela Sociedade Internacional de Patologias Urológicas.

Sistema de Graduação Patológica - Câncer de Próstata		
Nível	Escala Gleason	
Câncer nível 1	(3+3)	6
Câncer nível 2	(3+4)	7
Câncer nível 3	(4+3)	
Câncer nível 4	(4+4)	8
	(3+5)	
Câncer nível 5	(5+3)	9
	(4+5)	
	(5+4)	
	(5+5)	10

Fonte: <http://www.nomograms.org>

O tratamento para o câncer de próstata depende de alguns fatores individuais do paciente, como sua expectativa de vida e se o mesmo possui outras morbidades de risco mais eminente, que sobressaiam o risco provocado pelo CaP. Essa decisão é feita por um conselho médico junto do paciente. As técnicas de tratamento podem ser usadas concomitantemente e são, em casos não metastáticos, cirurgia, radioterapia e observação vigilante (LITWIN; TAN, 2017); em casos de metástase, no Brasil é recomendado a radioterapia associada a tratamento hormonal, com outros paliativos (FRASÃO; SAÚDE, [s.d.]). A observação vigilante acontece para pacientes com o nível de PSA sérico menor que 10 ng/ml, dentro da escala de Gleason em 6 e consiste em uma série de testes de PSA, exames físicos e biópsias, para analisar se o carcinoma evolui e se torna necessário outro tratamento.

1.3.Biomarcadores

O termo “biomarcador” é usado em publicações desde meados de 1970, mas nos últimos 15 anos o uso do termo aumentou exponencialmente. O Grupo de Trabalho em Definição de Biomarcadores (ou, em inglês: *Biomarkers Definitions Working Group*) definiu biomarcador como “uma característica que é objetivamente medida e avaliada como um indicador de processos biológicos normais, processos patogênicos ou respostas farmacológicas a uma intervenção terapêutica”, mas essa definição é criticada por alguns pesquisadores. Hoje, um biomarcador pode ser classificado como: molecular, celular, tecidual, de órgão e de corpo inteiro. Os biomarcadores são usados para a triagem de doenças, para o diagnóstico, para o desenvolvimento de intervenções terapêuticas individualizadas, para identificar tipos celulares específicos e outras muitas necessidades clínicas (ARONSON; FERNER, 2017).

A validação de um biomarcador para uso clínico é uma dificuldade discutida pelos pesquisadores que propõem novas moléculas para esta finalidade. Para ser considerado um bom biomarcador, a molécula precisa descrever com confiança se determinado paciente se encontra em determinado grupo diagnóstico, e assim permitir uma terapia adequada ao mesmo (FREIDLIN; KORN, 2013). Piccoli & Gaforolo (2018) escreveram uma discussão sobre os problemas de padronização métrica para desenvolvimento de novos biomarcadores, isso devido a diversidade de tipos moleculares, tamanho e finalidade, associado à falta de diretrizes regulatórias para esse fim. A FDA idealizou os testes de farmacocinética (PK)

como validadores para biomarcadores, mas se auto-refutou em 2013 afirmando a diferença inerente entre os ensaios. Com o avanço tecnológico, como as técnicas de cromatografia líquida

e espectrometria de massas, que identificam macromoléculas, como peptídeos, e *Ligand-binding assay* (ou ensaio de ligação ao ligante), que é usada quando as moléculas alvo são muito grandes, muito hidrofílicas e difíceis de detectar usando outras técnicas mais robustas (“Ligand Binding Assays – Certara”, [s.d.]), prospectar potenciais biomarcadores não é mais um problema. Porém, parâmetros para definir uma validação, considerando contexto do uso e técnica analítica disponível, são necessários para que se chegue aos ensaios clínicos.

Técnicas de proteômica tem possibilitado uma busca mais sensível para macromoléculas presentes em determinadas condições fisiológicas, como o câncer. Estudos primordiais com câncer de próstata encontraram diferenças na abundância de proteínas entre a condição e células pancreáticas normais, mas os ensaios ainda possuíam limitações (CAZARES et al., 2002). Diferentes moléculas são sintetizadas e excretadas por células em condições fisiológicas diferentes, e estudos proteômicos abriram a possibilidade de se encontrar biomarcadores para estas condições (WULFKUHLE; LIOTTA; PETRICOIN, 2003). Os primeiros peptídeos descobertos como agentes bioativos circulantes foram classificados como hormônio, mais especificadamente, a secretina, que é um neuropeptídeo excretado pelo pâncreas (SEKAR; CHOW, 2013), e logo sua sequência de aminoácidos foi decifrada. Uma busca por diferentes peptídeos presentes nos organismos fundou o termo peptidômica, mas na época, pouco se diferenciava da proteômica, porque muitos dos peptídeos encontrados eram fruto de proteínas degradadas, por isso, técnicas de isolamento dos peptídeos, como colunas de afinidade, foram desenvolvidos. Este conceito define a diferença entre as técnicas até a atualidade, pois para executar uma análise de proteômica, o extrato proteico da amostra deve ser digerido por uma enzima, comumente a tripsina, enquanto a peptidômica não requer digestão, além disso, a peptidômica propõem identificar peptídeos de ocorrência natural, endógenos à amostra (SCHRADER; SCHULZ-KNAPPE; FRICKER, 2014).

1.4. Peptídeos Biomarcadores na Urina

Entendendo que mudanças fisiológicas alteram o padrão de abundância de proteínas/peptídeos excretados pelas células, assim como de proteases, que podem atuar na clivagem de proteínas circulantes -gerando peptídeos, a aplicação da peptidômica em fluidos corporais de fácil acesso permitiria avaliar a presença de biomarcadores em análises não invasivas. Sangue/Soro/Plasma, urina, fluido lacrimal e, mais recentemente, cera de ouvido (BARBOSA et al., 2019), podem ser fonte de informações fisiológicas e patológicas em estados

pouco sintomáticos (BROZA et al., 2019). A urina é um fluido de fácil acesso e de obtenção abundante, além de ser descartado de resíduos biológicos, e já foi demonstrada como fonte de informações relativas a diferentes processos biológicos (BRÖKER et al., 2013; HARPOLE; DAVIS; ESPINA, 2016). Por ser armazenada na bexiga por horas, antes de ser eliminada, grande parte da proteólise acontece no órgão e, por isso, quando eliminada é bastante estável, facilitando o processo de armazenamento. Considerando a proteólise de proteínas urinárias na bexiga, excreção diferencial de enzimas proteolíticas de órgãos como a próstata, em condições patológicas, podem influenciar na presença de determinados peptídeos no fluido. Estudos já mostraram a presença de excreções prostáticas na urina, e puderam relacionar diferentes resultados em casos de CaP e HPB (BROZA et al., 2019). Além disso, muitas análises proteômicas compararam resultados a partir de amostras de urina contra plasma e puderam relacionar as abundâncias obtidas com qualidade de filtração renal, ademais, peptidômicas apresentam baixa obtenção amostral em amostras de plasma e de urina e uma baixa sobreposição dos resultados, demonstrando uma alta taxa de exclusividade de peptídeos em cada uma das amostras (MAGALHÃES et al., 2018). Até então, apenas três trabalhos na plataforma PubMed buscaram distinguir CaP e HPB usando espectrometria de massas em amostras de urina (DAVALIEVA et al., 2015; HAJ-AHMAD; ABDALLA; HAJ-AHMAD, 2014; JEDINAK et al., 2015).

Análises peptidômicas do mesmo indivíduo durante o período de um dia apresentam uma variabilidade de até 30% no tipo e abundância dos peptídeos obtidos pela urina (KOLCH et al., 2005), o que implica que em determinadas coortes de indivíduos haverá grande variedade entre os peptídeos encontrados nas análises. Uma possibilidade de biomarcador ideal neste caso, portanto, não se resumiria a apenas uma molécula, mas a uma associação entre biomarcadores e/ou condições que padronizem uma análise para um diagnóstico seguro.

1.5. Validação de Biomarcadores Clínicos

Como qualquer processo para validar um diagnóstico, biomarcadores moleculares precisam ser submetidos a análises estatísticas para confirmar sua capacidade de diferenciar dois grupos devido a fatores fisiológicos inerentes ao caso, não por questões meramente estocásticas. Isso porque a correlação entre dois fatores nem sempre está associada a um fator causal, e isso pode acarretar em perdas significativas de capacidade de diagnóstico ao se tratar de biomarcadores (ENSOR, 2014).

Quando o biomarcador precisa ser clinicamente validado, existem 3 requerimentos para a proposta: a função e a utilidade do marcador devem ser claras, precisa ser específico o suficiente para não causar dúvidas e sua sensibilidade deve ser confiável. Além disso, os ensaios de execução das análises devem possuir “Validação Analítica”, ou seja, todos os processos prévios das análises devem ser protocolados e passarem pelo mesmo processo, para que se mantenha um ensaio robusto e reproduzível. Por isso, quando se propõem um potencial biomarcador para ensaios clínicos, é preciso descrever e manter as condições de coleta, processamento, armazenamento e preparação da amostra, além dos reagentes e condições do protocolo. Quando os testes serão utilizados em amostras arquivadas, ou que não foram submetidas ao protocolo final, as análises devem ser executadas com ensaios cegos para os dados clínicos dos pacientes (SIMON; PAIK; HAYES, 2009). Ainda existem biomarcadores binários, que quando presentes correspondem a determinada condição exclusivamente, e não-binários, que são relativizados pela sua abundância na amostra. Para cada tipo de marcador, existe um cálculo amostral para sua validação mas algumas estatísticas procuram unificar a análise (POLLEY, 2017).

2. Justificativa

O câncer de próstata (CaP) é uma doença de alto impacto social e econômico no Brasil. Os métodos atualmente disponíveis para o diagnóstico da doença não atingem de forma satisfatória a população masculina no país. O biomarcador atualmente utilizado, o PSA, pode apresentar muitos falsos positivos, aumentando o custo para o Sistema Único de Saúde, além de submeter os pacientes a um tratamento que diminui a sua qualidade de vida desnecessariamente. Portanto, descobrir um biomarcador de maior precisão para o CaP geraria grande impacto nas tomadas de decisão em relação ao tratamento da doença, trazendo enorme benefício para a população masculina brasileira.

A urina é uma amostra de coleta fácil e não invasiva, cuja composição de moléculas reflete o estado fisiológico/patológico dos principais tecidos urológicos. Sendo assim, a urina possui grande potencial para ser utilizada como fonte de biomarcadores para CaP. O presente projeto propõe a análise em larga-escala de peptídeos circulantes na urina de pacientes com CaP, hiperplasia prostática benigna (HPB) e indivíduos saudáveis (sem diagnóstico de CaP ou HPB) para a identificação de potenciais biomarcadores para o CaP. A partir dos resultados obtidos, espera-se contribuir com moléculas candidatas para utilização no

desenvolvimento de futuros insumos diagnósticos, prognósticos e preditivos para o câncer de próstata.

3. Objetivos

3.1. Objetivo Geral

Avaliar o perfil qualitativo e quantitativo de peptídeos circulantes na urina de pacientes com câncer de próstata, hiperplasia prostática benigna e indivíduos saudáveis para a identificação de potenciais biomarcadores para o CaP.

3.2. Objetivos Específicos

- Coletar amostras de urina e dados clínicos e sociodemográficos de pacientes com CaP, HPB e indivíduos saudáveis;
- Purificar os peptídeos urinários a partir das amostras biológicas dos diferentes grupos de estudo;
- Analisar os peptídeos por espectrometria de massas;
- Identificar peptídeos diferencialmente abundantes entre os grupos de estudo;
- Avaliação preliminar do potencial diagnóstico dos peptídeos com presença/abundância diferencial entre os grupos.

4. Metodologia

4.1. Coleta de amostras biológicas

Foram convidados a participar da pesquisa indivíduos do sexo masculino, com idade a partir de 40 anos, com suspeita de diagnóstico de câncer de próstata ou hiperplasia prostática benigna, assim como voluntários fora do grupo de risco. Os sujeitos passaram por uma breve entrevista e realizaram uma coleta de urina, sem que fosse a primeira da manhã. Os pacientes foram recrutados no Hospital Ernesto Dornelles (Porto Alegre/RS) e no Hospital Ana Nery (Santa Cruz do Sul/RS). Os indivíduos que concordam em participar do estudo manifestam o seu consentimento através da assinatura de TCLE. O estudo foi aprovado pelo CEP das instituições participantes (CAAE 69852617.1.0000.5347). As amostras foram centrifugadas a 4000 rpm por 10 minutos e armazenadas a -80 °C.

4.2. Definição dos grupos de estudo

A partir dos dados informados pelos pacientes e dos laudos das biópsias de próstata, as amostras foram categorizadas em quatro grupos: adenocarcinoma de próstata (grupo 1), hiperplasia benigna da próstata (grupo 2), pacientes que apresentaram crescimento acinar atípico, inflamação ou neointraepitelial de alto grau focal (grupo 3) e controles (grupo 4) (Tabela 2).

Tabela 2: Tabela Referência para Definição dos Grupos. Tabela utilizada para definir os grupos de análise a partir dos resultados de biópsia, valores de PSA e resultado do exame de toque retal.

Tabela Referência para os Grupos					
Diagnóstico - Classificação	Sub-definições		Condições		
Adenocarcinoma (1)	1.A	Ou	Diagnóstico	1 (Adenocarcinoma)	
			Gleason	6 (3+3) 7 (3+4)	
		1.B	Ou	Diagnóstico	1 (Adenocarcinoma)
	Gleason			7 (4+3) 8 (4+4)	
	HPB (2)		2.A	Ou	Diagnóstico
		Toque Suspeito			0 (ignorado, ou não)
PSA >		4			
Diagnóstico		0 (Negativo)			
Ou		Toque Suspeito		1 Sim	
2.B		Ou	PSA <	4	
			Diagnóstico	0 (Negativo)	
		Ou	Toque Suspeito	1 (Sim)	
		Ou	PSA >	4	
			Toque Suspeito	0 (Não ou ignorado)	
Outros (3)	3	Ou	PSA >	10	
			Diagnóstico	2 (HPB)	
			Diagnóstico	3 (inflamação)	
Controle (4)	4		Diagnóstico	3 (ASAP)	
			Diagnóstico	3 (PIN-alto)	
			Diagnóstico	- Não suspeito	
			PSA <	4	

Os grupos “Adenocarcinoma (1)” e “HPB (2)” foram divididos em subcategorias. Para os casos de adenocarcinoma da próstata, foi considerado o grau de malignidade do câncer, sendo Gleason 6 (3+3) e 7 (3+4) -categorias 1 e 2, grupo A, e Gleason

7 (4+3) e 8 (4+4) -categoria 3 e 4, grupo B. Para categorizar os casos de hiperplasia benigna da próstata, foram considerados os diagnósticos da análise histopatológica para HPB, além da relação entre toque retal suspeito e valor de PSA total. Indivíduos que não informaram o resultado do exame de toque retal foram incluídos nas categorias daqueles que obtiveram resultado negativo para o mesmo exame. Pacientes com exame de toque retal negativo e nível sérico de PSA total acima de 4 ng/mL foram classificados no grupo 2A, ou de menor risco, enquanto os indivíduos com níveis de PSA total maiores que 10 ng/mL foram incluídos no grupo 2B, considerado grupo de maior risco. Indivíduos com exame de toque retal suspeito, mas com níveis séricos de PSA total abaixo de 4 ng/mL também foram classificados no grupo 2A, enquanto indivíduos com exame de toque retal suspeito e níveis séricos de PSA total acima de 4 ng/mL foram incluídos no grupo 2B.

4.3. Isolamento dos peptídeos urinários

O isolamento de peptídeos foi realizado conforme descrito por Parker et al. (2017), que comparou a recuperação de peptídeos por diferentes técnicas. As técnicas foram padronizadas no laboratório para amostras de urina, e aquela que recuperou maior abundância de peptídeos foi de PBS-TCA, o que condiz com os resultados no artigo de referência.

As amostras de urina foram misturadas com 1 volume de 20% de tricloroacético (TCA) e PBS e incubados durante 60 min a 4 °C para precipitação de proteínas. O precipitado foi centrifugado a 4.000 x g durante 10 min e o sobrenadante contendo peptídeos foi recolhido. Os peptídeos foram então dessalinizados em coluna HLB Oasis (Waters).

4.4. Análise do Peptidoma Urinário

Os peptídeos foram analisados por LC-MS/MS usando um sistema nanoACQUITY UPLC acoplado a um espectrômetro de massa Xevo G2-XS Q-Tof (Waters) com uma sonda de baixo fluxo na fonte, na Unidade de Química de Proteínas e Espectrometria de Massas do Centro de Biotecnologia da UFRGS. Os peptídeos foram separados por cromatografia analítica (Acquity UPLC BEH C18, 1,7 µm, 2,1 × 50 mm, Waters), a uma vazão de 8 µl/min, usando um gradiente linear de 7-85% de água/ACN 0,1% de ácido fórmico por 45 min. As amostras foram analisadas em duplicatas.

4.5. Processamento e análise dos dados de MS

Os dados foram processados e comparados usando o ProteinLynx Global version 2.5 (PLGS, Waters Corporation). As comparações/alinhamentos foram conduzidas usando como referência sequências de proteínas humanas obtidas pelo UniProt/Swiss-Prot, e a tolerância foi ajustada automaticamente (normalmente 10 ppm para precursor e 20 ppm para produtos iônicos). Nenhuma enzima foi selecionada, e foram consideradas as modificações hidroxil em resíduos de ácido aspártico, lisina, asparagina e prolina, e oxidação da metionina, como possíveis na análise.

Os peptídeos identificados foram filtrados baseados na classificação do *match* obtido a partir da busca contra do banco de dados (verde, OK; amarelo, *maybe*; e vermelho, *poor*). Os peptídeos classificados como “verde, OK” foram selecionados e suas intensidades relativas e tempos de retenção foram organizados em um tabela utilizando um *script* (**Anexo 1**) gerado em colaboração com a Unidade de Biologia Teórica e Computacional (UBTEC) do Centro de Biotecnologia da Universidade Federal do Rio Grande do Sul. Além disso, foram excluídos peptídeos presentes em ≤ 2 amostras e amostras/replicatas sem nenhuma identificação de peptídeos (*valid values*). Os dados foram normalizados utilizando a ferramenta online NormalyzerDE (<http://quantitativeproteomics.org/normalyzerde>; Willforss *et al.*, 2019) e os seguintes critérios: dados já transformados para log e *threshold* de baixa abundância igual a 2. Subsequentemente, as amostras foram classificadas nos diferentes grupos de estudo (grupos 1, 2, 3 e 4), conforme descrito na Tabela 2, e os dados normalizados foram submetidos a análises de expressão diferencial no NormalyzerDE.

4.6. Análises Estatísticas

A partir dos resultados de análise diferencial, os dados normalizados foram inseridos no programa *SPSS Statistics* e foi executada a função Curva ROC para os peptídeos com abundância ou presença diferencial entre os grupos comparados. O valor de especificidade e sensibilidade máxima é obtido a partir do ponto no qual a soma das duas variáveis apresenta o maior valor (YOU DEN, 1950). Os gráficos *BoxPlot* também foram gerados usando o programa *SPSS Statistics*.

5. Resultados e Discussão

Um total de cento e sete (107) amostras foram coletadas e classificadas nos grupos previamente definidos (Tabela 3). Como alguns dos subgrupos definidos, como, por exemplo, os subgrupos de adenocarcinoma (C1 e C2), possuíam um número muito díspar de amostras, as análises subsequentes foram realizadas utilizando as definições amplas dos grupos (grupos 1, 2, 3 e 4). Na Tabela 4 são mostrados os dados sociodemográficos dos participantes do estudo, de acordo com a classificação clínica.

Tabela 3: Tabela de Distribuição das amostras na classificação geral dos grupos de estudo. Classificação dos grupos de estudo e o número de amostras associado a cada um deles.

Tabela de Identificação Geral dos Grupos	
Grupo	Identificação
Adenocarcinoma (n=34)	1
HPB (n=40)	2
Outros (n=6)	3
Controle (n= 27)	4

Tabela 4: Dados Sociodemográficos dos Participantes da Pesquisa. Dados informados em questionário entregue aos participantes da pesquisa, sendo informados pela tabela: Média de idade, Autodeclaração Étnica, Escolaridade, Estado Civil e Histórico Familiar de Câncer de Próstata; de acordo com os grupos classificatórios de câncer de próstata (CaP), hiperplasia benigna da próstata (HPB), Outros e Controle.

Análises Descritivas

Grupo	Média de Idade	Autodeclaração Étnica					Escolaridade				
		Branca	Negra	Parda	Amarela	Não declarado	Ensino Fundamental	Ensino Médio	Ensino Superior	Analfabeto	Faltam dados
CaP (n= 34)	69	88%	0%	3%	0%	9%	35%	35%	24%	6%	0%
HPB (n= 40)	64	78%	18%	3%	0%	3%	30%	10%	33%	0%	28%
Outros (n= 6)	64	67%	17%	17%	0%	0%	17%	33%	50%	0%	0%
Controle (n= 27)	57	26%	0%	4%	0%	70%	4%	0%	26%	0%	70%

Grupo	Estado Civil					Histórico Familiar Câncer de Próstata		
	Casado	Separado	Viúvo	Solteiro	União Estável	Faltam dados	Sim	Não

CaP (n= 34)	85%	9%	6%	0%	0%	0%	29%	68%	3%
HPB (n= 40)	85%	5%	8%	3%	0%	0%	33%	65%	3%
Outros (n= 6)	33%	0%	17%	0%	50%	0%	33%	50%	17%
Controle (n= 27)	22%	4%	0%	4%	0%	70%	7%	22%	70%

A análise dos peptídeos urinários por LC-MS/MS resultou na identificação de 127 peptídeos. Os peptídeos identificados eram, majoritariamente, provenientes da proteína uromodulina. A uromodulina é uma das proteínas mais abundantes na urina, está associada com a proteção contra infecções urinárias e já foi proposta como biomarcador para danos renais (YOUHANNA et al., 2013). Peptídeos derivados de proteína kinase associada a morte celular (DAPK2) foram identificados em um pequeno número de amostras de CaP e HPB. Supreendentemente, identificamos poucos peptídeos derivados de colágeno, como descrito em estudos anteriores de peptídeos urinários (BRÖKER et al., 2013). A baixa abundância de peptídeos nas análises pode ser consequência de diferença metodológica, podendo ser na conservação e preparação das amostras. A *Human Proteome Organization* (do inglês: Organização de Proteoma Humano) sugere que a urina coletada seja a segunda da manhã, além de que caso a urina seja congelada por mais de 4 horas, se adicione ácido bórico (0,2 M) para preservação da amostra (HUMAN PROTEOME ORGANIZATION, [s.d.]), sendo, portanto, uma das possíveis abordagens para evitar degradação da amostra. Além disso, alguns estudos que identificaram peptídeos na urina prepararam a amostra por meio de ultrafiltração e análise por meio de CE-MS (GOOD et al., 2010; SIWY et al., 2011; THEODORESCU et al., 2008), e isso pode acarretar diferença na detecção de peptídeos. Ademais, estudos que detectaram peptídeos diferencialmente abundantes na urina de pacientes com doença renal não são comparáveis com o caso de CaP, pois o dano no rim acarreta diretamente na baixa filtração de proteínas que chegarão na bexiga. Devido ao risco de infecção pelo procedimento da biópsia, é recomendado aos pacientes a administração de antibióticos no período prévio à coleta do tecido (FIUK et al., 2015), o que também pode influenciar na abundância de peptídeos na urina.

Após a filtração dos dados, dados de MS referentes a 16 peptídeos e 95 amostras foram submetidos à normalização utilizando a ferramenta NormalyzerDE. Foram avaliados diferentes métodos de normalização de dados, incluindo métodos globais (mediana, média e quantile) e locais (log2, CyclicLoess e Regressão Linear Robusta – RLR) de normalização.

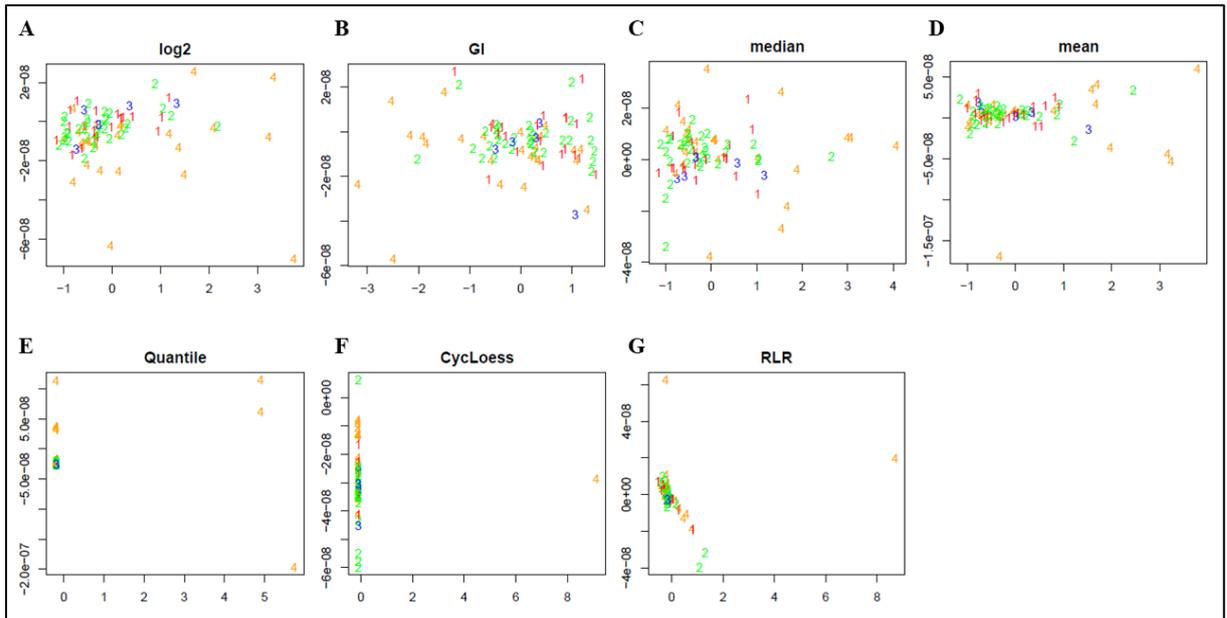


Figura 1: Gráficos de escalonamento multidimensional dos resultados de MS normalizados utilizando diferentes metodologias. Análises tipo MDS (*Multidimensional scaling*) dos dados normalizados utilizando log2 (A), Intensidade Global (B), Mediana (C), Média (D), Quantile (E), CycLoess (F) e RLR (G).

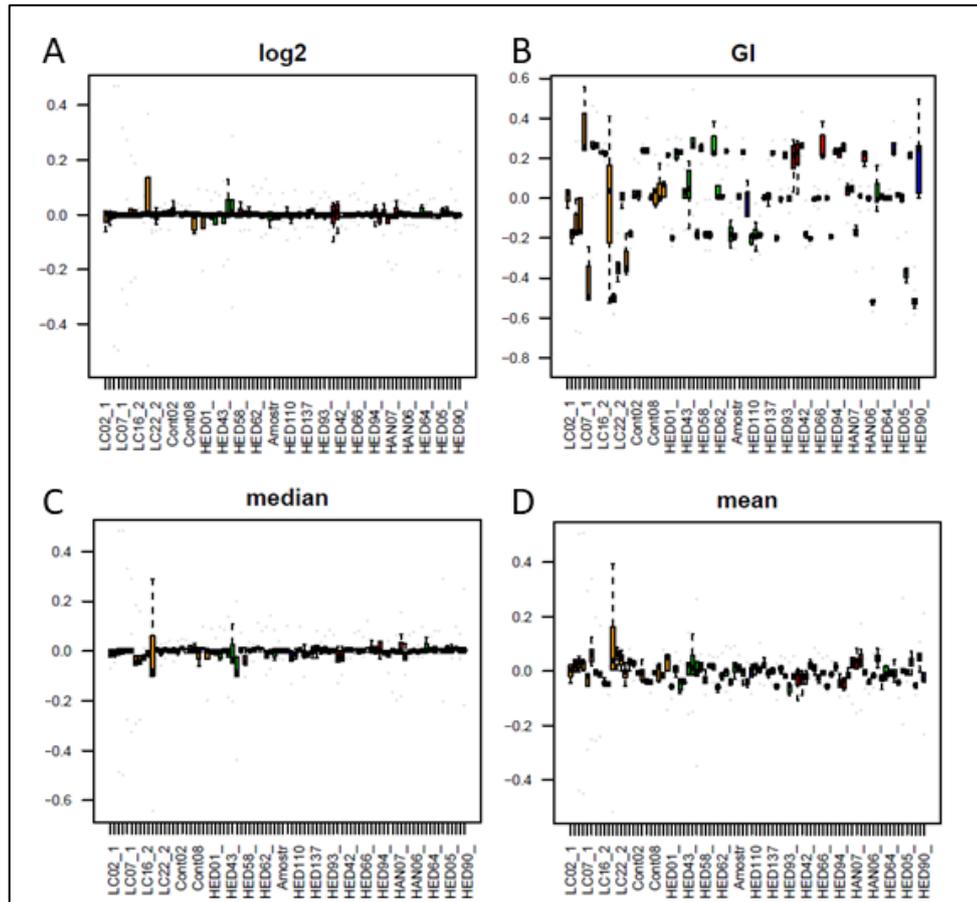


Figura 2: Gráficos de expressão relativa comparando os diferentes métodos de normalização dos dados de MS. Análises do tipo RLE (do inglês, *relative log expression*) a partir dos dados normalizados utilizando os métodos de log2 (A), GI (B), Mediana (C) e média (D).

A Figura 1 mostra os gráficos gerados a partir dos dados normalizados pelas diferentes metodologias. As análises de escalonamento multidimensional (MDS, do inglês *multidimensional scaling*) permitem representar graficamente n elementos em uma plotagem 2D levando em conta a distância ou a similaridade que os elementos têm entre si (WICKELMAIER; WICKELMAIER, 2003). Portanto, foram escolhidos os métodos de normalização que separaram de forma mais evidente as amostras ao longo do gráfico: logarítmica (log2), Intensidade Global (GI), Mediana (median) e Média (mean) (**Figura 1A, B, C e D**, respectivamente). Contudo, é evidente que em nenhum método de normalização utilizado foi possível observar a separação das amostras por grupo de estudo, demonstrando que o conjunto de peptídeos identificado (variáveis) não é capaz de promover a total diferenciação entre os grupos. Esse resultado é esperado devido ao baixo número de peptídeos utilizado na análise e a alta variação observada entre os indivíduos (amostras biológicas).

A análise dos gráficos dos valores de expressão relativa (RLE, do inglês *relative log expression*) mostraram que os dados normalizados pelos métodos log2 e mediana apresentam melhor centralização (valores alinhados próximo a zero) do que os dados normalizados pelos métodos GI e média, os quais apresentaram desvios sugestivos de discrepâncias nos dados (Figura 2). Os gráficos RLE surgiram a partir das análises de microarranjo para mostrar variações indesejadas na amostra, portanto em dados melhor normalizados, ou seja, que podem ser distribuídos em uma curva normal, a linearidade das amostras deve ser o mais homogênea possível dentro de seu grupo amostral (GANDOLFO; SPEED, 2018).

A análise de expressão diferencial foi realizada no programa NormalyzerDE a partir dos dados normalizados pelos métodos log2 e mediana. Para a realização dessas análises, foram acrescentadas ao conjunto de dados as amostras que não apresentaram nenhum peptídeo em ambas as duplicatas técnicas da análise de LC-MS/MS. O conjunto amostral para essas análises totalizou 101 amostras, sendo 32 CaP (grupo 1), 38 HPB (grupo 2), 6 outros (grupo 3) e 25 controles (grupo 4).

A análise de expressão diferencial realizada com os valores normalizados utilizando \log_2 identificaram 5 peptídeos diferencialmente representados ($p < 0,05$) entre os grupos de estudo, 3 deles entre CaP e controles e 2 deles entre HPB e controles (Figura 3 e Tabela 5). É possível observar que alguns peptídeos (peptídeo 13 e 15) foram identificados como diferencialmente expressos em mais de uma comparação entre os grupos (CaP x controles e HPB x controles). Nas análises realizadas com os dados normalizados pela mediana, 8 peptídeos diferencialmente representados entre os grupos de estudo foram identificados (Figura 4 e Tabela 5), sendo 2 entre CaP e controles, 2 entre CaP e HPB, 1 entre CaP e outros, 1 entre HPB e controles, 1 entre HPB e outros, e 1 entre outros e controles. Novamente é possível observar que alguns peptídeos (peptídeo 10, 13 e 15) tem presença ou abundância diferencial em mais de uma comparação entre os grupos. Sendo assim, no total, considerando os dois métodos de normalização de dados e todas as comparações entre os grupos, identificamos 4 peptídeos com potencial de diferenciação entre os grupos de estudo (peptídeos 3, 10, 13 e 15).

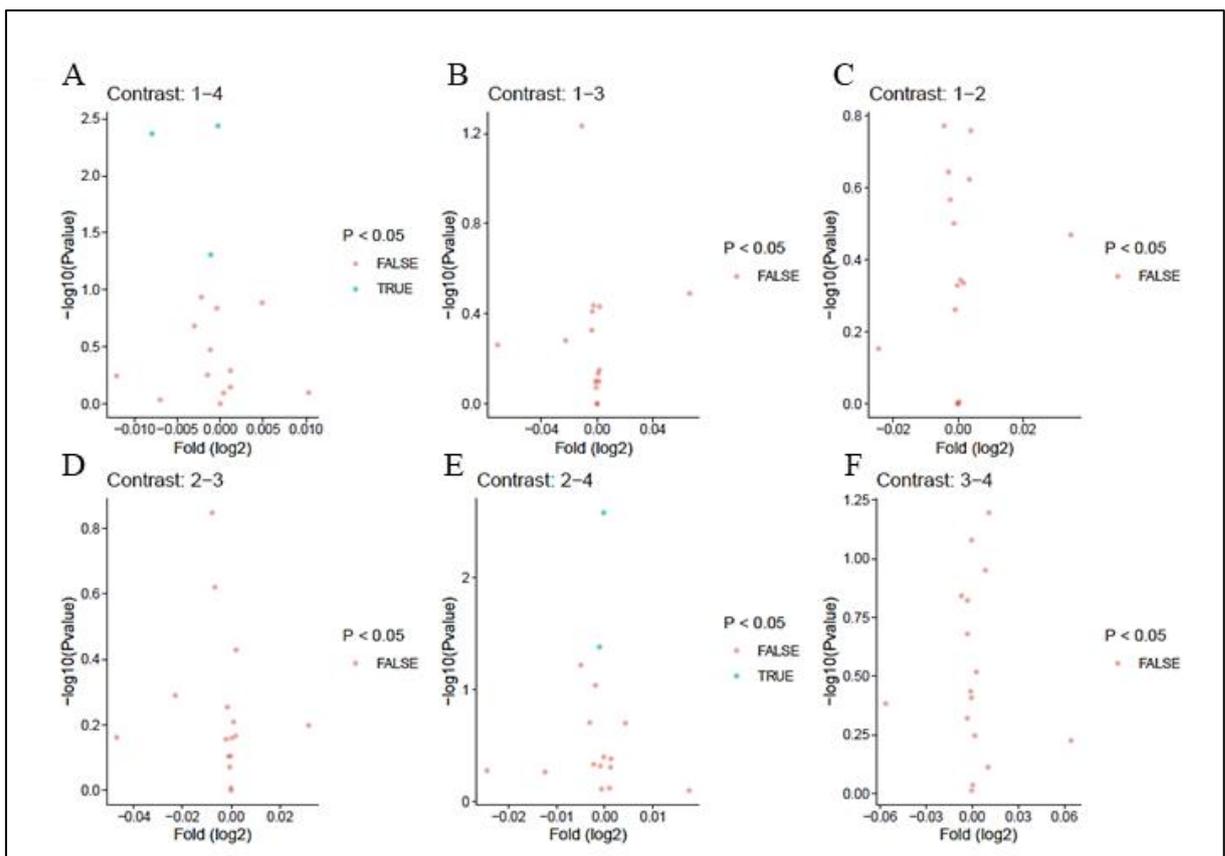


Figura 3: Gráficos do tipo Volcano mostrando a abundância dos peptídeos entre os diferentes grupos de estudo a partir dos dados normalizados por \log_2 . Em vermelho são mostrados os peptídeos sem diferença entre os grupos, enquanto em azul estão destacados os peptídeos com abundância/presença diferencial entre os grupos

de estudo. (A) CaP x controles; (B) CaP x outros; (C) CaP x HPB; (D) HPB x outros; (E) HPB x controles; (F) outros x controles.

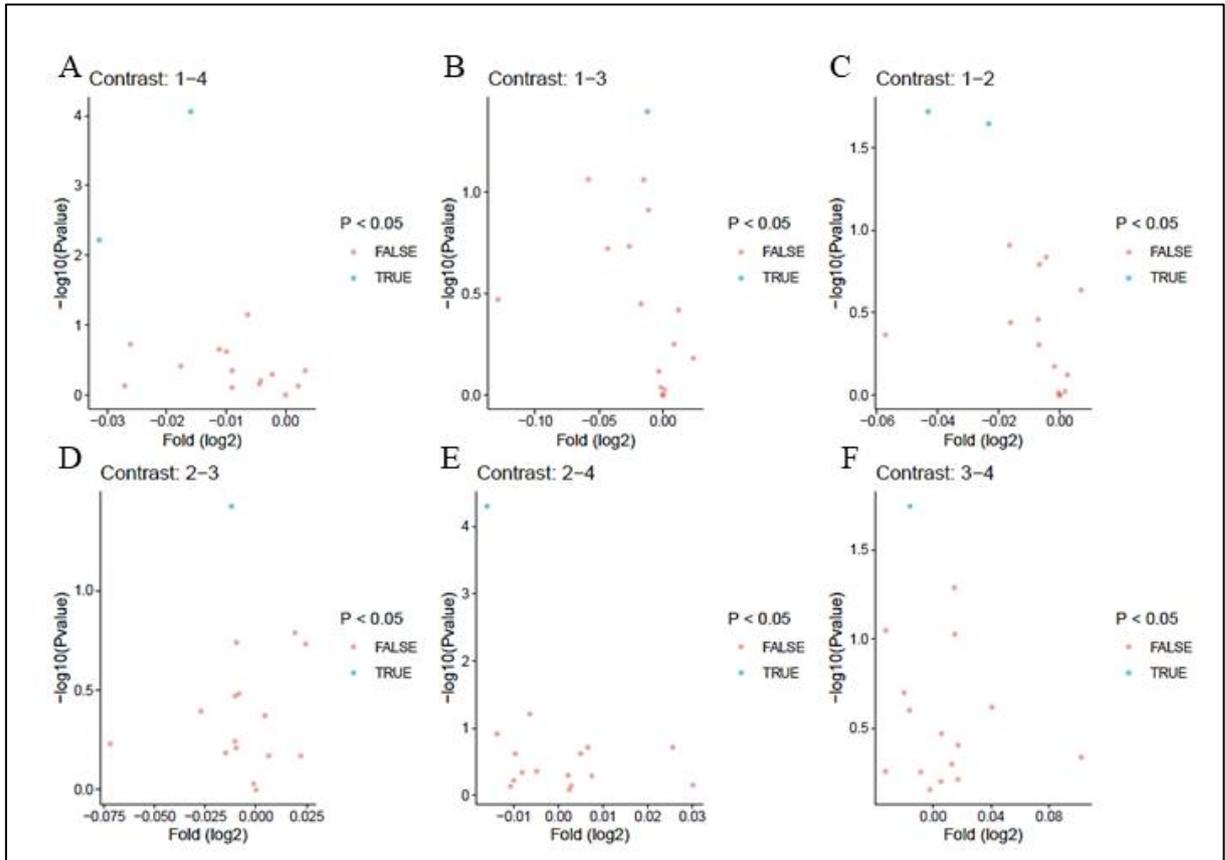


Figura 4: Gráficos do tipo Volcano mostrando a abundância dos peptídeos entre os diferentes grupos de estudo a partir dos dados normalizados por mediana. Em vermelho são mostrados os peptídeos sem diferença entre os grupos, enquanto em azul estão destacados os peptídeos com abundância/presença diferencial entre os grupos de estudo. (A) CaP x controles; (B) CaP x outros; (C) CaP x HPB; (D) HPB x outros; (E) HPB x controles; (F) outros x controles.

Na Figura 5 e 6 é possível observar que entre os peptídeos diferencialmente representados entre os grupos, alguns apresentam diferente abundância (intensidade), enquanto outros possuem presença específica em um determinado grupo. Por exemplo, enquanto os peptídeos 3 e 10 possuem abundância diferencial entre os grupos, os peptídeos 13 e 15 possuem presença específica nas amostras do grupo controle (peptídeo 13), e controle e outros (peptídeo 15). Como é possível observar nas Figuras 5B, 6B e 5C, 6C, esses peptídeos identificados exclusivamente nas amostras controle/outras estão presentes em um número limitado de amostras. O peptídeo 13 está presente em 6 amostras do grupo controle ($n = 25$), enquanto o peptídeo 15 está presente em 2 amostras do grupo controle ($n = 25$) e 1 amostra do grupo outros

(n = 6). Esses resultados refletem a variação biológica das amostras dentro de um mesmo grupo e podem ter um impacto negativo no potencial diagnóstico desses peptídeos. Porém, peptídeos exclusivamente identificados nas amostras controle podem ser muito interessantes para a discriminação entre indivíduos saudáveis e doentes (HPB, CaP e outros).

Tabela 5: Peptídeos com valor de p significativo em cada grupo de estudo. Tabela contendo os peptídeos encontrados diferencialmente expressos ($p < 0,05$) em cada grupo de estudo, de acordo com a normalização dos dados, sendo explicitado o valor de p e o valor de p ajustado específico de cada peptídeo. Os valores significativos estatisticamente foram sublinhados.

Peptídeos Estatisticamente Diferenciais			
Normalização pela Mediana			
Grupos Comparados	Peptídeo	Valor de p	Valor de p, ajustado
CaP e Controle	Peptídeo 10	<u>0,006035624</u>	<u>0,048284989</u>
	Peptídeo 13	<u>8,70253E-05</u>	<u>0,001392404</u>
CaP e HPB	Peptídeo 10	<u>0,019102467</u>	0,181176119
	Peptídeo 03	<u>0,022647015</u>	0,181176119
CaP e Outros	Peptídeo 15	<u>0,040185172</u>	0,465311366
HPB e Controle	Peptídeo 13	<u>4,98519E-05</u>	<u>0,00079763</u>
HPB e Outros	Peptídeo 15	<u>0,037769893</u>	0,604318288
Outros e Controle	Peptídeo 13	<u>0,018097008</u>	0,289552126
Normalização por log2			
Grupos Comparados	Peptídeo	Valor de p	Valor de p, ajustado
CaP e Controle	Peptídeo 10	<u>0,004269639</u>	<u>0,03415711</u>
	Peptídeo 13	<u>0,003644951</u>	<u>0,03415711</u>
	Peptídeo 15	<u>0,049176249</u>	0,26227333
CaP e HPB	-	-	-
CaP e Outros	-	-	-
HPB e Controle	Peptídeo 13	<u>0,002622317</u>	<u>0,041957076</u>
	Peptídeo 15	<u>0,041606054</u>	0,322681573
HPB e Outros	-	-	-
Outros e Controle	-	-	-

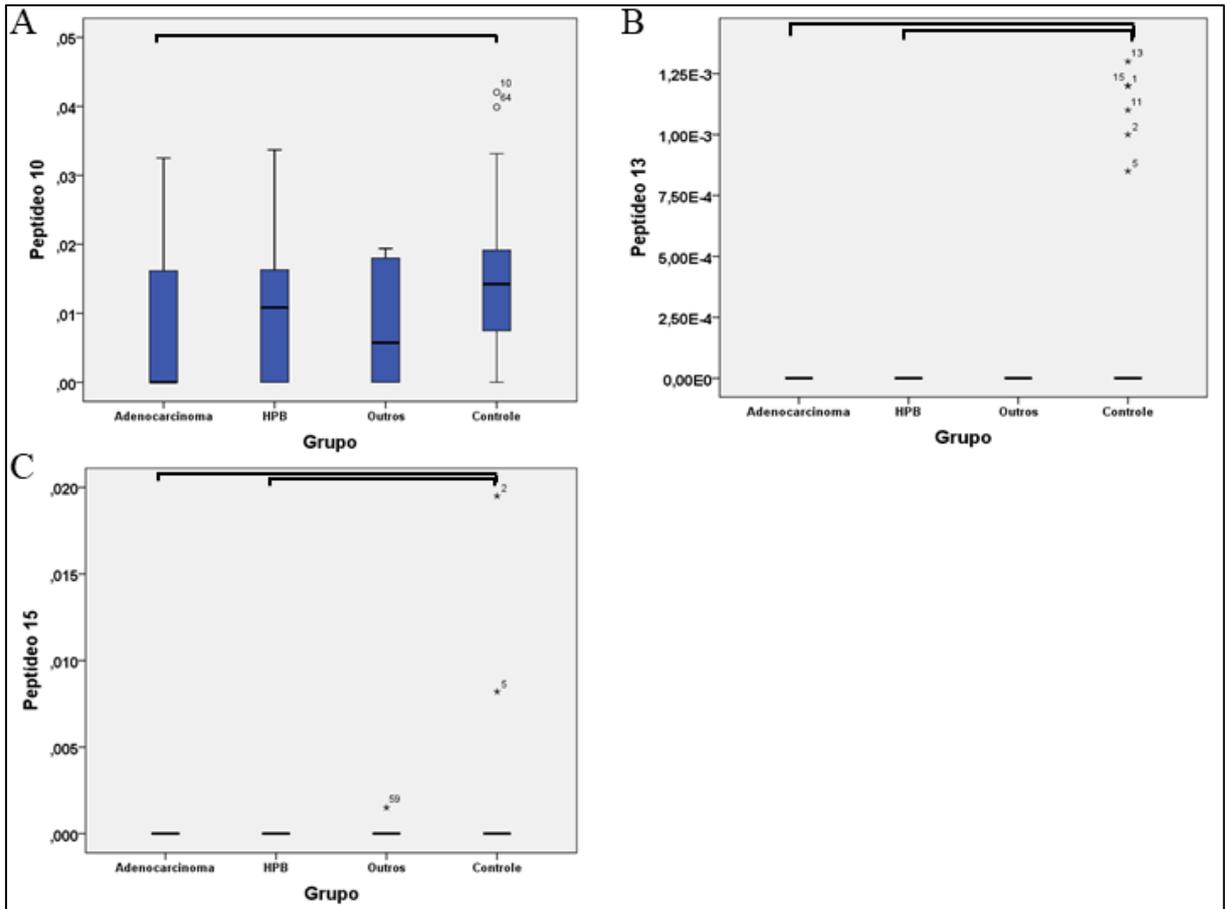


Figura 5: Boxplot dos valores de log2 para os peptídeos diferencialmente expressos. Gráfico Boxplot dos valores de log2 dos peptídeos com valor de $p < 0,05$, sendo destacados pelas barras comparativas entre os grupos quando diferencialmente expressos. A: Peptídeo 10, B: Peptídeo 13, C: Peptídeo 15.

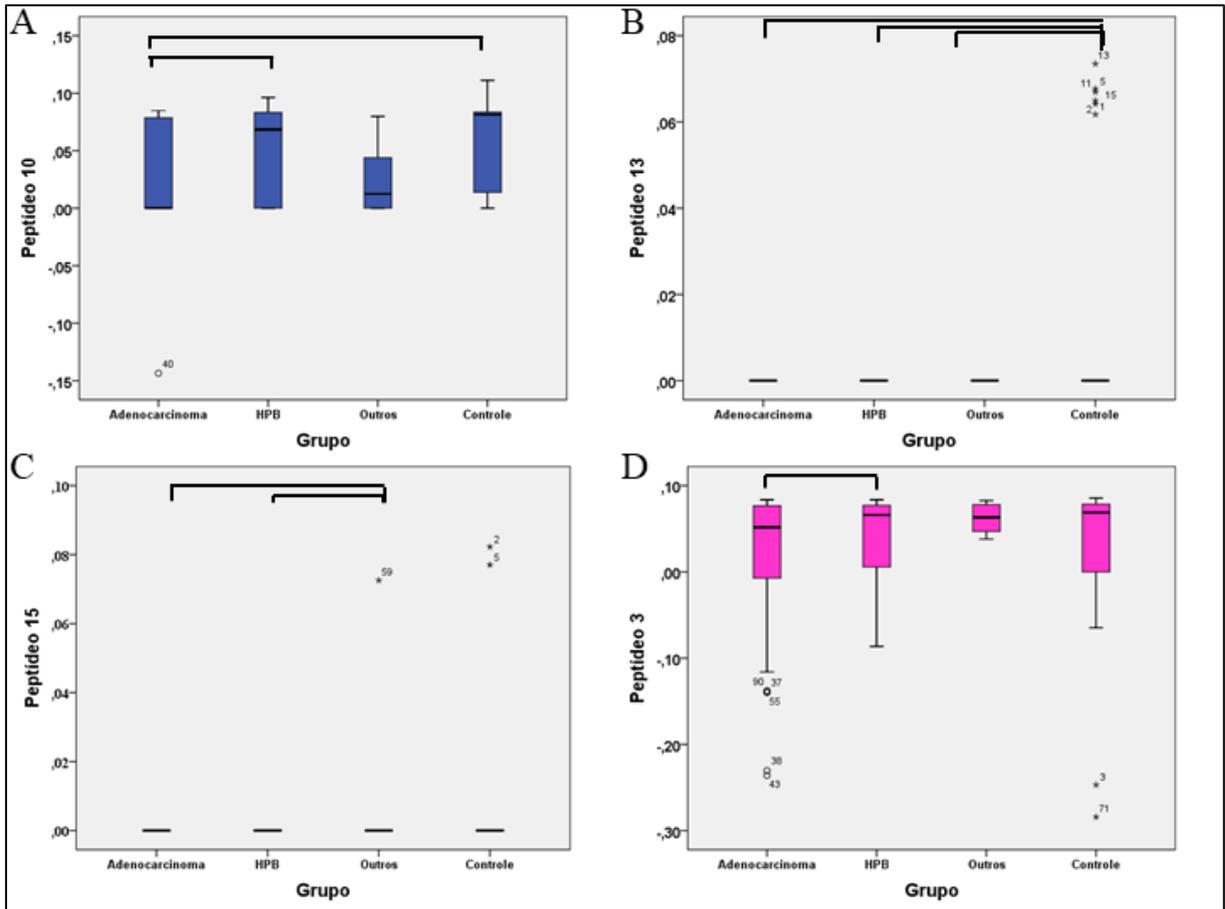


Figura 6: Boxplot dos valores de mediana para os peptídeos diferencialmente expressos. Gráfico Boxplot dos valores de \log_2 dos peptídeos com valor de $p < 0,05$, sendo destacados pelas barras comparativas entre os grupos quando diferencialmente expressos. A: Peptídeo 10, B: Peptídeo 13, C: Peptídeo 15 e D: Peptídeo 3.

Curvas ROC (do inglês, *Receiver Operating Characteristic*) foram calculadas para cada peptídeo identificado como diferencialmente representado entre os grupos utilizando os valores de intensidade relativa normalizada por \log_2 ou mediana para cada um destes peptídeos em cada amostra. O resultado dessa avaliação preliminar do valor diagnóstico de cada peptídeo está apresentado na Figura 7 e os resultados obtidos estão sumarizados na Tabela 6.

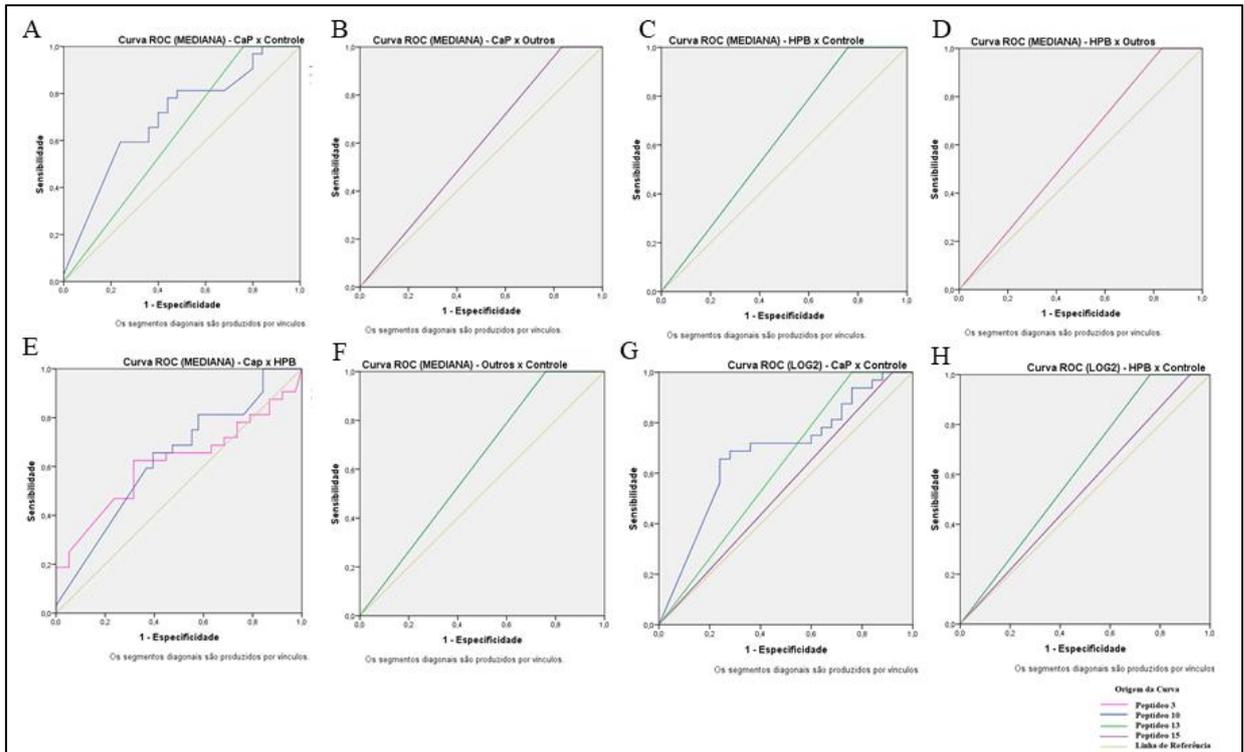


Figura 7: Curvas ROC com peptídeos diferencialmente expressos. Análise de curva ROC com os peptídeos diferencialmente expressos ($p < 0,05$) em cada grupo de estudo. De A a F, valores normalizados pela mediana, G e H, valores normalizados por log2. A e G: CaP x Controle; B: CaP x Outros; C e H: HPB x Controle; D: HPB x Outros; E: CaP x HPB; F: Outros x Controle.

A curva ROC é um gráfico bidimensional no qual o eixo x é definido pela especificidade e o eixo y pela sensibilidade. Portanto, a curva ROC é muito usada para analisar a capacidade de determinado biomarcador identificar corretamente uma característica determinada por um grupo. A área sob a curva (AUC) é o valor determinante da força de predição do marcador, por avaliar a discordância entre as amostras (HSU; CHANG; HSUEH, 2014). Os peptídeos urinários mostraram valores de AUC de 0,540 a 0,697 (Tabela 6), com o peptídeo 10 resultando no melhor valor diagnóstico na comparação CaP x controles. Além dos valores demonstrados na comparação CaP x controle, os peptídeos 13 e 15 mostraram também potencial para a diferenciação dos grupos HPB e controle. Os peptídeos 3 e 10 mostraram potencial valor diagnóstico para a diferenciação entre CaP e HPB. Como esperado, o peptídeo 15 apresentou, no geral, pior performance diagnóstica (menores valores de AUC) entre os peptídeos avaliados, uma vez que ele foi identificado em um número bastante limitado de amostras. Contudo, apenas o peptídeo 15 mostrou algum potencial diagnóstico para a diferenciação entre CaP x outros e HPB x outros. Embora esse peptídeo tenha sido identificado

em apenas uma amostra do grupo outros (grupo 3), esse grupo é composto por apenas 6 amostras. Peptídeos capazes de diferenciar condições pré-malignas, como PIN, ou condições diagnósticas duvidosas, como ASAP, são importantes ferramentas diagnósticas. Portanto, esse peptídeo não deve ser descartado de futuras análises.

Embora o número amostras utilizado na nossa análise seja baixo, o que afeta negativamente a análise das curvas ROC e, conseqüentemente, o valor diagnóstico dos peptídeos, os resultados obtidos são comparáveis a outros biomarcadores urinários descritos na literatura, que podem possuir valores de AUC próximos a 0,61 a 0,75, e valores de sensibilidade e especificidade bastante variadas (FUJITA; NONOMURA, 2018; WANG et al., 2017). Alguns biomarcadores propostos pela literatura são identificados apenas após protocolo de massagem na próstata, mas tal manipulação pode ser empecilho para obtenção da amostra e, por isso, esta metodologia não foi usada no trabalho (FUJITA et al., 2009a, 2009b).

Além disso, construímos curvas ROC com os valores de PSA total informados pelos pacientes para as diferentes comparações entre os grupos (Figura 8). A curva ROC do PSA pode ser executada utilizando a razão entre os valores de PSA livre e PSA total (EROL et al., 2014), mas o primeiro raramente foi informado pelos pacientes, o que inviabilizou o cálculo para esta curva. As curvas ROC calculadas a partir dos valores de PSA total para os grupos CaP x controles e HPB x controles geraram altos valores de AUC (0,998 e 0,994, respectivamente). Porém, é importante ressaltar que parte da seleção das amostras foi baseada nos valores de PSA. Além disso, as amostras foram coletadas no momento prévio à biópsia de próstata, sendo este procedimento solicitado a partir do aumento do valor de PSA. Portanto os dados utilizados para a realização da curva ROC contêm um viés importante. Análises com curva ROC com valores de PSA, utilizando diferente número amostral apresentaram AUC de 0,68-0,89, mas que dependiam do nível de avanço da doença (THOMPSON et al., 2005; WEI et al., 2016). De qualquer forma, ao observar a curva ROC para os valores de PSA total para os grupos CaP e HPB (Figura 8C), percebe-se que os valores de PSA não são suficientes para diferenciar a condição maligna da benigna, ressaltando a necessidade do desenvolvimento de um teste marcador mais específico. Na comparação CaP x HPB, os peptídeos 3 e 10 mostraram valores de AUC superiores aos apresentados pelo PSA total, demonstrando o potencial diagnóstico dos peptídeos urinários.

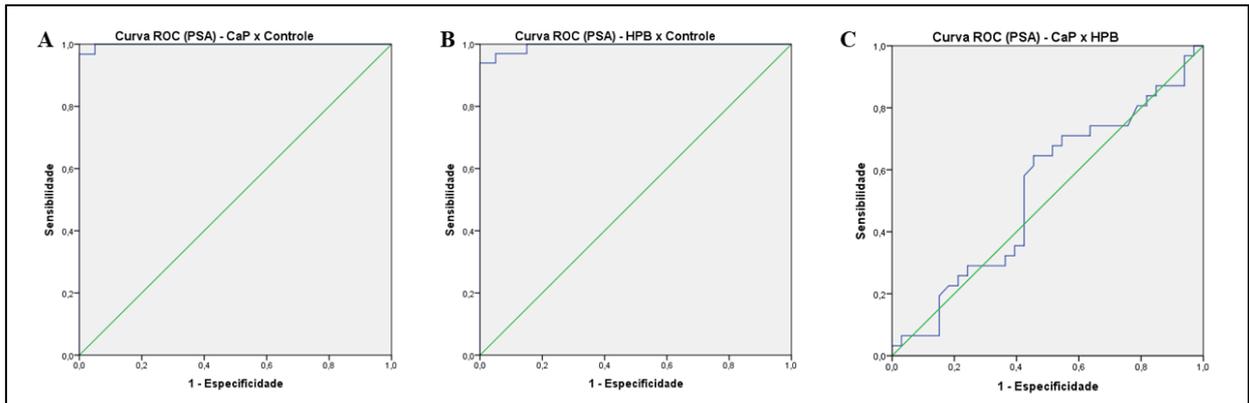


Figura 8: Curvas ROC usando os valores de PSA. As amostras utilizadas para gerar este gráfico foram as mesmas utilizadas para gerar os gráficos anteriores. Sendo A: comparação entre os grupos CaP e Controle; B: comparação entre os grupos HPB e Controle; e C: comparação entre os grupos CaP e HPB.

Tabela 6: Resumo dos dados obtidos pelas Curvas ROC. Dispõem dos cálculos de área sob a curva, sensibilidade máxima e especificidade máxima.

Resultados - Curva ROC					
Grupos Comparados	Peptídeo Analisado	Dados Utilizados	Área Sob a Curva (AUC)	Sensibilidade Máxima	Especificidade Máxima
CaP e Controle	Peptídeo 13	Normalização LOG2	0,62	100%	24%
		Normalização Mediana	0,62	100%	24%
	Peptídeo 10	Normalização LOG2	0,684	66%	76%
		Normalização Mediana	0,697	59%	76%
	Peptídeo 15	Normalização LOG2	0,54	100%	8%
CaP e HPB	Peptídeo 10	Normalização Mediana	0,632	66%	61%
	Peptídeo 3	Normalização Mediana	0,614	63%	68%
CaP e Outros	Peptídeo 15	Normalização Mediana	0,583	100%	17%
HPB e Controle	Peptídeo 13	Normalização LOG2	0,62	100%	24%

		Normalização Mediana	0,62	100%	24%
	Peptídeo 15	Normalização LOG2	0,54	100%	8%
HPB e Outros	Peptídeo 15	Normalização Mediana	0,583	100%	17%
Outros e Controle	Peptídeo 13	Normalização Mediana	0,62	100%	24%
CaP e Controle	<i>(Não Aplicável)</i>	Valores de PSA	0,998	97%	100%
HPB e Controle	<i>(Não Aplicável)</i>	Valores de PSA	0,994	94%	100%
CaP e HPB	<i>(Não Aplicável)</i>	Valores de PSA	0,53	65%	55%

Apesar das limitações do estudo, peptídeos com potencial para uso no diagnóstico do CaP foram identificados. Avaliações devem ser feitas utilizando maior número de amostras e metodologias mais sensíveis e específicas (target MS) para determinar o valor diagnóstico dos peptídeos identificados. A fim de aumentar a força diagnóstica dos biomarcadores, a análise em conjunto é mais eficiente em diversos casos (GILQUIN et al., 2017; HERMANN et al., 2017), e por isso, é importante considerar o uso associado dos potenciais peptídeos biomarcadores encontrados neste trabalho para melhor diagnosticar as patologias prostáticas.

6. Conclusão e Perspectiva

Neste trabalho foi possível estabelecer um fluxo metodológico para detecção e análise de peptídeos circulantes na urina, assim como discorrer sobre possíveis erros associados a determinados processos. Encontramos quatro peptídeos com diferentes potenciais para biomarcador das condições patológicas estudadas (Câncer de Próstata, Hiperplasia Benigna da Próstata e condições “pré-câncer”). Sugerimos que um estudo com um número amostral maior seja executado para confirmar a eficácia dos peptídeos apresentados neste trabalho usando técnicas mais sensíveis de detecção, além de um estudo mais aprofundado comparando os subgrupos definidos para cada condição.

7. Referências

ARONSON, J. K.; FERNER, R. E. Biomarkers-A General Review. **Current protocols in pharmacology**, v. 76, p. 9.23.1-9.23.17, 17 mar. 2017.

BARBOSA, J. M. G. et al. Cerumenogram: a new frontier in cancer diagnosis in humans. **Scientific reports**, v. 9, n. 1, p. 11722, ago. 2019.

Biópsia para Diagnóstico do Câncer de Próstata - Instituto Oncoguia. Disponível em: <<http://www.oncoguia.org.br/conteudo/biopsia-para-diagnostico-do-cancer-de--prostata/1201/289/>>. Acesso em: 23 out. 2019.

BRIZMOHUN APPAYYA, M. et al. National implementation of multi-parametric magnetic resonance imaging for prostate cancer detection – recommendations from a UK consensus meeting. **BJU International**, v. 122, n. 1, p. 13–25, 1 jul. 2018.

BRÖKER, M. E. E. et al. Collagen Peptides in Urine: A New Promising Biomarker for the Detection of Colorectal Liver Metastases. **PLOS ONE**, v. 8, n. 8, p. 1–7, 2013.

BROWN, A. M. et al. Recent advances in image-guided targeted prostate biopsy. **Abdominal Imaging**, v. 40, n. 6, p. 1788–1799, 12 ago. 2015.

BROZA, Y. Y. et al. Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. **Chemical Reviews**, p. acs.chemrev.9b00437, 2019.

BRUNER, D. W. et al. Relative risk of prostate cancer for men with affected relatives: Systematic review and meta-analysis. **International Journal of Cancer**, v. 107, n. 5, p. 797–803, 10 dez. 2003.

CAZARES, L. H. et al. Normal, benign, preneoplastic, and malignant prostate cells have distinct protein expression profiles resolved by Surface Enhanced Laser Desorption/Ionization mass spectrometry. **Clinical Cancer Research**, v. 8, n. 8, p. 2541–2552, 2002.

DAVALIEVA, K. et al. Proteomics analysis of urine reveals acute phase response proteins as candidate diagnostic biomarkers for prostate cancer. **Proteome Science**, v. 13, n. 1, 29 jan. 2015.

ENSOR, J. E. Outcomes Research Biomarker Validation: Common Data Analysis Concerns. **The Oncologist**, v. 61, p. 886–891, 2014.

EROL, B. et al. The cutoff level of free/total prostate specific antigen (f/t PSA) ratios in the

diagnosis of prostate cancer: A validation study on a Turkish patient population in different age categories. **Kaohsiung Journal of Medical Sciences**, v. 30, n. 11, p. 545–550, 1 nov. 2014.

FARASHI, S. et al. Post-GWAS in prostate cancer: from genetic association to biological contribution. **Nature Reviews Cancer**, v. 19, n. 1, p. 46–59, 1 jan. 2019.

FIUK, J. V. et al. Antibiotics prophylaxis before prostate biopsy in practice: Review of online clinical guidelines. **Urology Annals**, v. 7, n. 2, p. 279–280, 1 abr. 2015.

FRASÃO, G.; SAÚDE, /MINISTÉRIO DA. **Câncer de próstata: causas, sintomas, tratamentos e prevenção**. Disponível em: <<http://www.saude.gov.br/saude-de-a-z/cancer-de-prostata#tratamento>>. Acesso em: 20 out. 2019.

FREIDLIN, B.; KORN, E. L. Biomarker enrichment strategies: matching trial design to biomarker credentials. **Nature Publishing Group**, v. 11, p. 81–90, 2013.

FUJITA, K. et al. Endoglin (CD105) as a urinary and serum marker of prostate cancer. **International Journal of Cancer**, v. 124, n. 3, p. 664–669, 1 fev. 2009a.

FUJITA, K. et al. Specific detection of prostate cancer cells in urine by multiplex immunofluorescence cytology. **Human Pathology**, v. 40, n. 7, p. 924–933, jul. 2009b.

FUJITA, K.; NONOMURA, N. Urinary biomarkers of prostate cancer. **International Journal of Urology**, v. 25, n. 9, p. 770–779, set. 2018.

GANDOLFO, L. C.; SPEED, T. P. RLE plots: Visualizing unwanted variation in high dimensional data. **PLOS ONE**, v. 13, n. 2, p. e0191629, 5 fev. 2018.

GILQUIN, B. et al. Multiplex and accurate quantification of acute kidney injury biomarker candidates in urine using Protein Standard Absolute Quantification (PSAQ) and targeted proteomics. **Talanta**, v. 164, p. 77–84, 1 mar. 2017.

GLOBAL BURDEN OF DISEASE CANCER COLLABORATION. **Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-Years for 29 Cancer Groups, 1990 to 2017: A Systematic Analysis for the Global Burden of Disease Study**. Disponível em: <<https://doi.org/10.1001/jamaoncol.2019.2996>>.

GOOD, D. M. et al. Naturally occurring human urinary peptides for use in diagnosis of

chronic kidney disease. **Molecular and Cellular Proteomics**, v. 9, n. 11, p. 2424–2437, 2010.

HAJ-AHMAD, T. A.; ABDALLA, M. A. K.; HAJ-AHMAD, Y. Potential urinary protein biomarker candidates for the accurate detection of prostate cancer among benign prostatic hyperplasia patients. **Journal of Cancer**, v. 5, n. 2, p. 103–114, 2014.

HARPOLE, M.; DAVIS, J.; ESPINA, V. Current state of the art for enhancing urine biomarker discovery. **Expert Review of Proteomics**, v. 13, n. 6, p. 609–626, 2 jun. 2016.

HEIJNSDIJK, E. A. M. et al. Cost-effectiveness of Prostate Cancer Screening: A Simulation Study Based on ERSPC Data. **JNCI: Journal of the National Cancer Institute**, v. 107, n. 1, jan. 2015.

HERMANN, N. et al. Diagnostic relevance of a novel multiplex immunoassay panel in breast cancer. **Tumor Biology**, v. 39, n. 6, p. 101042831771138, 15 jun. 2017.

HSU, M. J.; CHANG, Y. C. I.; HSUEH, H. M. Biomarker selection for medical diagnosis using the partial area under the ROC curve. **BMC Research Notes**, v. 7, n. 1, 10 jan. 2014.

HUMAN PROTEOME ORGANIZATION. **Standard Protocol for Urine Collection and Storage**. Disponível em: <[http://www.hkupp.org/Urine collection Documents.htm](http://www.hkupp.org/Urine%20collection%20Documents.htm)>. Acesso em: 17 nov. 2019.

JACOB, L. et al. Impact of tobacco smoking on the risk of developing 25 different cancers in the UK: A retrospective study of 422,010 patients followed for up to 30 years. **Oncotarget**, v. 9, n. 25, p. 17420–17429, 1 abr. 2018.

JEDINAK, A. et al. Novel non-invasive biomarkers that distinguish between benign prostate hyperplasia and prostate cancer. **BMC Cancer**, v. 15, n. 1, p. 259, 11 dez. 2015.

KOLCH, W. et al. Capillary electrophoresis-mass spectrometry as a powerful tool in clinical diagnosis and biomarker discovery. **Mass Spectrometry Reviews**, v. 24, n. 6, p. 959–977, nov. 2005.

LANDER, E. S.; SCHORK, N. J. Genetic dissection of complex traits. **Science**, v. 265, n. 5181, p. 2037–2048, 1994.

LICHTENSTEIN, P. et al. Environmental and heritable factors in the causation of cancer: Analyses of cohorts of twins from Sweden, Denmark, and Finland. **New England Journal of**

Medicine, v. 343, n. 2, p. 78–85, 13 jul. 2000.

Ligand Binding Assays – Certara. Disponível em:

<<https://www.certara.com/2014/07/22/ligand-binding-assays/>>. Acesso em: 5 nov. 2019.

LILJA, H.; ULMERT, D.; VICKERS, A. J. Prostate-specific antigen and prostate cancer: Prediction, detection and monitoring. **Nature Reviews Cancer**, v. 8, n. 4, p. 268–278, abr. 2008.

LITWIN, M. S.; TAN, H. J. The diagnosis and treatment of prostate cancer: A review. **JAMA - Journal of the American Medical Association**, v. 317, n. 24, p. 2532–2542, 27 jun. 2017.

LOPES, P. M. et al. O papel da ecografia transretal no diagnóstico do câncer da próstata: Novas contribuições. **Radiologia Brasileira**, v. 48, n. 1, p. 7–11, 2015.

MAGALHÃES, P. et al. Comparison of Urine and Plasma Peptidome Indicates Selectivity in Renal Peptide Handling. **PROTEOMICS - Clinical Applications**, v. 12, n. 5, p. 1700163, set. 2018.

MOREIRA, N. M. **O preconceito em relação ao exame de toque retal como forma de rastreamento do câncer de próstata**. [s.l.] Universidade Federal de Minas Gerais. Faculdade de Medicina. Núcleo de Educação em Saúde Coletiva . Conselheiro Lafaiete, 2012.

PARKER, B. L. et al. Multiplexed Temporal Quantification of the Exercise-regulated Plasma Peptidome. **Molecular & cellular proteomics : MCP**, v. 16, n. 12, p. 2055–2068, dez. 2017.

POLLEY, M.-Y. C. Power estimation in biomarker studies where events are already observed. **Clinical Trials**, v. 14, n. 6, p. 621–628, 4 dez. 2017.

SANTOS, A. C. **Avaliação do teste para o gene 3 do câncer da próstata (PCA3) como estratégia para melhorar a especificidade das biópsias em homens com suspeita deste tumor**. [s.l.] Hospital do Câncer de Barretos. Fundação Pio XII (FP). Barretos , SP, Brasil, 2015.

SCHRADER, M.; SCHULZ-KNAPPE, P.; FRICKER, L. D. Historical perspective of peptidomics. **EuPA Open Proteomics**, v. 3, p. 171–182, 2014.

SEKAR, R.; CHOW, B. K. C. Metabolic effects of secretin. **General and Comparative Endocrinology**, v. 181, n. 1, p. 18–24, 5 jan. 2013.

SHEN, M. M.; ABATE-SHEN, C. Molecular genetics of prostate cancer: new prospects for

- old challenges. **Genes & development**, v. 24, n. 18, p. 1967–2000, 15 set. 2010.
- SIMON, R. M.; PAIK, S.; HAYES, D. F. Use of Archived Specimens in Evaluation of Prognostic and Predictive Biomarkers. **JNCI: Journal of the National Cancer Institute**, v. 101, n. 21, p. 1446–1452, 2009.
- SIVARAMAN, A. et al. Evolution of prostate biopsy techniques. Looking back on a meaningful journey. **Actas Urologicas Espanolas**, v. 40, n. 8, p. 492–498, 1 out. 2016.
- SIWY, J. et al. Human urinary peptide database for multiple disease biomarker discovery. **PROTEOMICS - Clinical Applications**, v. 5, n. 5–6, p. 367–374, jun. 2011.
- TABAYOYONG, W.; ABOUASSALY, R. Prostate Cancer Screening and the Associated Controversy. **Surgical Clinics of North America**, 2015.
- THEODORESCU, D. et al. Discovery and validation of urinary biomarkers for prostate cancer. **PROTEOMICS - Clinical Applications**, v. 2, n. 4, p. 556–570, 2008.
- THOMPSON, I. M. et al. Operating characteristics of prostate-specific antigen in men with an initial PSA level of 3.0 ng/mL or lower. **Journal of the American Medical Association**, v. 294, n. 1, p. 66–70, jul. 2005.
- VERZE, P.; CAI, T.; LORENZETTI, S. The role of the prostate in male fertility, health and disease. **Nature reviews. Urology**, v. 13, n. 7, p. 379–86, 2016.
- WANG, T. et al. Diagnostic significance of urinary long non-coding PCA3 RNA in prostate cancer. **Oncotarget**, v. 8, n. 35, p. 58577–58586, 2017.
- WEI, J. T. et al. Can urinary PCA3 Supplement PSA in the early detection of prostate cancer? **Journal of Clinical Oncology**, v. 32, n. 36, p. 4066–4072, 20 dez. 2014.
- WEI, R. J. et al. Serum levels of PSA, ALP, ICTP, and BSP in prostate cancer patients and the significance of ROC curve in the diagnosis of prostate cancer bone metastases. **Genetics and Molecular Research**, v. 15, n. 2, 3 jun. 2016.
- WICKELMAIER, F.; WICKELMAIER, F. An introduction to MDS. 2003.
- WULFKUHLE, J. D.; LIOTTA, L. A.; PETRICOIN, E. F. Proteomic applications for the early detection of cancer. **Nature Reviews Cancer**, v. 3, n. 4, p. 267–275, 2003.
- YOU DEN, W. J. Index for rating diagnostic tests. **Cancer**, v. 3, n. 1, p. 32–35, 1950.

YOUHANNA, S. et al. Determination of uromodulin in human urine: influence of storage and processing. **Nephrology Dialysis Transplantation**, v. 29, n. 1, p. 136–145, 2013.

Anexo 1

```

#Função ArqList gera uma lista das tabelas a serem usados na análise
def ArqList (path):
    temp = os.listdir (path)
    dirs = []
    for x in temp:
        temp2 = os.listdir (path + '/' + x)
        for y in temp2:
            if (len(y.split("_final_"))>1:
                if y.split("_final_")[1] == "peptide.csv":
                    dirs = dirs + [path + '/' + x + '/' + y]
    return (dirs)

#Função Group_Table recebe a matriz resultante de ArqList e gera a primeira tabela e
a lista de grupos
def Group_Table (dirs,Grupos):
    res = []
    for x in range(len(dirs)):
        for y in dirs[x]:
            temp = y.split('/')
            if (len(temp)) > 2:
                temp2 = Sample_rename(Grupos[x],temp[-1])
                res = res + [temp2+'/'+'t'+str(x+1)]
    return (res)

#Função Sample_rename recebe o nome do grupo e renomeia das amostras com base no nome
da tabela
def Sample_rename (group,sample):
    if sample[0]!='_':
        temp = sample.split('_')
        temp2 = group+'_'+temp[0]+'_'+temp[3]
    else:
        temp2 = 'ERROR'
    return (temp2)

#Função read_file gera uma lista com as linhas contidas no arquivo filename
def read_file (filename):
    with open (filename) as file:
        return file.read().split('\n')[0:]

#Função AutoCurate_Filter seleciona apenas as informações de peptídeos com valor de
AutoCurate 'Green'
def AutoCurate_Filter (table,sample):
    res = []
    for x in table:
        if x!='':
            temp = x.split(',')
            if temp[17] == 'Green':
                res = res +[temp[25]+'_'+temp[41]+'_'+temp[51]+'_'+sample]
    return(res)

```

```
#Função Tab2_oldPep adiciona informações de um peptídeo já encontrado em outra amostra
```

```
def Tab2_oldPep(table2,info):
    #print 'Old'
    pep = info.split(',')[0]
    rt = info.split(',')[1]
    value = info.split(',')[2]
    sample = info.split(',')[3]
    temp = len(table2[0])
    res = []
    tempcont = 0 #coluna da amostra
    temppep = 0 #linha do peptideo
    for x in range(len(table2)):
        if table2[x][0] == pep:
            temppep = x
    for x in range(len(table2)):
        if x == 0:
            if sample in table2[0]:
                for y in range(len(table2[0])):
                    if table2[0][y] == sample:
                        tempcont = y
                res = [table2[0]]
            else:
                res = [table2[0]+[sample]]
        else:
            if tempcont == 0:
                temp2 = table2[x]+'0'
                res = res + [temp2]
            else:
                temp2 = table2[x]
                res = res + [temp2]
    if tempcont == 0:
        res[temppep][-1] = value
    else:
        res[temppep][tempcont] = value
    return (res)
```

```
#Função Tab2_newPep adiciona informações de um peptídeo não encontrado em outra amostra
```

```
def Tab2_newPep(table2,info):
    #print 'New'
    pep = info.split(',')[0]
    rt = info.split(',')[1]
    value = info.split(',')[2]
    sample = info.split(',')[3]
    temp = len(table2[0])
    res = []
    tempcont = 0
    if len(table2) == 1:
        res = table2[0]
        res = res+[sample]###
        res = [res] + [[pep,rt,value]]
    else:
        for x in range(len(table2)):
            if x == 0:
                if sample in table2[0]:
                    for y in range(len(table2[0])):
                        if table2[0][y] == sample:
                            tempcont = y
```

```

res = [table2[0]]
    else:
        res = [table2[0]+[sample]]
else:
    if tempcont == 0:
        temp2 = table2[x]+'0'
        res = res + [temp2]
        temp3 = [pep,rt]
        for y in (range(len(table2[0])-2)):
            if y < (len(table[0])-3):
                temp3 = temp3 + ['0']
            temp3 = temp3 + [value]
            if x == ((len(table2))-1):#
                res = res + [temp3]#
    else:
        temp2 = table2[x]
        res = res + [temp2]
        temp3 = [pep,rt]
        for y in (range(len(table2[0])-2)):
            temp3 = temp3 + ['0']
            temp3[tempcont] = value
            if x == ((len(table2))-1):
                res = res + [temp3]

#print 'res final'
#print res
return (res)

#Gera Tabela2 com as médias dos valores das amostras, se uma tiver valor igual a
zero, mantém o outro valor
def ValueMed (table):
    temp = table
    return (table)

Grupos = ['grupo1','grupo2','grupo3']
tables = []
for x in Grupos:
    temp = ArqList ('pasta'+x)
    tables = tables +[temp]
table1 = Group_Table (tables,Grupos)
table1 = ['Sample/tGroup']+table1

Tabl=open('D:/GroupsTable.txt','w')
for x in table1:
    Tabl.write('%s\n'%(x))
Tabl.close()

resultlist = []
for x in range(len(tables)):
    for y in tables[x]:
        table = read_file(y)
        temp = y.split('/')
        if (len(temp)) > 2:
            temp2 = Sample_rename(Grupos[x],temp[-1])
            if temp2 != 'ERROR':
                temp3 = AutoCurate_Filter(table,temp2)
                if temp3!='':
                    resultlist=resultlist+[temp3]

pep = []
table2 = [['Peptide','RT']]

```

```

for x in resultlist:
    for y in x:
        if y.split(',')[0] in pep:
            temp = Tab2_oldPep(table2,y)
        else:
            pep = pep + [y.split(',')[0]]
            temp = Tab2_newPep(table2,y)
        table2 = temp
        #print table2

Tab2=open('D:/TablePep.txt','w')
for x in table2:
    for y in x:
        Tab2.write('%s\t'%(x))
    Tab2.write('\n')
Tab2.close()

Med = read_file('D:/TableSamples.txt')
for x in range(len(Med)):
    temp = Med[x].split('\t')
    if temp[0] == 'x':
        Med[x] = [Med[x]] + [0]
    else:
        for y in range(len(table2[0])):
            if table2[0][y] == temp[0]:
                Med[x] = [Med[x]] + [y]
    if temp[1] == 'x':
        Med[x] = Med[x] + [0]
    else:
        for y in range(len(table2[0])):
            if table2[0][y] == temp[1]:
                Med[x] = Med[x] + [y]

#Tab3=open('D:/TablePep2.txt','w')
#for x in Med:
#    Tab3.write('%s\t'%(x))
#Tab3.close()

#print Med

#res_med

#for x in range(len(table2)):
#    if x == 0:
#        temp = []
#        for y in range(len(Med)):

```