

MINISTÉRIO DA SAÚDE
FUNDAÇÃO OSWALDO CRUZ
INSTITUTO OSWALDO CRUZ

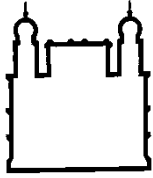
Doutorado em Biologia Parasitária

CARACTERIZAÇÃO DE SERINA PEPTIDASES E IDENTIFICAÇÃO DO
PERFIL PROTEICO EM INTESTINO DE *Culex quinquefasciatus*,
Anopheles albitarsis s.s e *Aedes aegypti*
(DIPTERA: CULICIDAE)

ANDRÉ BORGES VELOSO

Rio de Janeiro

Agosto de 2015



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Parasitária

André Borges Veloso

Caracterização de serina peptidases e identificação do perfil proteico em intestino de *Culex quinquefasciatus*, *Anopheles albitarsis* s.s e *Aedes aegypti* (Diptera: Culicidae)

Tese de doutorado apresentada à coordenação do Programa de Biologia Parasitária do Instituto Oswaldo Cruz – Fiocruz, como parte dos requisitos necessários à obtenção do título de Doutor em Ciências, área de concentração em Bioquímica e Genética. A tese foi desenvolvida no Laboratório de Biologia Molecular e Doenças Endêmicas do Instituto Oswaldo Cruz – Fiocruz, com suporte financeiro do CNPQ.

Orientador: Prof. Dr. José Batista de Jesus

RIO DE JANEIRO

Agosto de 2015

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

V443 Veloso, André Borges

Caracterização de serina peptidases e identificação do perfil proteico em intestino de *Culex quinquefasciatus*, *Anopheles albitarsis s.s* e *Aedes aegypti* (Diptera: Culicidae) / André Borges Veloso. – Rio de Janeiro, 2015.

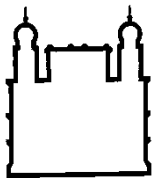
xxi, 183 f. : il. ; 30 cm.

Tese (Doutorado) – Instituto Oswaldo Cruz, Pós-Graduação em Biologia Parasitária, 2015.

Bibliografia: f. 99-118

1. Serina peptidases. 2. Tripsina. 3. Culicidae. 4. *Aedes aegypti*. 5. *Anopheles albitarsis*. 6. *Culex quinquefasciatus*. 7. LC-MS/MS. I. Título.

CDD 572.65



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Programa de Pós-Graduação em Biologia Parasitária

AUTOR: ANDRÉ BORGES VELOSO

Caracterização de serina peptidases e identificação do perfil proteico em intestino de *Culex quinquefasciatus*, *Anopheles albitarsis* s.s e *Aedes aegypti* (Diptera: Culicidae)

ORIENTADOR: Prof. Dr. José Batista de Jesus

Aprovado em: ____/____/____

EXAMINADORES:

Prof. Dr. Reginaldo Brazil – Presidente (Fundação Oswaldo Cruz/RJ)

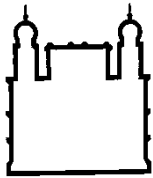
Prof. Dr. Gabriel Padrón (Centro de Engenharia Genética e Biotecnologia/Cuba)

Prof. Dr. Adeilton Brandão (Fundação Oswaldo Cruz/RJ)

Prof. Dra. Daniele Pereira de Castro (Fundação Oswaldo Cruz/RJ)

Prof. Dr. André Luís Souza dos Santos (Universidade Federal do Rio de Janeiro/RJ)

Rio de Janeiro, 28 de agosto de 2015



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Anexar a cópia da Ata que será entregue pela SEAC já assinada.

AGRADECIMENTOS

Ao meu orientador, Dr. José Batista de Jesus, pela orientação e amizade. Por ter me aceito em seu grupo de pesquisa e proporcionado as condições para o desenvolvimento do trabalho.

À Dra. Patrícia Cuervo Escobar, por ter sido muito mais que co-orientadora, esteve sempre disponível para discutir sobre os experimentos e também para refletirmos sobre a vida.

À Dra. Constança Britto, por me receber em seu laboratório de maneira matriarcal. Obrigado pelo apoio e pelo doce convívio.

À Dra. Cláudia M. D'Ávila Levi, por estar sempre disponível para conversar e discutir assuntos referentes ao trabalho, ou não.

Ao Dr. José Bento e Dra. Denise Valle, do Laboratório de Fisiologia e Controle de Artrópodes Vetores-LAFICAVE, pelo fornecimento dos mosquitos utilizados para o desenvolvimento da tese.

Aos responsáveis pela e manutenção do insetário - LAFICAVE, que foram sempre pacientes e comprometidos com o fornecimento dos insetos nas condições necessárias para o desenvolvimento da tese.

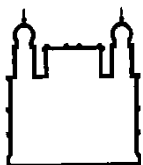
Aos camaradas de trabalho Geovane, Léo e Camilinha. Foram importantes para o desenvolvimento dos experimentos da maneira correta e divertida.

Aos companheiros do LABIMDOE que tornaram as rotinas laboratoriais mais humanas, em especial Otacílio, Bernardo, Franklin, Myllena, Ícaro, Natália, Raquel e Thaís e Vítor.

Às agências de fomento e fontes de financiamento, cujo apoio permitiu a realização desta pesquisa FAPEMIG (J.B.J Edital Universal Processo No. APQ-01070-12), CNPq (J.B.J. PQ Processo No. 308679/2012-1), FAPERJ (C.B. Processo E-26/110.594/2012 e CNE E-26/102.775/2012), FIOCRUZ-IOC e CAPES.

À minha família, que fomentaram irrestritamente os meus sonhos. Não existe em minha memória resquícios de momentos em que falharam para comigo.

À minha mulher, Renata de Cássia Pires, por ser uma pessoa especial em minha vida.
Obrigado por me apoiar e ser uma valente companheira.



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Caracterização de serina peptidases e identificação do perfil proteico em intestino de *Culex quinquefasciatus*, *Anopheles albitarsis* s.s e *Aedes aegypti* (Diptera: Culicidae)

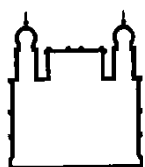
André Borges Veloso

RESUMO

Os dípteros *Aedes aegypti*, *Anopheles albitarsis* e *Culex quinquefasciatus* pertencem à família Culicidae e são considerados um problema de saúde pública quando atuam como vetores de parasitos e/ou arboviroses. O regime de alimentação de fêmeas dos mosquitos anautógenos influencia os processos que ocorrem no intestino e regula os genes envolvidos na reprodução. Além disso, apesar dos ciclos de vida dos parasitos veiculados por essas espécies de vetores serem distintos, todos eles são ingeridos durante a hematofagia e expostos ao ambiente do intestino médio (estômago) para em seguida atravessar o epitélio intestinal e chegarem ao tecido apropriado para seu desenvolvimento e/ou transmissão para um novo hospedeiro vertebrado. Neste trabalho, utilizamos uma abordagem proteômica para a identificação de proteínas totais, bem como ensaios em solução e enzimografia em gel copolimerizado com substrato para identificação de peptidases ativas presentes no intestino de fêmeas de *Cx. quinquefasciatus* e *An. albitarsis* s.s sob regime de alimentação com açúcar. As técnicas de enzimografia também foram utilizadas aqui para a caracterização de peptidases ativas de intestinos de fêmeas de *Ae. aegypti* alimentadas com açúcar ou sangue. Foi observado que o intestino de fêmeas das três espécies alimentadas com açúcar apresenta um complexo perfil de peptidases ativas, do tipo tripsina, composto por bandas migrando nas regiões entre ~ 24 a 40 kDa e em regiões de alto peso molecular. Atividades proteolíticas foram detectadas entre pH 3,5 – 10, revelando diferenças quantitativas e qualitativas entre os perfis proteolíticos das três espécies. Em fêmeas de *Ae. aegypti*, o regime de alimentação com açúcar ou sangue alterou o perfil de SPtrip ativas do intestino de forma qualitativa e quantitativa. Um total de oito

tripsinas em *Cx. quinquefasciatus* e dez tripsinas em *An. albitarsis*, foram identificadas por SDS-PAGE acoplado a LC-MS/MS. Apesar dessas SPtrip identificadas apresentarem características comuns às tripsinas digestivas de invertebrados, foram observadas diferenças nas sequências de aminoácidos, como por exemplo, em regiões de especificidade ao substrato e de ativação do zimogênio. Foi observado que os genes codificadores para SPtrip possuem diferentes tamanhos e organização, tais como, diferenças no número de éxons/introns. Com relação à composição de proteínas totais solúveis do intestino de fêmeas de *Cx. quinquefasciatus* e *An. albitarsis*, análise por LC-MS/MS e bioinformática revelou a presença de proteínas envolvidas em processos fisiológicos importantes, como exopeptidases, glicosidases, enzimas detoxificantes e proteínas relacionadas com a interação parasito/vetor. Em *Cx. quinquefasciatus*, foram identificadas 1397 proteínas, distribuídas em 1090 grupos, representando um total de ~ 7,5% das proteínas codificadas pelo genoma dessa espécie. Em *An. albitarsis* foram identificadas 916 proteínas distribuídas em 748 grupos. As centenas de proteínas presentes no intestino de fêmeas alimentadas com açúcar, de ambas as espécies, foram classificadas segundo sua ontologia gênica e o papel de algumas famílias de proteínas historicamente importante em mosquitos foram discutidas. Finalmente, foi demonstrado que abordagem proteômica combinada à enzimografia e bioinformática é uma ferramenta que pode auxiliar na anotação funcional dos genes de tripsinas expressos no intestino de fêmeas alimentadas com açúcar, bem como auxilia no mapeamento do repertório de proteínas totais presentes no intestino de *Cx. quinquefasciatus* e *An. albitarsis*.

Palavras-chave: Serina peptidases, Tripsina, Culicidae, *Aedes aegypti*, *Anopheles albitarsis* e *Culex quinquefasciatus*, LC-MS/MS



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

INSTITUTO OSWALDO CRUZ

Serine peptidase characterization and protein profiling of the midgut of *Culex quinquefasciatus*, *Anopheles albitarsis* s.s and *Aedes aegypti* (Diptera: Culicidae)

André Borges Veloso

ABSTRACT

Culex quinquefasciatus, *Anopheles albitarsis* and *Aedes aegypti* are hematophagous insect from the Culicidae family that feeds on the blood of humans, dogs, birds and livestock. These species transmit a wide variety of pathogens between humans and animals. The midgut environment is the first location of pathogen-vector interaction for blood-feeding mosquitoes and the expression of specific peptidases in the early stages of feeding could influence the outcome of the infection. Trypsin-like serine peptidases belong to a multi-gene family that can be expressed in different isoforms under distinct physiological conditions. However, the confident assignment of the trypsin genes that are expressed under each condition is still a challenge due to the large number of trypsin-coding genes in the Culicidae family and most likely because they are low abundance proteins. We used zymography for the biochemical characterization of the peptidase profile of the midgut from *Cx. quinquefasciatus*, *An. albitarsis*, and *Ae. aegypti* females fed on sugar. We also characterized the peptidases profile of the midgut from *Ae. aegypti* fed on blood, during different moments after feeding. Protein samples were also submitted to SDS-PAGE followed by liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis for peptidase identification. The peptidases sequences were analyzed by bioinformatics tools to assess their distinct features. Zymography revealed that trypsin-like serine peptidases were responsible for the proteolytic activity in the midgut of females fed on sugar or blood diet. In addition, we observed that the profile is influenced by the blood ingestion. After fractionation in

SDS-PAGE, eight and ten trypsin-like serine peptidases were identified by LC-MS/MS in *Cx. quinquefasciatus* and *An. albitarsis*, respectively. Peptidases from *Cx. quinquefasciatus* were also analysed with bioinformatic tools revealing that they have structural features typical of invertebrate digestive trypsin peptidases but exhibited singularities at the protein sequence level such as: the presence of different amino acids at the autocatalytic motif and substrate binding regions as well as different number of disulfide bounds. Data mining revealed a group of trypsin-like serine peptidases that are specific to *C. quinquefasciatus* when compared to the culicids genomes sequenced so far. We demonstrated that proteomics approaches combined with bioinformatics tools and zymographic analysis can lead to the functional annotation of trypsin-like serine peptidases coding genes and aid in the understanding of the complexity of peptidase expression in mosquitoes.

Key words: Culicidae, Trypsin, Zymography, *Aedes aegypti*, *Anopheles albitarsis*, *Culex quinquefasciatus*, LC-MS/MS

ÍNDICE

LISTA DE FIGURAS	xv
LISTA DE TABELAS	xviii
LISTA DE SIGLAS E ABREVIATURAS	xviii
LISTA DE ANEXOS	xx
1. INTRODUÇÃO	1
Mosquitos como vetores biológicos.....	1
Intestino dos mosquitos.....	5
Peptidases	8
Biossíntese e secreção das tripsinas em mosquitos	14
Proteômica.....	18
2. OBJETIVOS	26
Objetivos específicos	26
3. MATERIAIS E MÉTODOS.....	27
Insetos	27
Dissecção dos órgãos.....	27
Enzimografia	28
Extração de proteínas	28
Eletroforese.....	28
Ensaio enzimático em solução	29
Espectrometria de massas.....	30
Extração de proteínas de intestino de fêmeas para LC-MS/MS	30
Análise no espectrômetro de massas.....	31
Análise dos dados.....	32
Análises de alinhamento múltiplo e bioinformática	32

4. RESULTADOS.....	34
Atividade proteolítica em intestino de fêmeas de <i>Cx. quinquefasciatus</i> alimentadas com açúcar.....	34
Curso temporal	34
Influência do pH.....	35
Efeito dos inibidores.....	36
Atividades proteolíticas em intestinos de machos de <i>Cx. quinquefasciatus</i> alimentados com açúcar	38
Peptidases identificadas por espectrometria de massas	39
Análise de bioinformática e alinhamento das tripsinas identificadas.....	41
Identificação e classificação funcional do perfil proteico do intestino fêmeas <i>Cx. quinquefasciatus</i> alimentadas com açúcar.....	49
Grupo de proteínas com funções fisiológicas importantes.....	52
Enzimas proteolíticas	53
Inibidores de serina peptidases.....	53
Glicosidases.....	53
Detoxificação e resistência a inseticidas	54
Interação parasito/vetor.....	54
Atividade proteolítica em intestino de fêmeas de <i>An. albitarsis</i> s.s alimentadas com açúcar.....	55
Curso temporal	55
Influência do pH.....	56
Efeito dos inibidores.....	58
Atividade proteolítica em intestinos de machos de <i>An. albitarsis</i> s.s alimentados com açúcar.....	59
Peptidases identificadas por espectrometria de massas	60
Identificação e classificação funcional do perfil proteico do intestino fêmeas <i>An. albitarsis</i> s.s alimentadas com açúcar	60

Caracterização e comparação do perfil proteolítico do intestino de fêmeas de <i>Ae. aegypti</i> alimentadas com açúcar e sangue.....	64
Curso temporal de fêmeas alimentadas com açúcar.....	64
Influência do pH na atividade proteolítica de fêmeas alimentadas com açúcar	65
Efeito dos inibidores sobre a atividade proteolítica de intestino de fêmeas alimentadas com açúcar	66
Atividades proteolíticas em intestinos de machos de <i>Ae. aegypti</i>	67
Curso temporal de fêmeas de <i>Ae. aegypti</i> alimentadas com sangue	68
Efeito dos inibidores em fêmeas alimentadas com sangue	71
5. DISCUSSÃO	74
6. CONCLUSÕES	98
7. REFERÊNCIAS BIBLIOGRÁFICAS	99
8. ANEXOS	119
Anexo I.....	120
Anexo II.....	144

LISTA DE FIGURAS

- Figura 1.** Fêmea de *Culex quinquefasciatus*.
- Figura 2.** Fêmea de *Aedes aegypti*.
- Figura 3.** Fêmea de *Anopheles* sp.
- Figura 4.** Esquema ilustrativo dos órgãos internos dos mosquitos.
- Figura 5.** Esquema ilustrativo geral do intestino dos mosquitos e suas compartimentalizações.
- Figura 6.** Esquema das diferentes etapas envolvidas na ativação e atividade das tripsinas.
- Figura 7.** Rotas migratórias e locais de desenvolvimento de parasitos no intestino de mosquitos.
- Figura 8.** Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar.
- Figura 9.** Efeito do pH sobre as atividades proteolíticas de extratos de intestinos de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar.
- Figura 10.** Efeito dos inibidores de peptidases no perfil proteolítico de extrato do intestino médio de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar
- Figura 11.** Efeito dos inibidores de peptidases no perfil proteolítico de intestino médio de machos de *Cx. quinquefasciatus* alimentados com açúcar.
- Figura 12.** Perfil enzimográfico representativo e SDS-PAGE das proteínas totais solúveis de intestino fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar.
- Figura 13.** Alinhamento das tripsinas em fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar
- Figura 14.** Localização do gene da Tripsina1 de *Cx. quinquefasciatus* em seu supercontig.

Figura 15. Classificação funcional das proteínas totais de intestino de *Cx. quinquefasciatus* identificadas por espectrometria de massas, segundo os termos de ontologia gênica para componente e processo celular.

Figura 16. Classificação funcional das proteínas totais de intestino de *Cx. quinquefasciatus* identificadas por espectrometria de massas, segundo os termos de ontologia gênica para função molecular.

Figura 17. Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *An. albitarsis* alimentadas com açúcar.

Figura 18. Efeito do pH sobre as atividades proteolíticas de extratos de intestinos de fêmeas de *An. albitarsis* alimentadas com açúcar.

Figura 19. Efeito dos inibidores de peptidases no perfil proteolítico de intestino médio de fêmeas de *An. albitarsis* alimentadas com açúcar.

Figura 20. Efeito dos inibidores de peptidases sobre o perfil proteolítico de intestino de machos de *An. albitarsis* alimentados com açúcar

Figura 21. Componente celular e processos biológicos das proteínas identificadas em intestino de fêmeas de *An. albitarsis* segundo o Gene Ontology.

Figura 22. Função molecular das proteínas identificadas em intestino de fêmeas de *An. albitarsis* segundo o Gene Ontology.

Figura 23. Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar.

Figura 24. Efeito do pH sobre as atividades proteolíticas de extratos de intestinos de fêmeas de *Ae. aegypti* alimentadas com açúcar.

Figura 25. Efeito dos inibidores de peptidases no perfil proteolítico de extrato do intestino médio de fêmeas de *Ae. aegypti* alimentadas com açúcar.

Figura 26. Efeito dos inibidores de peptidases no perfil proteolítico de intestino médio de machos de *Ae. aegypti* alimentados com açúcar.

Figura 27. Perfil proteolítico de extrato de intestino de fêmeas de *Ae. aegypti* alimentadas com sangue e açúcar.

Figura 28. Efeito dos inibidores de peptidases sobre o perfil proteolítico de intestino de fêmeas de *Ae. aegypti* pós alimentados com sangue (PAS).

Figura 29. Análise da inibição da atividade proteolítica em solução de intestino de fêmeas de *Aedes aegypti* alimentadas com sangue

LISTA DE TABELAS

Tabela 1. Serina peptidases do tipo tripsina, identificadas por LC-MS/MS, em intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar.

Tabela 2. Características gerais das serina peptidases do tipo tripsina identificadas em *Cx. quinquefasciatus*.

Tabela 3. Caracterização *in silico*, das serina peptidases do tipo tripsina identificadas no intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar.

Tabela 4. Peptídeos proteotípicos das tripsinas identificadas por LC-MS/MS, preditos pelo software SKYLINE.

LISTA DE SIGLAS E ABREVIATURAS

- ATP** - Fosfato de Adenosina
- AaLT** - Enzima Tripsina Tardia
- Arg** - Arginina
- Asp** - Aspargina 189
- BApNA** - N α Benzoil-DL-Arginina-p-nitroanilida
- BOD** - Incubadora com demanda bioquímica de Oxigênio
- DNA** - Ácido desoxirribonucleico
- DMSO** - Dimetilsulfóxido
- E-64** - Trans-Epoxy succinyl-L-Leucylamido-(4-guanidino) Butano
- ET**- Tripsinas iniciais (*Early Trypsin*)
- EDTA** - Ácido etilenodiaminotetracético
- ESI** - Ionização por eletrospray (*Electrospray ionization*)
- GO** - Gene Ontology
- His 57**- Histidina 57
- IBEx** - Instituto de Biologia do Exército
- LC-MS/MS** - Cromatografia líquida acoplada a espectrometria de massas *in tandem*
- kDA** - Kilodaltons
- Lys** - Lisina
- LT** - Tripsinas tardias (*Late Trypsin*)
- MALDI** - Dessorção a laser assistida por matriz (*Matrix-assisted laser desorption/ionization*)
- mg** – Miligramas
- mM** - Milimolar
- MP** - Matriz Peritrófica
- MS** - Espectrometria de massas
- MS/MS** - Espectrometria de massa *in tandem*
- MS1** - Espectros de Massa 1
- MS2** - Espectros de Massa 2
- m/z** – Razão entre a massa e a carga de um íon
- NaCl** - Cloreto de Sódio
- Ng** – Nanogramas
- nm** - Nanômetros

- OMS** - Organização Mundial da Saúde
- Orbi-Trap** - Armadilha orbital
- PAS** - Pós Alimentado com Sangue
- PBS** - Tampão solução salina (*Phosphate buffered saline*)
- pH** - Potencial Hidrogeniônico
- pI** - Ponto isoelétrico
- PMSF** - Fluoreto de fenilmetanosulfonila (*Phenyl-methyl sulfonyl-fluoride*)
- PSM** - Encontro de espectro de peptídeos (*Peptide spectrum match*)
- proPO** – Profenoloxidase
- RNA** - Ácido ribonucleico
- RNAi** - RNA de interferência
- RNA_m** - RNA mensageiro
- Ser 195** - Serina 195
- SDS** - Dodecil sulfato de sódio (*Sodium Dodecyl Sulfate*)
- SDS-PAGE** - Eletroforese em gel de poliacrilamida (*Polyacrylamide Gel Electrophoresis*)
- SPs** - Serina Proteases
- SPTrips** - Serina Proteases do tipo Tripsina
- STI** - Inibidor de Tripsina de Soja
- TLCK** - N α Tosil L Lisina Clorometil Cetona (*Chloromethyl Ketone Hydrochloride*)
- TPCK** - N p Tosil L Fenilalanina Clorometil Cetona (*Np Tosyl Lphenylalanine Chloromethyl Ketone*)
- TOF** - Tempo de voo (*time to fly*)
- Tris-HCl** - Cloridrato de Trizma
- μ g** - Microgramas
- UIBBM** - União Internacional de Bioquímica e Biologia Molecular
- μ L** - Microlitros
- Z-phe-arg-MCA** - Benzoil-Fenilalanina-Arginina-Aminometilcumarina

LISTA DE ANEXOS

ANEXO I – Tabela das proteínas de intestino de fêmeas de *Culex quinquefasciatus* alimentadas com açúcar envolvidas em processos biológicos importantes como digestão, detoxificação e interação parasito/vetor.

ANEXO II – Manuscrito reunindo os resultados referentes as peptidases expressas no intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar intitulado: *In-depth characterization of trypsin-like serine peptidases in the midgut of the sugar fed Culex quinquefasciatus.*

MATERIAL SUPLEMENTAR

O material suplementar encontra-se nos arquivos de mídia removível anexado na última página deste trabalho.

Suplementar I.

Suplementar II.

Suplementar III.

1. INTRODUÇÃO

Mosquitos como vetores biológicos

Os dípteros *Culex quinquefasciatus*, *Aedes aegypti* e *Anopheles albitalis* s.s pertencem a família Culicidae e são considerados um problema de saúde pública quando atuam como vetores de parasitos e/ou arboviroses. A família inclui 3.546 espécies, classificadas em duas subfamílias: Anophelinae com três gêneros e Culicinae que possui 109 gêneros, segregados em 11 tribos. A família Culicidae é monofilética, sendo a subfamília Anophelinae a linhagem monofilética basal às demais. Os mosquitos, assim denominados, são reconhecidos pelo corpo delgado, presença de probóscide e pernas longas, além do corpo coberto por cerdas e escamas (Forattini, 2002; Harbach e Besansky, 2014).

Os mosquitos apresentam um padrão de voo que provoca um zumbido irritante e uma saliva que podem causar reações alérgicas onde picam. São reportados em ambientes urbanos desde o século XVII, e desde essa época, vêm se adaptando aos ambientes antropizados (Redi, 1688; Réaumur, 1740; Forattini, 2002). A antropofagia, característica do comportamento alimentar destas espécies, provoca uma série de transtornos aos seres humanos. Desta forma, as últimas décadas foram recheadas de estudos sobre os culicídeos que vêm consolidando a importância das espécies desta família nos dinâmicos ciclos epidemiológicos. Apesar dos esforços para o controle das populações de vetores, a incidência de arboviroses, protozooses e helmintoses têm aumentado, indo em um sentido contrário as consequências esperadas de um controle eficaz. Atualmente, mais da metade da população global habita regiões onde ocorrem doenças transmitidas por mosquitos e, para piorar este cenário, a maior parte destas regiões é composta por países com dificuldades socioeconômicas. A incapacidade do estado em prover as condições ideais para uma vigilância sanitária eficaz, é magnificada pela falta de conhecimento sobre as causas, dificultando ainda mais a conscientização popular sobre as medidas de controle e comportamentos que, em muitos casos, ajudam no controle dos vetores.

Estimativas da Organização Mundial de Saúde (OMS) indicam que as doenças transmitidas pelos mosquitos destes três gêneros estão entre as principais causas de

morbidade e mortalidade nos países em desenvolvimento. Os agentes etiológicos da malária humana são transmitidos por mosquitos do gênero *Anopheles*; os vírus da febre amarela e do dengue são transmitidos pelo gênero *Aedes*, e as filariose linfáticas e encefalites virais por mosquitos dos gêneros *Culex*, *Aedes* e *Anopheles* (Bartholomay et al., 2010). Tal importância colocaram estes mosquitos na primeira onda de genomas sequenciados - *An. gambiae* em 2002, *Ae. aegypti* em 2007 e *Cx. quinquefasciatus* em 2010.



Figura 1. Fêmea de *Culex quinquefasciatus*

Culex (Culex) quinquefasciatus Say, 1823 (Figura 1) pertence ao complexo de espécies *Culex pipiens*. O gênero *Culex* com ~1200 espécies descritas é o mais diverso e amplamente distribuído, tanto em regiões tropicais como subtropicais. Apesar de possuir um comportamento antropofílico e endofílico, as fêmeas adultas exibem uma alta plasticidade em seu comportamento alimentar, que as caracterizam como uma espécie oportunista, capaz de se alimentarem de sangue humano, de cães, de aves e de gado. Esta característica faz desta espécie um importante vetor na transmissão de zoonoses entre homens e animais (Farajollahi et al., 2011; Simpson et al., 2012). Apesar das semelhanças entre os mosquitos do complexo *Pipiens*, existem diferenças vetoriais consideráveis entre eles (Turell, 2012). *Culex quinquefasciatus* é considerado vetor biológico de vários agentes etiológicos no Brasil e no Mundo, dentre eles destacam-se: Filariose Bancroftiana (*Wuchereria bancrofti*) e diversas encefalites de origem arboviral tais como a febre do Nilo Ocidental (West Nile Vírus), Oropuche, encefalite Equina Ocidental e Venezuelana, Chikungunya, encefalite Japonesa, encefalite de São Luís (St. Louis) (Holder et al., 1999; Sardelis et al., 2001; Forattini, 2002; Goddard et al., 2002). Também tem sido sugerida a participação de *Cx. quinquefasciatus* na transmissão de parasitos que acometem os animais, como a malária aviária, provocada por protozoários do gênero *Plasmodium* e *Dirofilaria immitis*, um nematóide conhecido como verme do coração do cão (Ahid et al., 2000). O genoma no *Culex quinquefasciatus* possui um repertório de 18,883 genes codificadores de proteínas, o qual é 22% maior que o número de genes codificantes do *Aedes aegypti* e 52% maior que o do *Anopheles gambiae*. Foram observadas múltiplas expansões de algumas famílias gênicas, incluindo receptores gustatórios e olfatórios, genes da glândula

salivar e genes associados com a detoxificação de xenobióticos (Arensburger et al., 2010).



Figura 2. Fêmea de *Aedes aegypti*

Aedes (Stegomyia) aegypti Linnaeus, 1762 (Figura 2) pertence ao subgênero *Stegomyia*, assim como outro importante vetor da dengue, *Aedes (Stegomyia) albopictus*. Atualmente, tem sido sugerido a alteração de seu nome para *Stegomyia (Stegomyia) aegypti* (Harback, 2015). *Ae. aegypti* é originariamente uma espécie africana descoberta no Egito e assim como todo o subgênero *Stegomyia*, deveriam ter sua biogeografia restrita às regiões do chamado Velho Mundo. Entretanto, devido ao desenvolvimento das comunicações comerciais esta espécie foi introduzida na região neotropical. Sua distribuição geográfica atual inclui a maior parte das áreas tropicais e subtropicais da África, Ásia, Américas e Oceania (Forattini, 2002). Essencialmente, *Ae. aegypti* se distribui nas áreas urbanas e suburbanas, onde as alterações antrópicas proporcionam a sua proliferação. Trata-se, provavelmente, da espécie de mosquito mais sinantrópica, coexistindo quase estritamente com os humanos.

Acredita-se que *Ae. aegypti* tenha sido introduzido no Brasil em dois momentos. O primeiro, durante o tráfico negreiro entre as costas africanas e brasileiras no período de colonização; e o segundo, no início dos anos 70, após ter sido considerado erradicado do país em 1955. Essa invasão se deu, provavelmente, em decorrência do trânsito de humanos e mercadorias originárias de países que não erradicaram o mosquito como a Venezuela, Guianas, Estados Unidos e vários países centro-americanos. A falta de planejamento durante os projetos de urbanização, principalmente em países em desenvolvimento, juntamente com o incremento dos meios de transporte, a falta de controle vetorial eficaz e a globalização, têm provocado uma expansão geográfica das epidemias de dengue no mundo.

Atualmente, a dengue é endêmica em cerca de 100 países na Ásia, nas Américas, na África, no Pacífico e no Caribe, sendo a doença infecciosa mais importante e que afeta áreas urbanas tropicais, atingindo mais de 390 milhões de pessoas a cada ano e com cerca de 22 mil mortes (OMS, 2015). O surgimento e disseminação de novas arboviroses como o Zika vírus e o Chykungunya vírus,

reforçam ainda mais a importância deste gênero como vetor de doenças urbanas e silvestres.

Ae. aegypti tem sido objeto de pesquisas frequentes no campo da genética e é o culicíneo melhor caracterizado. As facilidades de manutenção em condições de laboratório, aliadas a importância epidemiológica, fomentaram os estudos sobre esse inseto. No mundo, *Ae. aegypti* existe em pelo menos dois tipos, denominados *Ae. aegypti formosus* (tipo selvagem original encontrada na África) e *Ae. aegypti aegypti* (forma urbana amplamente distribuída). No Brasil, a variabilidade enzimática permitiu a identificação de três populações naturais de *Ae. aegypti aegypti*, com 1,8% de diferenciação (Dinardo-Miranda e Contel, 1996).

O genoma de *Ae. aegypti* foi sequenciado em 2005 pelo “Broad Institute” e “The Institute for Genomic Research”, utilizando uma população de larvas recém eclodidas, extraídas da cepa “Liverpool”. O genoma desta espécie é o maior dentre os culicídeos já sequenciados, com 1376 Mb, o qual é cerca de 5 vezes maior que o genoma do *An. gambiae* que possui 273 Mb. Entretanto, o genoma de *Ae. aegypti* possui várias regiões repetidas, sendo metade considerado elementos transponíveis. Desta forma, a diferença entre o número de genes codificantes não é discrepante entre as duas espécies – 15.419 em *Aedes aegypti* e 12.457 em *Anopheles gambiae* (Sinkins et al. 2007).



Figura 3. Fêmea de *Anopheles* sp. espécies não podem ser facilmente distinguidas na fase adulta, porém, análises morfológicas minuciosas, izoenzimáticas e genéticas demonstram claramente a existência deste complexo. (Sinka et al., 2010).

Anopheles (Nyssorhynchus) albitarsis Lynch Arribálzaga, 1878 (Figura 3) é membro de um complexo que inclui pelo menos cinco espécies crípticas – *An. albitarsis* (espécies A, B e E); *An. marajoara* (espécie C) e *An. deaneorum* (espécie D) (Sinka et al., 2010). Estas

Membros deste complexo possuem grande distribuição geográfica além de alta variabilidade comportamental, apresentando diferentes papéis epidemiológicos. Os representantes do complexo *An. albitarsis* são bem distribuídos na região meridional do Brasil, com exceção do *An. marajoara* que é encontrado apenas nas regiões litorâneas destas áreas. Os mosquitos do complexo *Albitarsis* estão fortemente associados com a monocultura de arroz (Forattini et al., 1994). Os adultos são

geralmente exofílicos mas podem se alimentar dentro e das habitações humanas. O comportamento alimentar destes animais parece ser variado uma vez que se alimentam de humanos e animais durante o dia e a noite (Sinka et al., 2010).

Aproximadamente, 460 espécies de anofelinos são conhecidas e ~60 podem transmitir os diferentes agentes etiológicos da malária (*Plasmodium vivax*, *P. falciparum*, *P. malariae*, *P. knowlesi* e *P. ovale*) na natureza (Zhou et al., 2014; Neafsey et al., 2015). As espécies de anofelinos neotropicais estão classificadas em cinco subgêneros: *Nyssorhynchus*, *Kerteszia*, *Anopheles*, *Stethomyia* e *Lophopodomyia* (Faran e Linthicum 1981; Wilkerson e Sallum 1999), porém, os vetores dos agentes causadores de malária humana são encontrados apenas nos três primeiros. Além da malária, alguns anofelinos podem transmitir arbovírus como o O'nyong-nyong vírus e nematoides como a *Wuchereria bancrofti* (Zhou et al., 2014, Neafsey et al., 2015).

An. albitarsis s.s (ou espécie A), utilizado neste trabalho, não possui o genoma completo sequenciado, tendo apenas depositada a sequência completa do DNA mitocondrial. Cerca de 20 espécies de anofelinos, entre elas espécies neotropicais como o *Anopheles darlingi*, possuem o genoma completamente sequenciado e já depositados no *VectorBase* (<http://biomart.vectorbase.org>), um banco de dados com ferramentas de bioinformática dedicado aos vetores. Adicionalmente, dados de transcriptoma também estão depositados em bancos de dados como o UNIPROT (<http://uniprot.org>), fazendo com que o banco de dados do gênero *Anopheles* atinja um número expressivo de 79 mil sequências depositadas. O banco de dados dos culicídeos está em constante atualização. Recentemente, o genoma de 16 anofelinos foi sequenciado e disponibilizado nestes bancos de dados (Neafsey et al., 2015).

Intestino dos mosquitos

O tubo digestivo dos mosquitos é formado por uma camada de células epiteliais e uma camada acelular de lâmina basal, voltada para a hemocele. Este tubo se divide em três regiões: (i) intestino anterior, envolvido na ingestão, condução e estoque de alimento, sendo revestido por quitina e não secreta enzimas digestivas; (ii) intestino médio (estômago), responsável pela síntese e secreção das enzimas digestivas e

componentes da membrana peritrófica – uma delgada membrana que separa o intestino médio em dois compartimentos e pela absorção do alimento digerido; (iii) intestino posterior, envolvido na regulação da composição da hemolinfa e na excreção e eliminação do material não digerido proveniente do intestino médio e tubos de Malpighi (Romoser, 1996) (Figura 4).

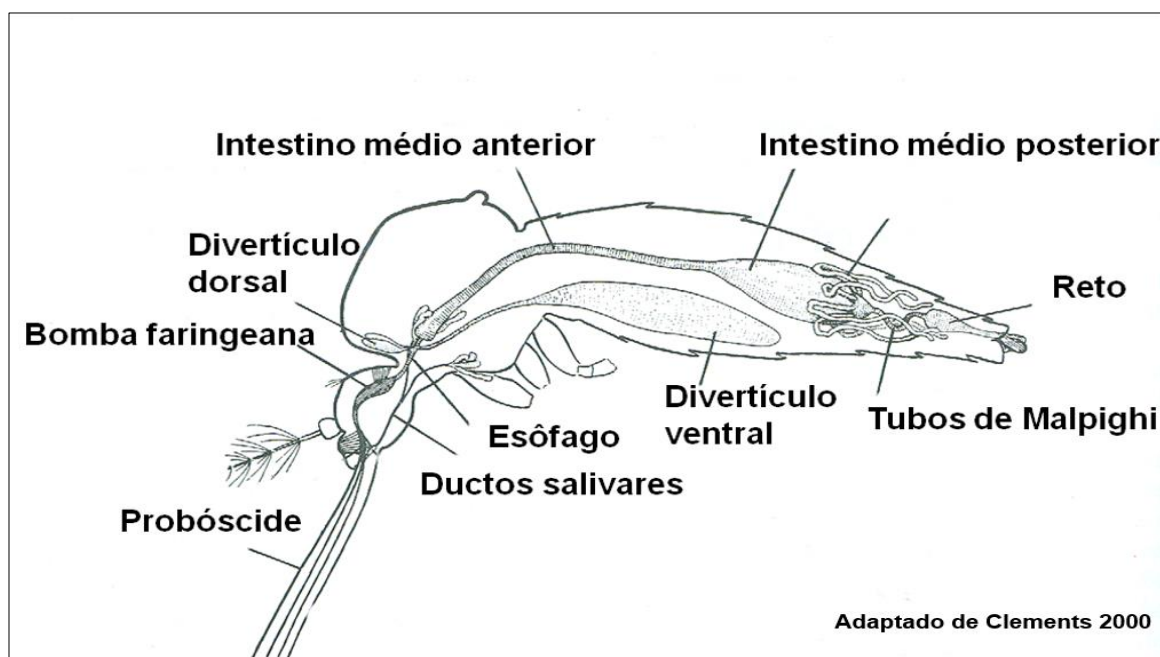


Figura 4. Esquema ilustrativo dos órgãos internos dos mosquitos. Adaptado de Clements, 2000.

O sangue é ingerido somente pelas fêmeas e vai direto para o intestino médio, onde ocorre, em sua região posterior, a digestão e absorção dos nutrientes (Romoser, 1996; Terra e Ferreira, 1994). As fêmeas se alimentam de sangue a fim de obter os nutrientes adequados para a realização do ciclo gonotrófico. Sem o repasto sanguíneo, as fêmeas dos mosquitos anautógenos não conseguem os aminoácidos necessários para a produção das proteínas do vitelo e a energia metabólica suficiente para prover o completo desenvolvimento dos oócitos. A composição do sangue é principalmente de proteína (~ 500 ng em 2 μ l de peso seco de sangue), das quais 80% é composto de hemoglobina (~ 300 ng), albumina do soro (~ 50 μ g) e imunoglobulina (~ 15 μ g), com baixas quantidades de lipídios e carboidratos (Isoe et al., 2009). O néctar vegetal é ingerido pelos adultos, machos e fêmeas, sendo estocado no intestino anterior em uma região impermeável (divertículos = papo). Quando necessário, o conteúdo do divertículo é liberado para o intestino médio para ser digerido e absorvido.

Nesse caso, o processo ocorre com mais intensidade na região anterior do intestino médio (Lopes et al. 2006)

Para uma digestão eficiente, o intestino médio dos mosquitos secreta grande quantidade de enzimas proteolíticas, carboidrases e esterases (Clements, 2000). Estas enzimas digestivas atuam ao longo de toda a superfície do bolo alimentar (Mackenzie et al., 2004), o qual é separado do epitélio do intestino médio por uma camada de glicoproteína e quitina, conhecida como a membrana peritrófica (MP). A MP protege o epitélio das enzimas digestivas e serve como uma barreira contra as partículas abrasivas, toxinas e parasitos ingeridos (Figura 5).

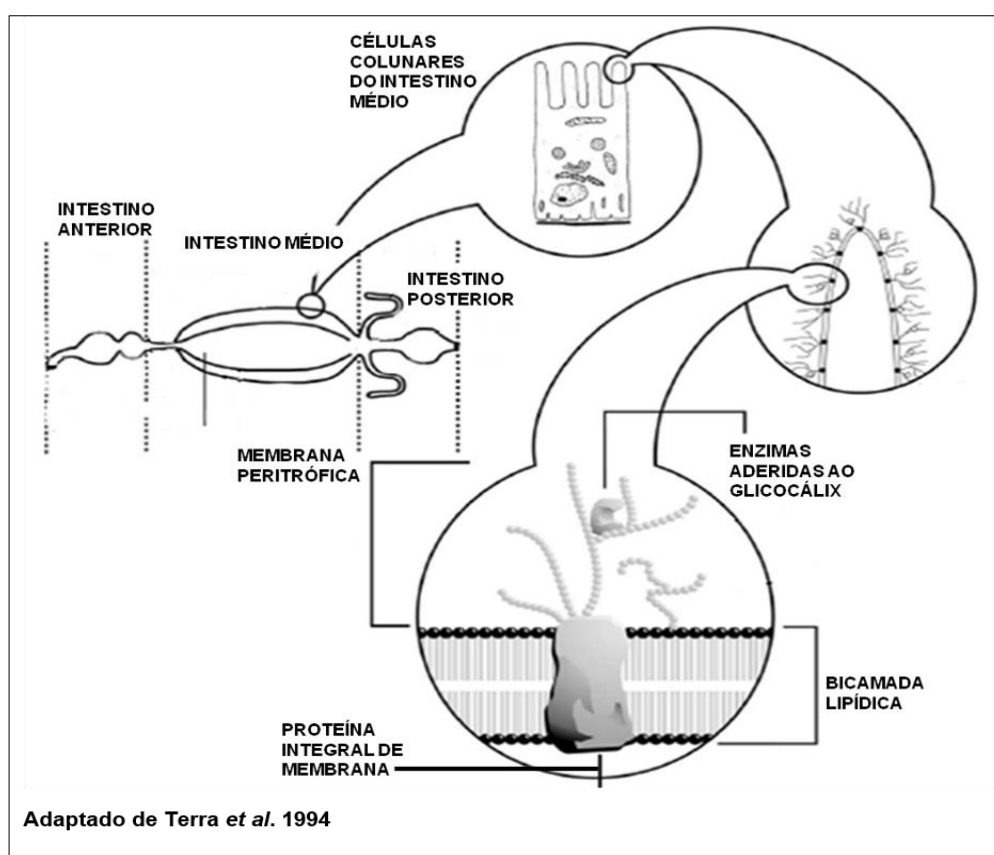


Figura 5. Esquema ilustrativo geral do intestino dos mosquitos e suas compartimentalizações. As ampliações mostram as células colunares intestinais, as microvilosidades e enzimas aderidas ao glicocálice. A membrana peritrófica é representada pelas duas linhas mais internas no intestino médio. Adaptado de Terra et al., 1994.

A MP também proporciona um processo de compartimentalização da digestão, onde as variadas peptidases trabalham em um processo concatenado: (i) ocorre a digestão inicial pela ação das endopeptidases em uma região chamada de espaço endoperitrófico; (ii) ocorre a digestão final pela ação das exopeptidases em um exíguo compartimento, entre a membrana peritrófica e as células epiteliais, chamado de espaço ectoperitrófico (Figura 5).

As moléculas de proteínas adquiridas durante a alimentação são clivadas inicialmente em peptídeos grandes pela ação de endopeptidases, as quais devem apresentar um tamanho suficiente para possibilitar a passagem através da membrana peritrófica. Por sua vez, os grandes peptídeos são reduzidos pela ação de exopeptidases e dipeptidases. Essas enzimas envolvidas na digestão intermediária atuam livres no fluido do espaço ectoperitrófico, enquanto as enzimas responsáveis pela fase final da digestão agem ligadas às células da microvilosidades do intestino médio. De fato, insetos holometábolos como *Musca domestica* (Diptera), *Spodoptera frugiperda* (Lepidoptera) e *Aedes aegypti* (Diptera) apresentam atividades de tripsinas majoritariamente no intestino médio, podendo estas serem solúveis, aderidas ao glicocálice ou à membrana das células do epitélio (Figura 5). (Terra e Ferreira, 1994; Clements, 2000).

Peptidases

As peptidases são enzimas que hidrolisam ligações peptídicas. A União Internacional de Bioquímica e Biologia Molecular (UIBBM) tem recomendado o uso do termo peptidase para o subgrupo de hidrolases de ligações peptídicas desde 1984. O termo protease é um sinônimo de peptidase que englobam dois tipos de enzimas - endopeptidases e exopeptidases (classificadas de acordo com o banco de dados *MEROPS*, <http://merops.sanger.ac.uk/index.shtml>). O termo proteinase é também usado como sinônimo para endopeptidase (Barret et al., 2004). As endopeptidases (E.C. 3.4.21-24) são subdivididas de acordo com seus mecanismos catalíticos em quatro principais subclasses: serina peptidases, cisteína peptidases, aspártico peptidases e metalo peptidases (Barret et al., 2004; Rawlings et al., 2012). As exopeptidases (E.C. 3.4.11-19) podem clivar ligações peptídicas na região amino ou

carboxi terminal. Quando cliva um único resíduo na região N terminal é chamada de aminopeptidases (E.C. 3.4.11), um dipeptídeo recebe o nome de dipeptidil-peptidases e quando cliva um tripeptídeo é nomeada de tripeptidil-peptidases. Quando cliva na região C terminal, liberando um único aminoácido, denominam-se carboxipeptidases (E.C. 3.4.16-18); e quando libera um dipeptídeo, são as peptidil-dipeptidases. Algumas exopeptidases são específicas para dipeptídeos (dipeptídeo-hidrolases, E.C.3.4.13) e existem ainda as omega peptidases que clivam ligações isopeptídicas (Barrett et al., 2004).

Os primeiros estudos sobre peptidases começaram aproximadamente no início do século 20, envolvendo principalmente as tripsinas e quimiotripsinas de vertebrados. Naquele momento, algumas enzimas foram isoladas e propriedades cinéticas, características moleculares e a natureza evolutiva dessas enzimas e zimogênios foram exploradas (Muhlia-Almazán et al., 2008). A tripsina do pâncreas bovino foi a primeira a ser cristalizada em 1930 (Northrup e Kunitz, 1931) e quase toda sequência do zimogênio foi conhecida pelo sequenciamento de proteínas nos anos 60 (Walsh et al., 1964). Com o surgimento da cristalografia por raios X, a estrutura tridimensional pode ser revelada (Huber et al., 1974). Já as tripsinas de invertebrados foram descritas apenas nos anos 60, quando mostraram que invertebrados terrestres e aquáticos sintetizam e secretam tripsinas e outras peptidases (Muhlia-Almazán et al., 2008). Atualmente, está bem documentado que invertebrados sintetizam parálogos e ortólogos dessas enzimas (Zhu et al., 2000; Heger e Ponting 2008; Wu et al., 2009). O sequenciamento gênico de algumas espécies tem possibilitado identificar quais são os genes que codificam para estas enzimas, como estão organizados no genoma e como podem ser regulados (Muhlia-Almazán et al., 2008). Entretanto, existe uma dificuldade em se estudar a cinéticas de atividade e as estruturas tridimensionais de serina peptidases de insetos, devido, principalmente, à dificuldade técnica em purificá-las e expressá-las de maneira recombinante. Na literatura, há poucos casos descritos de expressão recombinante de tripsinas de *Anopheles gambiae* e *Aedes aegypti* em bactérias (Müller et al., 1993; Brackney et al., 2008; Rascón et al., 2011). A quimiotripsina digestiva de *Solenopsis invicta* (Botos et al., 2000) é a única serina peptidase de inseto com estrutura tridimensional resolvida.

Em algumas famílias como a dos mosquitos, as principais enzimas digestivas responsáveis são as serina peptidases do tipo tripsinas e quimiotripsinas, bem como

as carboxi- e amino-peptidases. As serina peptidases do tipo tripsinas (SPtrips), quimiotripsinas e elastases pertencem ao clã SA da família S1 e possuem três domínios clássicos conservados: (i) catalítico, (ii) de reconhecimento ao substrato (Muhlia-Almazán et al., 2008) e (iii) de ativação do zimogênio (Hedstrom, 2002). A tríade catalítica conservada é composta pelos resíduos His57, Asp102 e Ser195, e a de reconhecimento ao substrato pelo resíduo Asp189 (numeração de acordo com a quimiotripsina bovina CTRA_BOVIN_P00766 - UNIPROT).

As serina peptidases do tipo tripsina (SPtrips) (EC 3.4.21.4) são as mais abundantes em intestino de muitas ordens de insetos, inclusive em mosquitos. São conhecidas como enzimas de alta especificidade, uma vez que elas clivam ligações peptídicas preferencialmente no lado carboxila de dois tipos de aminoácidos: lisina e arginina (Craick et al., 1985). Enquanto isso, quimiotripsinas (EC 3.4.21.1) apresentam maior variabilidade em seus substratos preferenciais e clivam ligações peptídicas no lado carboxila de resíduos de: fenilalanina, tirosina, triptofano e leucina (Hedstrom et al., 1992).

As tripsinas encontradas em animais superiores possuem pouca identidade de sequência quando comparadas as tripsinas de microrganismos, mas as estruturas terciárias de todas tripsinas conhecidas são extremamente conservadas. Utilizando métodos imunológicos, foi observado um alto grau de reação cruzada entre as tripsinas de vertebrados e invertebrados, sugerindo grande similaridade na estrutura tridimensional das tripsinas desses organismos. As partes funcionais da enzima são mais conservadas que as demais regiões. Além da estrutura tridimensional conservada e os resíduos da tríade catalítica inteiramente conservados, os aminoácidos adjacentes ao resíduo do sítio ativo também são conservados - His57 (TAAHC), Asp102 (DIAL) e Ser195 (GDSGGP). Apesar da mesma tríade catalítica, essas enzimas possuem diferenças em sua sequência de aminoácidos, as quais permitem variações estruturais e na especificidade a diferentes substratos, garantido o desempenho de diferentes funções (Neurath, 1984; Perona e Craik, 1995). Mutações que afetam a sequência primária de aminoácidos das enzimas ocorreram em demasia, entretanto, a atividade catalítica e especificidade foram pouco afetadas, o que indica o surgimento das tripsinas antes da divergência entre vertebrados e invertebrados (Arnon e Neurath, 1969).

A estrutura terciária das tripsinas é caracterizada pela presença de dois β -barris muito similares, compostos por seis folhas betas ligadas entre si e uma α -hélice C-terminal (Baptista et al., 1998). A conformação tridimensional das enzimas se dá de tal forma que os resíduos do sítio catalítico ficam próximos e se posicionam entre as duas cadeias β -barris. Com relação aos aminoácidos do sítio catalítico, a Ser195, atua como o nucleófilo durante a hidrólise. A His57 atua como uma base e ASP102 estabiliza o correto emparelhamento da His57. O substrato forma uma folha β -antiparalela com o sítio de ligação à proteína. A especificidade está fortemente relacionada com a cadeia lateral da Asp189 altamente conservada, a qual é positivamente carregado e interage com os resíduos negativos Lys ou Arg do substrato. Além disso, outros resíduos próximos ao domínio catalítico interagem com o substrato e interferem na especificidade das enzimas ao substrato. O conjunto de resíduos da enzima que interage com uma única posição do substrato é chamado “subsítio” sendo indicado como S1, S2, ..., Sn no sentido N-terminal, e S1', S2', ..., Sn' no sentido C-terminal do substrato. Já os resíduos do substrato recebem uma numeração P1, P2, ..., Pn no sentido N-terminal, e P1', P2', ..., Pn' no sentido C-terminal do substrato. O sítio de ligação ao substrato está localizado entre as posições P1 e P1' (Schechter e Berger 1968; Bode et al., 1989; Hedstrom, 2002) (Figura 6). Desta forma, a especificidade da clivagem das tripsinas depende do volume/tamanho, estrutura/conformação e polaridade/carga/hidrofobicidade de partes específicas da superfície da proteína que ficará em contato com o substrato.

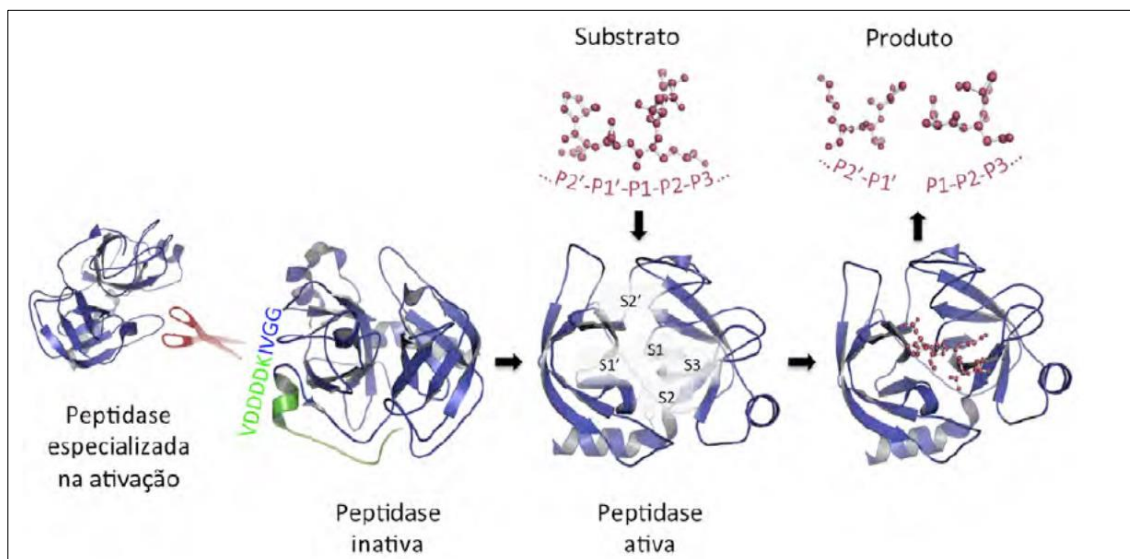


Figura 6. Esquema das diferentes etapas envolvidas na ativação e atividade das tripsinas. O motivo de ativação do zimogênio representado corresponde ao consenso encontrado em tripsinas de mamíferos. As posições dos resíduos do substrato (P2', ..., P3) e dos subsítios da enzima (S2', ..., S3) seguem a nomenclatura proposta por Schechter e Berger, 1968.

A composição de aminoácidos de cada sub-sítio (sequência primária) e as suas posições espaciais (estrutura terciária) são determinantes na especificidade enzima-substrato. Esses subsítios de ligação entre tripsinas-substratos devem divergir entre as enzimas dos insetos, e apesar da alta similaridade das sequências e estrutura tridimensional, devem haver diferenças na estrutura primária suficientes para que as enzimas apresentem diferentes especificidades e propriedades cinéticas. Por exemplo, as tripsinas de quatro ordens distintas de insetos (Dictyoptera: *Periplaneta americana*; Coleoptera: *T. molitor*; Diptera: *M. domestica*; Lepidoptera: *Diatraea saccharalis*) apresentam diferentes propriedades cinéticas (Lopes et al., 2006). Tais enzimas podem apresentar características espécie-específicas desde que haja diferenças significativas em hábitos alimentares, composição do alimento ingerido e no processo de digestão proteica. Essas singularidades são respostas adaptativas das espécies a diferentes estilos de vida e ambientes (Muhlia-Almazán et al., 2008).

As SPs que atuam no lúmen do intestino são direcionadas para este órgão via secreção. Sabe-se que peptídeos sinais tendem a ser mais similares em genes próximos e drasticamente diferentes em ortólogos distantes (Veitia e Carburet, 2009). Outra característica relevante dessas SPs é o fato delas necessitarem passar por um

processo de ativação. Tripsinogênios e quimiotripsinogênios, os quais compreendem os zimogênios (formas inativas) de tripsinas e quimiotripsinas, respectivamente, são ativados pela clivagem de uma sequência N-terminal de tamanho variável. A presença de um aminoácido básico (arginina ou lisina), imediatamente anterior ao resíduo de isoleucina inicial das peptidases de invertebrados (R/L-IVGG), faz dessa região um substrato potencial para as próprias tripsinas. De fato, uma grande parte dos zimogênios das serina peptidases (incluindo os quimiotripsinogênios) são ativados por tripsinas (Rinderknecht, 1986), o que dá a essas enzimas um papel chave na cascata de ativação de diversas peptidases. No entanto, o processo de ativação dos zimogênios ainda não está claramente entendido, sendo que a região de pré-ativação como um todo, incluindo o peptídeo sinal e a via de secreção dessas enzimas, permanece pouco explorada.

As SPtrips digestivas também têm sido descritas participando de processos de interação entre patógenos e vetores. Dentre várias isoformas de SPtrips, apenas uma parece limitar a infectividade do vírus Dengue (DENV-2) no intestino de *Ae. aegypti* (Brackney et al., 2008). Apesar do ambiente proteolítico do intestino médio dos mosquitos degradar muitos microrganismos e vírus que ali chegam, limitando a infecção, arbovírus de diferentes famílias e protozoários como *Plasmodium* spp. fazem bom uso dessa situação. Por exemplo, DENV-2 (Flaviviridae), La Crosse vírus (Bunyaviridae) e Blue tongue vírus (Reoviridae) utilizam as peptidases digestivas para o processamento proteolítico das proteínas da superfície viral, aumentando a adesão à células do intestino (Ludwig et al., 1989; Ludwig et al., 1991; Mertens et al., 1996; Molina-Cruz et al., 2005; Brackney et al., 2008). Para conseguirem atravessar a membrana peritrófica, o *Plasmodium gallinaceum* secreta uma quitinase que precisa ser ativada pelas peptidases digestivas (Shahabuddin et al., 1996).

Além de atuarem no intestino dos mosquitos, as serina proteases têm sido descritas participando das repostas imunes, como a coagulação da hemolinfa, síntese de melanina e de peptídeos antimicrobianos (Gorman e Paskewitz, 2001). A maioria das serina proteases reguladas após a infecção parasitária pertencem à categoria de domínio CLIP. No genoma de *An. gambiae* representam 41 dos 305 genes preditos de serina peptidases (Christophides et al., 2002; Zdobnov et al., 2002). Proteases CLIP são importantes na melanização (Volz et al., 2006; Sriwichai et al., 2008), uma vez que possuem um papel pivotal na ativação da profenoloxidase (proPO) que culminará

na formação das cápsulas de melanina que envolvem bactérias, protozoários e nematoides que possam estar infectando os mosquitos (Paskewitz et al., 1988; Hillyer et al., 2003; Shiao et al., 2001). As proteases também são importantes na reprodução dos mosquitos (Mancini e colaboradores 2011).

Biossíntese e secreção das tripsinas em mosquitos

Em insetos que se alimentam continuamente (larvas de dípteros), a síntese e secreção de peptidases parece ser constitutiva, ao passo que em insetos que não tem hábito alimentar contínuo (predadores ou hematófagos), a síntese parece ser finamente regulada (Lehane e Billingsley, 1996).

A síntese e expressão de SPtrips estão mais bem caracterizadas em *Aedes aegypti* e *Anopheles gambiae*. Os primeiros relatos feitos em 1950, relacionavam a atividade proteolítica do intestino de *Aedes aegypti* com a ingestão do sangue (Fisk, 1950). Foi observado que endopeptidases do tipo tripsina eram responsáveis por 75% da atividade proteolítica total, sendo induzidas após a alimentação (Gooding 1966 a; Gooding 1966 b; Briegel e Lea, 1975; Graf e Briegel, 1989). Em 1989, foi demonstrado que um grupo “iniciais” de tripsinas (*early trypsins- ET*) eram detectadas no intestino médio dentro de 2 h após a alimentação, enquanto um outro grupo “tardio” (*late trypsins- LT*), eram detectados 12 h após a alimentação. Nesta época, as tripsinas tardias foram apontadas como a responsável por quase toda a digestão endoproteolítica do sangue (Graf e Briegel, 1989). Nenhuma enzima proteolítica foi detectada em fêmeas de *Ae. aegypti* em jejum, sendo estas, sintetizada somente após o repasto sanguíneo (Graf e Briegel, 1989; Kalhok et al., 1993). Em 1991, utilizando extratos brutos ou intestinos dissecados de mosquitos alimentados com sangue e outros tipos de proteínas, Felix e colaboradores mostraram que a digestão do sangue é de fato bifásica: a primeira fase começando nas primeiras horas após a alimentação (tradução imediata de RNAm já existentes nas células do intestino); a segunda fase com início cerca de 7-9 horas após a alimentação com sangue (com a tradução de RNAm recém-sintetizados).

No fim dos anos 80, Graf e colaboradores descreveram um gene de tripsina da fase tardia (AaLT) que acreditavam ser a principal endopeptidase envolvida na fase

tardia da digestão do sangue (12 h PAS). Em 1991, esse mesmo gene foi clonado e caracterizado por Barrillas-Mury e colaboradores. Dois anos depois, mais dois genes de tripsinas de *Ae. aegypti* foram clonados e caracterizados (Kalhok et al., 1993). Estes dois novos genes clonados (3A1 e 5G1) demonstraram ser diferentes do gene da tripsina tardia (AaLT). O transcrito do gene da tripsina 3A1 mostrou-se presente em mosquitos em jejum, mas a síntese das enzimas para as quais este codifica não é iniciada até que se dê o início da alimentação com sangue (Kalhok et al., 1993). Atualmente, esta enzima é conhecida como tripsina inicial AaET (Ráscon, et al., 2011). Por outro lado, o transcrito do gene 5G1 (AaSPVI) foi expresso somente na fase tardia, com pico de expressão ocorrendo entre 18-24 h PAS (Graf e Briegel, 1989; Kalhok et al., 1993). Nesta época, em *An. gambiae* foi demonstrado que dois genes (Antryp1 e Antryp2) são induzidos após o repasto sanguíneo (Müller et al., 1995).

Os estudos em SPtrips foram acompanhando o desenvolvimento de novas ferramentas de biologia molecular. Foi observado que o RNAm de tripsina iniciais estava presente em fêmea de *Ae. aegypti* em jejum, três dias após emergirem das pupas (Kalhok et al., 1993; Noriega et al., 1996). Após a alimentação, o nível de RNAm de tripsina inicial diminui, correlacionando com um aumento nos níveis iniciais da enzima no intestino, indicando que a AaET é regulada ao nível da tradução (Noriega et al., 1996). Uma vez que o gene de AaLT foi demonstrado ser expresso somente após um repasto sanguíneo, foi proposto que a regulação do gene se dá ao nível da transcrição (Barrillas-Mury e Wells, 1993).

Em 1995, Barrillas-Mury e colaboradores propuseram um mecanismo de *feedback* que envolve a atividade da enzima da fase inicial AaET e a transcrição dos genes da enzima da fase tardia AaLT. Neste estudo, os autores alimentaram as fêmeas de *Ae. aegypti* com refeições proteicas que continham um inibidor de tripsina conhecido como inibidor de tripsina de soja (STI). Ensaios de inibição da atividade da AaET foram realizados, em diferentes tempos após a alimentação, e os transcritos de RNAm de AaLT foram analisados. Foi observado que quando a atividade da AaET era inibida, os níveis de RNAm das AaLT também não eram observados. Como experimento controle para mostrar que o efeito se devia à inibição de tripsina inicial e não algum outro efeito provocado pelo STI, uma refeição de sangue pré-digerido com tripsina bovina, na presença do inibidor de tripsina de soja (que não inibe a tripsina bovina) foi dada para fêmeas de *Ae. aegypti*. Esta refeição de sangue pré-digerido

restaurou os níveis de RNAm das tripsinas tardias e conseqüentemente o de proteínas traduzidas. Desta forma, concluíram que os produtos iniciais da digestão realizada pelas enzimas da fase inicial eram necessários para a transcrição do gene de AaLT.

A partir desses estudos, os autores propuseram que a ingestão de sangue ativa a tradução do RNAm de AaET, resultando em uma digestão inicial limitada do sangue e que os produtos desta digestão, talvez aminoácidos ou peptídeos, ativariam a transcrição e tradução dos genes de AaLT, levando à completa digestão do alimento (Barrillas-Mury et al., 1995). Este modelo de regulação das proteases no intestino médio dos mosquitos, em especial de *Ae. aegypti*, foi aceito por muitos anos. Entretanto, não foi possível a reprodutibilidade desses resultados, isto é, não se conseguiu demonstrar a inibição da atividade da AaET e conseqüente transcrição do gene AaLT (Ráscon, et al., 2011).

Em 2006, Lu et al., revisou o trabalho feito por Barrillas-Mury em 1995, e determinou que o inibidor de tripsina de soja utilizado para os experimentos estava contaminado com flavonas vegetais. Estes contaminantes estariam interferindo no metabolismo dos mosquitos, o que explicaria a inibição observada da expressão do gene da AaLT. Lu e colaboradores, repetiram os experimentos utilizando diferentes inibidores, incluindo um inibidor de tripsina de soja ultrapuro e bloqueios do gene da AaET com RNAi. Com esta abordagem, a expressão dos genes de AaLT foi observada mesmo com a inibição de AaET. Recentemente, estudos de bloqueios por RNAi revelaram que a AaLT não é a principal responsável pela degradação do substrato BApNA observada em extratos de intestino médio de mosquitos alimentados com sangue. Além disso, mostraram que outra peptidase da fase tardia, conhecida como AaSPVI (antiga 5G1), é, na verdade, a principal tripsina responsável pela atividade proteolítica no intestino médio do *Aedes aegypti* (Brackney et al., 2008; Isoe et al., 2009). As técnicas de biologia molecular abriram novas possibilidades para entender as famílias multigênicas das peptidases, assim como novas funções para os genes de tripsinas puderam ser descobertos (Heger e Ponting 2008).

A recente descoberta que a atividade da tripsina inicial não é requerida para a expressão dos genes das peptidases de fase tardia, e que a tripsina tardia AaLT não é uma SPtrip clássica, indicam que, mesmo com auxílio das ferramentas de biologia molecular, muitas características sobre a fisiologia do intestino permanecem obscuras. Porém, ficou claro que a função e regulação das tripsinas é um processo

complexo e que conta com a expressão de um repertório de genes. Outras enzimas, que não as tripsinas, também parecem ser induzidas pela alimentação com sangue. As aminopeptidases, carboxipeptidase A e carboxipeptidase B, tiveram sua ação estimulada pela ingestão de sangue, apresentando maior atividade durante o pico de proteólise da tripsina, que ocorre entre 20 e 24 horas após a ingestão de sangue pelas fêmeas de *Ae. aegypti* (Noriega et al., 1996; Noriega et al., 2002). Em *Anopheles stephensi* e *An. darlingi*, serina peptidases e aminopeptidases identificadas no intestino de mosquitos adultos provavelmente estão relacionadas à digestão das proteínas do sangue (Rosenfeld e Vandenberg, 1998; Okuda et al., 2005). Desta forma, vários genes estariam envolvidos na síntese desses tipos de tripsinas nas distintas espécies de mosquitos e parecem possuir diferentes padrões de expressão. Enquanto a expressão de alguns genes de SPtrips é constitutiva, a expressão de outros genes é induzida após a alimentação com sangue. Um rico repertório de enzimas promove um eficiente mecanismo para a degradação de proteínas e vantagens adaptativas para os insetos que se alimentam de fontes ricas em inibidores de proteínas (Terra e Ferreira, 1994), como é o caso do sangue.

Por apresentarem um ciclo de vida holometábolo, é possível que muitas serina proteases sejam fundamentais para outros estágios evolutivos nos mosquitos. Já foi demonstrado que a expressão dessas peptidases em culicídeos pode ser sexo-estágio e órgão-específica (Venâncio et al., 2009, Mesquita-Rodrigues et al., 2011; Borges-Veloso et al., 2012; Saboia-Vahia et al., 2013). Isto mostra que a diversidade de enzimas pode estar relacionada com atuação dessas em distintas vias fisiológicas e em diferentes tecidos durante o ciclo de vida de um mosquito.

Grande parte do conhecimento produzido sobre as peptidases digestivas de mosquitos foram conduzidas anteriormente ao sequenciamento gênico dessas espécies. Durante esta época, estudos bioquímicos clássicos e análises enzimáticas *in vitro* foram realizados e muito se explorou sobre as cinéticas de atividade frente a diferentes substratos e inibidores. Com o uso da eletroforese observou-se que várias isoformas de tripsinas são sintetizadas. Após o sequenciamento de DNA de algumas espécies de mosquitos, pode-se observar que o repertório de genes que codificam para estas enzimas, em especial para as serina peptidases do tipo tripsina, é variado e contém centenas de genes preditos. Entretanto, a confirmação da identidade das SPtrips que são de fato traduzidas nas diferentes condições fisiológicas é um desafio

devido ao grande número de cópias de genes dessa família de enzimas encontrado em mosquitos. Por exemplo, de 380 genes preditos para SPtrips no genoma de *Ae. aegypti*, apenas seis tem sido caracterizados ao nível de proteína (Dana et al., 2005; Isoe et al., 2009; Brackney et al., 2010).

A dificuldade em purificar algumas isoformas de SPs é outro obstáculo para o estudo dessas enzimas, o que impede análises bioquímicas das frações puras e a cristalização para análise de sua estrutura tridimensional. O estudo de várias serina peptidases isoladas de mamíferos têm sido realizadas com amostras adquiridas pela maneira clássica, ou seja, diretamente do tecido de estudo. Por exemplo, a primeira quimiotripsina bovina isolada foi diretamente do pâncreas (Kunitz e Northrop, 1935). Para os mosquitos, tem sido feito o mesmo, já que as enzimas de interesse são retiradas dos homogenatos de intestino. Porém, o tamanho dos mosquitos de aproximadamente 2 a 3 mm atrapalha o bom rendimento das extrações. O grande repertório de genes que codificam para SPtrips também dificulta o trabalho de clonagem e expressão de todos esses genes. Estudos que contemplem essas questões são necessários para entender a expressão e síntese de tripsinas nos distintos estágios e tecidos dos insetos. Por apresentarem papel fundamental em diversos processos fisiológicos, essas enzimas são importantes alvos para o desenvolvimento de estratégias de controle de insetos vetores. Desta forma, a proteômica pode ser uma poderosa ferramenta para a detecção e identificação das peptidases presentes em baixa quantidade e em amostras complexas.

Proteômica

Apesar dos ciclos de vida de parasitos e arbovírus transmitidos por mosquitos vetores serem distintos, todos eles são ingeridos e expostos ao ambiente do intestino médio dos mosquitos. Em seguida, devem atravessar o epitélio intestinal, atingir a hemolinfa e chegar ao tecido apropriado para seu desenvolvimento. Os locais mais comuns de desenvolvimento desses parasitos incluem o intestino médio, túbulos de Malpighi, músculos torácicos, glândulas salivares e gônadas (Beerntsen et al., 2000). Conseqüentemente, a composição desses tecidos nas diferentes espécies de mosquitos é um dos determinantes para sua competência vetorial. Entretanto, pouco

se sabe a respeito da composição proteica do intestino e interação molecular entre as células do hospedeiro e os agentes infecciosos (Figura 7). Assim, a identificação de alvos promissores para o desenvolvimento de tecnologias que interrompam o desenvolvimento do parasito nos vetores, depende, em parte, do conhecimento das proteínas sintetizadas no intestino, o qual pode ser amplamente explorado por proteômica e análises de bioinformática (Beerntsen et al., 2000; Biessmann et al., 2005; David et al., 2005; Marinotti et al., 2005; Calvo et al., 2007; Waterhouse et al., 2007; Strode et al., 2008).

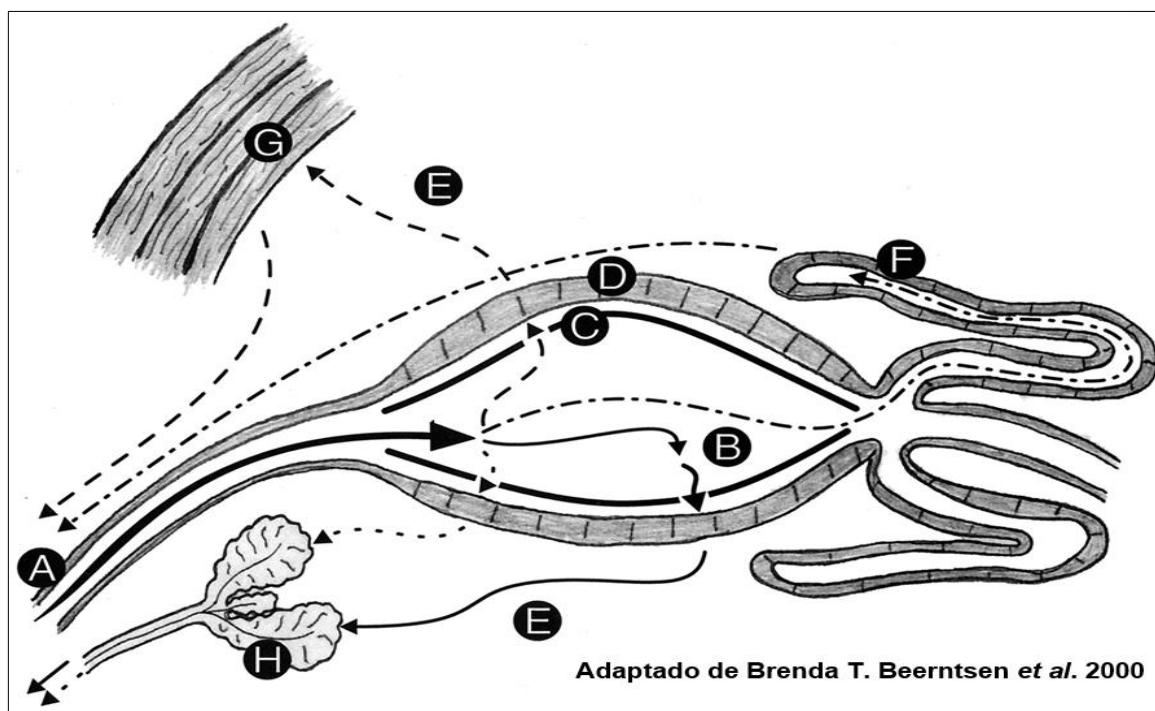


Figura 7. Rotas migratórias e locais de desenvolvimento de vírus, parasitos da malária e helmintos filarióides. Os sítios de desenvolvimento dentro do mosquito são definidos pelas letras de A-H e a rotas migratórias são representadas pelas linhas. (A): Após se alimentarem de um sangue infectado. (B): Todos os patógenos entram pelo intestino. Vírus (representados por —) entram nas células epiteliais (D), replicam e deixam as células, atravessam a hemolinfa (E) até as glândulas salivares (H), onde podem infectar um novo hospedeiro vertebrado. *Plasmodium* sp. (representados por) permanecem no lúmen do intestino por algumas horas, onde realizam a reprodução sexuada e ocorre a formação do oocineto. Este atravessa a matriz peritrófica formada (C), passa através do epitélio do intestino médio e se aderem a membrana basal. Quando as filárias que provocam doenças em humanos (-----) e cães (-.-.-) são ingeridas junto com o sangue, aquelas atravessam o epitélio do intestino e migram para os locais de desenvolvimento no músculo torácico (G), estas migram pelo intestino e se desenvolvem intracelularmente nos tubos de malpighi (F). Após o período de desenvolvimento, as larvas infectivas de terceiro estágio deixam as células do músculo torácico ou túbulos de Malpighi, atingem a hemocele e migram para a região da cabeça.

Em um encontro sobre eletroforese bidimensional realizado em Siena em 1994, o termo proteoma foi cunhado por Marc Wilkins pela primeira vez, em referência ao completo repertório de proteínas expressas pelo genoma de uma célula, tecido ou organismo em um determinado momento. Em 1997, o termo proteômica foi utilizado de forma análoga ao termo genômica, referindo-se as tecnologias envolvidas no

estudo do proteoma. Desde então, o termo tem aparecido de forma crescente na literatura. Atualmente, as diversas abordagens proteômicas desenvolvidas permitem identificar e quantificar as proteínas e suas modificações pós-traducionais, bem como a expressão diferencial destas e as interações com as demais proteínas em um organismo, tecidos ou células específicas (Twyman, 2014). Os estudos proteômicos em insetos e em outros organismos vêm sendo facilitados e estimulados graças ao sequenciamento dos genomas de diversas espécies (Ribeiro et al., 2007).

Análises de genômica comparativa têm demonstrado que o repertório genético de *Cx. quinquefasciatus* é 22% maior quando comparado ao *Ae. aegypti* e 52% maior quando comparado ao *An. gambiae*. São observadas expansões em algumas famílias de genes de *Cx. quinquefasciatus* como as dos receptores gustatórios e olfatórios, os genes da glândula salivar e de alguns genes associados à detoxificação de xenobióticos (Arensburger et al., 2010). Entretanto, a sequência genômica não é suficiente para explicar todas as variações fenotípicas de um organismo e descrever a estrutura, função e localização celular das proteínas. Algumas informações funcionais importantes, modificações pós traducionais como glicosilação e fosforilação, não podem ser observadas olhando apenas a composição dos genes. A análise do RNAm também não informa essas características e não permite prever o momento da tradução da proteína. Nem sempre a alteração na transcrição de um determinado gene se consolida na tradução imediata da proteína, uma vez que os processos pós traducionais são complexos e podem ser alterados a qualquer momento (Bartholomay et al., 2004; Biron et al., 2005; Cox e Mann, 2011).

A espectrometria de massas tem se consolidado como a principal tecnologia para a profunda caracterização da composição proteica em amostras biológicas. O estudo de moléculas grandes, como as proteínas, está diretamente relacionado ao desenvolvimento tecnológico. O estudo de proteínas por espectrometria só foi possível após o desenvolvimento das técnicas de ionização suave, MALDI (*Matrix-assisted laser desorption/ionization*) e ESI (*Electrospray Ionization*), na década de 80. Com o desenvolvimento dos métodos experimentais, da instrumentação e das abordagens de bioinformática para análise de dados, foi possível a identificação de centenas de proteínas presentes em uma amostra biológica. Hoje é possível encontrar uma diversidade de espectrômetros de massas como TOF (*time to fly*), Orbi-Trap,

Quadri-polos e até mesmo híbridos, mas a forma de ionização ainda se dá por MALDI ou ESI.

Nos primórdios dos estudos proteômicos, era possível identificar poucas proteínas por cada análise, já que dependia da laboriosa técnica de degradação de Edman (Patterson e Abersold, 2003; Biron et al., 2005). Já na era pós-genômica foi possível a identificação de várias proteínas ao mesmo tempo. Atualmente, é possível aplicações de abordagens como: descrever cada proteína que o compõe (proteômica descritiva), determinar a abundância das proteínas (proteômica quantitativa) e verificar variações induzidas pelo ambiente, condição fisiológica ou desenvolvimento biológico (proteômica diferencial ou comparativa) (Fields, 2001; Graves e Haystead, 2002; Zhu et al., 2003; Valledor e Jorrín, 2011). A utilização de técnicas de marcação de proteínas com isótopos estáveis, combinadas com a espectrometria de massas, tem promovido análises comparativas e quantitativas cada vez mais robustas. Dentre essas técnicas se destacam o ICAT, iTRAQ e SILAC. (Gygi et al., 1999; Ong et al., 2004; Ross et al., 2004; Mann 2006; Gan et al., 2007). Os estudos proteômicos são conhecidos por serem dirigidos pela descoberta, e não por hipótese, como ocorre com a maioria dos desenhos experimentais científicos (Graves et al., 2002). Entretanto, o desenvolvimento de metodologias como *Selected Reaction Monitoring* –SRM, tem mudado essa característica, pois possibilita identificar uma proteína pré-determinada. O SRM é uma metodologia utilizada em estudos proteômicos de hipótese-dirigida, que permite monitorar e analisar peptídeos alvos em uma amostra complexa, possibilitando a identificação de peptídeos pouco abundantes em uma amostra (Prakash et al., 2009; Gallien et al., 2015).

Com relação a forma de preparação de extratos proteicos para análise no espectrômetro de massas, três abordagens têm sido utilizadas: (i) *bottom-up*, um extrato bruto de proteínas é submetido a um processo de digestão enzimática por peptidases e os peptídeos produzidos são ionizados e direcionados para o detector do espectrômetro de massas; (ii) *top-down*, a amostra de proteínas é diretamente enviada ao espectrômetro de massas, e posteriormente ocorre a fragmentação da proteína; (iii) *middle-down*, na qual as proteínas são clivadas enzimaticamente em peptídeos grandes e depois submetidas à análise por espectrometria de massas (Bruce et al., 2013).

A abordagem *top-down* é boa para estudos que envolvam modificações pós-traducionais, sendo a *bottom-up* a mais utilizada. Resumidamente, a amostra biológica com as proteínas pode ser diretamente digerida ou opcionalmente pré-fracionada (eletroforese ou cromatografia) e as frações submetidas a clivagem enzimática em peptídeos. Os peptídeos são ionizados e no analisador, os íons são separados de acordo com a razão massa/carga (m/z). No primeiro analisador são obtidos os perfis de espectros de massa (MS1). Íons podem ser selecionados e submetidos à fragmentação para a obtenção dos espectros de massa *in tandem* (MS2), o resultado é a relação m/z de cada peptídeo ionizado e seus fragmentos. Os dados são organizados sob a forma de um espectro de massas, que é um gráfico em que os valores de m/z estão no eixo X e sua abundância relativa no eixo Y. As sequências dos peptídeos são utilizadas para identificar as proteínas correspondentes.

Após a obtenção dos espectros (MS1 e MS2), duas principais abordagens são utilizadas para a identificação automatizada das proteínas: (i) PSM (*peptide spectrum match*), no qual os espectros de massas experimentais são comparados com espectros teóricos gerados a partir da sequência de peptídeos oriundos de uma base de dados (sequências de genes, transcritos ou de proteínas depositadas). Este método foi utilizado inicialmente por Jonh Yates e Jimmy Eng, desenvolvedores do programa SEQUEST, e hoje é utilizado por outros softwares. Alguns programas que utilizam essa abordagem são o COMET, Mascot e X! Tandem (Eng et al., 1994; Junqueira e Carvalho, 2012). Apesar de ser considerada o padrão ouro, a dependência do banco de dados não permite que sequências com variações (polimorfismos, mutações ou erros de anotação) sejam identificadas (Junqueira e Carvalho, 2012); (ii) o *de novo sequencing* é outra abordagem utilizada na identificação e consiste na interpretação dos espectros sem a utilização de um banco de dados de espectros (Yates, 1998). É utilizada principalmente para organismos que não possuem genoma sequenciado. Os programas que utilizam essa abordagem são pNovo, PepNovo e PEAKS (Junqueira e Carvalho, 2012). Posteriormente às duas abordagens, uma análise estatística deve ser realizada para garantir a confiabilidade das identificações. O desenvolvimento das ferramentas de bioinformática também permite a identificação de proteínas de espécies desconhecidas através de buscas por homologia. Atualmente, uma série de *softwares*, muitos destes reunidos no

Expasy (<http://www.expasy.org>) foram desenvolvidos para auxiliar na interpretação da enorme quantidade de dados gerados pelos ensaios de proteoma.

Desta forma, com a espectrometria de massas é possível identificar indiretamente os genes que são expressos dentre todo o repertório disponível, ajudando na anotação funcional do genoma destes insetos. Apesar da era pós-genômica trazer simplificações nos caminhos para analisar dados do repertório gênico, as análises manuais que garantam anotações acuradas são necessárias para que se obtenha informações corretas do genoma sobre funções específicas de determinados genes. Em função da enorme quantidade de dados de sequências genômicas a anotação automática é utilizada. Porém, isso resulta em diferentes tipos de erros que podem ser resolvidos através da combinação de anotação manual e evidência experimental (Pandey e Mann, 2000; Kalume et al., 2005).

Neste trabalho, utilizamos diferentes abordagens para a identificação de proteínas totais e de peptidases ativas presentes no intestino de *Cx. quinquefasciatus*, *An. albitarsis* s.s e *Ae. aegypti*. As peptidases foram detectadas e caracterizadas por ensaios em solução e enzimografia em gel copolimerizado com substrato proteico. Em conjunto, a aplicação das duas técnicas permitiu a detecção acurada e o mapeamento do perfil de serina peptidases ativas do intestino de fêmeas alimentadas com açúcar. A influência da alimentação com sangue sobre a atividade proteolítica também foi investigada em *Aedes aegypti*.

A análise proteômica - associadas a separação por SDS-PAGE - e de bioinformática foram usadas para identificação do perfil proteico de extratos de intestino de *Cx. quinquefasciatus* e *An. albitarsis* s.s. Tais técnicas permitiram também identificar e mapear as isoformas de tripsinas ativas presentes no intestino destas espécies, e demonstrar quais dos genes preditos para estas enzimas, no genoma do mosquito, são de fato constitutivamente expressos naquele órgão nas condições de alimentação com açúcar. Análises de bioinformática revelaram as principais características das sequencias das tripsinas identificadas, tais como, o sítio ativo, motivo de ativação do zimogênio, especificidade ao substrato e presença de pontes dissulfeto. Os resultados obtidos permitiram elaborar uma ampla discussão sobre função e estrutura das peptidases. Adicionalmente, as centenas de proteínas presentes no intestino de fêmeas alimentadas com açúcar foram classificadas

segundo sua ontologia gênica e o papel de algumas famílias de proteínas historicamente importante em mosquitos foram discutidas.

2. OBJETIVOS

O objetivo geral deste projeto foi caracterizar o perfil de peptidases ativas em intestino de adultos machos e fêmeas de *Cx. quinquefasciatus*, *An. albitarsis* s.s e *Ae. aegypti*, bem como identificar por espectrometria de massas a composição do perfil proteico.

Objetivos específicos

- Caracterizar as peptidases presentes no intestino de adultos machos e fêmeas de *Cx. quinquefasciatus*, *An. albitarsis* s.s e *Ae. aegypti* alimentados com açúcar, utilizando enzimografia unidimensional e substrato fluorogênico Z-phe-arg-MCA;
- Caracterizar por enzimografia e ensaios em solução, o perfil proteolítico do intestino de fêmeas de *Aedes aegypti* em diferentes tempos após a alimentação com sangue, e padronização para a identificação por espectrometria de massas;
- Identificar, por espectrometria de massas, o perfil de proteínas totais do intestino de fêmeas de *Cx. quinquefasciatus* e *An. albitarsis* s.s alimentadas com açúcar;
- Caracterizar as sequências das peptidases identificadas por espectrometria de massas utilizando ferramentas de bioinformática;
- Classificar as proteínas totais identificadas por espectrometria de massas a partir de sua ontologia gênica e mineração dos dados.

3. MATERIAIS E MÉTODOS

Insetos

Todos os experimentos foram realizados com adultos de *Ae. aegypti* (linhagem *Rockefeller*), *An. albitarsis* s.s, e *Cx. quinquefasciatus* (linhagem *Colônia*) mantidos em colônias no Laboratório de Fisiologia e Controle de Artrópodes Vetores do Instituto Oswaldo Cruz (FIOCRUZ), situado no Instituto de Biologia do Exército (IBEx) no Rio de Janeiro. Os adultos utilizados nos experimentos possuíam 4-5 dias de idade após emergirem das pupas e eram mantidos em gaiolas na presença de solução de sacarose a 10% “*ad libitum*”. Essas gaiolas eram colocadas no interior de uma câmara de crescimento tipo BOD a 28°C, com umidade relativa em $75 \pm 5\%$ e fotoperíodo de 12 horas claro: 12 horas escuro. As fêmeas eram alimentadas, quando necessário, em cobaias anestesiadas.

Dissecção dos órgãos

Para a obtenção dos órgãos isolados (intestino e gônadas), os mosquitos foram imobilizados no freezer e transferidos para uma placa de petri no gelo. A dissecção foi realizada em estereomicroscópio com estiletes entomológicos em solução de PBS (tampão fosfato de sódio 10 mM, NaCl 0,15 M, pH 7.4). O tubo digestivo (intestino anterior, médio e posterior), órgãos anexos (divertículos, túbulos de Malphigi) e órgãos reprodutores eram obtidos fixando o terceiro segmento abdominal com o estilete e puxando o tórax na direção oposta. Os órgãos de interesse, intestino médio e as gônadas, eram removidos e lavados na solução de PBS, e por fim, colocadas no tampão de lise nos eppendorffs devidamente etiquetados.

Enzimografia

Extração de proteínas

Para a realização dos ensaios enzimográficos em gel e em solução, foram extraídos 20 intestinos de fêmeas e machos. Os órgãos foram lisados utilizando um pistilo plástico acoplado em um homogeneizador mecânico. O tampão de lise usado era composto de 100 mM Tris-HCl pH 8,0, 150 mM NaCl, 10% Glicerol, 0,6% Triton X-100 (Galán et al., 1992). O homogenato foi centrifugado à 14.000 x g por 15 min a 4 °C para remover o material insolúvel e o sobrenadante era colocado em eppendorffs devidamente etiquetados. Em seguida, as proteínas foram quantificadas usando o *Qubit*®. Os homogeneizados foram aliquoteados e mantidos à -20°C até o uso.

Eletroforese

Os extratos protéicos foram diluídos em tampão de amostra (125 mM Tris-HCl, pH 6,8; 4% SDS; 20% glicerol e 0,002% azul de bromofenol) e submetidos à separação por eletroforese, a 4 °C e 110 V constantes, em géis de poliacrilamida com dodecil sulfato de sódio (SDS-PAGE) 10% co-polimerizados com 0,1% de gelatina de porco (Lockwood et al., 1987). Foram aplicados equivalentes a 1/10 de intestino por poço em cada gel. Após a eletroforese os géis foram lavados duas vezes por 30 min a 4° C em tampão de acetato de sódio 0,1 M contendo 2,5% Triton X-100 (pH 3,5 e 5,5) ou em tampão de Tris-HCl 0,1 M contendo 2,5% Triton X-100 (pH 7,5 e 10,0) para remoção do SDS e renaturação das peptidases. Posteriormente, os géis eram incubados em tampão de reação acetato de sódio 0,1 M (pH 3,5 e 5,5), ou em tampão Tris-HCl 0,1 M (pH 7,5 e 10,0) com intuito de observar as atividades proteolíticas. As bandas de atividade proteolítica, resultantes da degradação do substrato foram visualizadas após a coloração dos géis com Azul de Coomassie (0,25% azul de Coomassie R-250, 40% metanol e 10% de ácido acético) e posterior descoloração com solução de ácido acético 10%. O peso molecular das peptidases foi calculado por comparação com peso molecular comercial (PageRuler®).

Para determinar a classe de peptidase responsável pela atividade proteolítica foram usados distintos inibidores de peptidases. Soluções estoque dos inibidores 1,10-fenantrolina (200 mM) e pepstatina A (1 mg/ml) foram preparadas em etanol, enquanto que *Trans-epoxysuccinyl L-leucylamido-(4-guanidino) butano* (E-64, 10 μ M) foi solubilizado em água; fluoreto de fenil-metil-sufonil (*Phenyl-methyl sulfonyl-fluoride*, PMSF, 250 mM) em isopropanol; N α -Tosil-L-lisina clorometil hidrocloreto cetona (*N- α -Tosyl-L-lysine chloromethyl ketone hydrochloride*, TLCK, 100 mM) e N-p-Tosyl-L-fenilalanina clorometil cetona (*N-p-Tosyl-L-phenylalanine chloromethyl ketone*, TPCK, 100 mM) em metanol. Todas as soluções foram armazenadas a -20°C.

Extratos proteicos obtidos como anteriormente descrito, foram pré-incubados por 30 minutos à 4°C, antes de serem aplicados no gel, com um dos seguintes inibidores de peptidases: 10 μ M E-64, um inibidor irreversível de cisteína-peptidases; 1 mM PMSF, inibidor irreversível de serino-peptidases; 100 μ M TLCK; 100 μ M TPCK; 10 μ M pepstatina-A, um inibidor reversível de aspártico-peptidases e 10 mM 1,10-fenantrolina. A amostra controle foi submetida às mesmas condições, porém na ausência de inibidores.

Todos os resultados são derivados de pelo menos cinco experimentos independentes realizados em triplicata (cinco replicatas biológicas e três replicatas experimentais). As imagens dos géis foram obtidas utilizando um densitômetro GS-800™ (Bio-Rad).

Ensaio enzimático em solução

A determinação fluorimétrica da atividade de tripsina nas preparações dos homogeneizados foi realizada com a utilização de substrato cumarínico: carbobenzoxi-phearg7-amino-4-metil coumarina (Z-FR-MCA-Sigma) na concentração de 100 μ M em tampão de reação acetato de sódio 0,1 M (pH 3,5 e 5,5), ou em tampão Tris-HCl 0,1 M (pH 7,5 e 10,0). Metil coumarina é liberada quando o substrato é hidrolisado e, desta forma, pode-se detectá-la através de sua fluorescência, a qual é medida com excitação a 380 nm e emissão a 460 nm (Alves et al., 1996). A detecção dos produtos de hidrólise foi realizada em um espectrofluorímetro (SpectraMax Gemini xps, Molecular Devices, CA, USA) termo-

estabilizado a 37°C. Foram aplicados equivalentes a 1/10 de intestino (aproximadamente 10 µg de proteína por poço) nos ensaios de *Cx. quinquefasciatus* e *Ae. aegypti* alimentados com açúcar. Para os ensaios de amostras provenientes de intestino de fêmeas de *An. albitarsis* s.s e *Ae. aegypti* alimentadas com sangue as amostras precisaram ser diluídas, pois a alta atividade consumia todo o substrato antes da leitura ser iniciada pelo espectrofluorímetro. Nesses casos, equivalentes a 1/100 de intestino foram utilizados na análise.

O substrato foi preparado na concentração estoque de 3 mM em 100% de dimetilsulfóxido (DMSO) e antes de cada ensaio o substrato era diluído em água mili-Q a fim de ficar em uma concentração de 100 µM.

Para os ensaios em solução ensaios em solução foram utilizados os seguintes inibidores: 10 µM E-64, 1 mM PMSF, 100 µM TLCK; 100 µM TPCK. Neste caso, as amostras eram pré-incubadas na presença do inibidor por 15 min. à temperatura ambiente. Após a incubação, foi adicionado o substrato e a leitura foi feita no espectrofluorímetro após 40 minutos de reação. A atividade foi interrompida pela adição de ácido acético 30%. O controle foi tratado da mesma forma, porém na ausência dos inibidores. Os experimentos foram realizados em triplicatas experimentais.

Espectrometria de massas

Extração de proteínas de intestino de fêmeas para LC-MS/MS

Para cada espécie estudada, 100 intestinos médio de fêmeas alimentadas com açúcar foram lizados em 200 µL de (4% SDS, 20% glicerol, 10% mercaptoetanol, 0.004% de azul de bromofenol e 0.125M Tris-HCl, pH 6.8). Para remover o material insolúvel, foram utilizadas duas centrifugações à 14.000 x g por 10 minutos à 4°C e o sobrenadante coletado. As amostras foram dosadas utilizando o reagente Pierce 660 nm (Thermo Scientific). Posteriormente, as amostras foram aquecidas por 5 minutos em banho maria e separadas por SDS-PAGE 12%, acrilamida 30% e bis-acrilamida 0,8%. As proteínas foram coradas com azul de Coomassie e fotodocumentadas. Foram feitos três géis com amostras provenientes de três lises independentes. As

proteínas foram digeridas seguindo protocolos previamente estabelecidos (González-Caballero et al., 2014) com algumas modificações. Brevemente, os pedaços dos géis foram manualmente excisados e descorados com 400 μL de acetomitrila 50% e 25 mM NH_4HCO_3 pH 8.0 por 15 minutos. As proteínas foram reduzidas e alquiladas utilizando 65mM de DTT e 200 mM de iodoacetamida, respectivamente. Os pedaços dos géis foram lavados com 100 mM de NH_4HCO_3 , seguido de desidratação por acetomitrila. Os pedaços de géis foram incubados por 24 horas a 37°C para reidratação e digestão em solução contendo 20 ng/ μL de tripsina de porco modificada (Promega, USA) e 50mM de NH_4HCO_3 . Após a digestão tríptica, os peptídeos foram extraídos utilizando ácido fórmico 0,1% e acetomitrila 50% v/v. Os peptídeos foram purificados e concentrados com resina POROS oligo R3 C18 (Applied Biosystems, USA). Os peptídeos eluídos foram analisados em nano-HPLC acoplados em linha com espectrômetro híbrido linear quadripolo (LTQ) Orbitrap.

Análise no espectrômetro de massas

Para cada amostra, 4 μL da solução contendo os peptídeos eluídos (0,1% ácido fórmico) foram aplicados no sistema de nano-HPLC acoplado em linha a um espectrômetro de massas de alta resolução LTQ Orbitrap Velos, equipado com interface de ionização por nanoelectrospray - nanoESI (Thermo Fisher Scientific). A coluna para eluição (150 $\mu\text{m} \times 2$ cm) foi empacotada in-house com resina ReproSil 5 μm (Dr. Maisch) e a coluna analítica (100 $\mu\text{m} \times 15$ cm) também foi empacotada in-house com resina ReproSil 3 μm (Dr. Maisch). Foi utilizada como fase móvel A solução de ácido fórmico 0.1 % (v/v) em água e como fase móvel B, ácido fórmico 0.1 % (v/v) em acetomitrila. Os espectros de massas foram adquiridos no modo positivo utilizando busca de MS dependente de dados (DDA). Cada DDA consiste em um escaneamento de busca em uma faixa de 300-2000 m/z e resolução de 60000 com valor alvo de 1×10^{-6} ions. Cada escaneamento de busca foi composto dos 10 íons mais intensos detectados após a dissociação induzida por colisão (CID).

Análise dos dados

Para *An. albitarsis* s.s e *Cx. quiquefasciatus*, a abordagem utilizada foi a de *peptide spectrum match* – PSM. Dois diferentes motores de busca foram utilizadas para a identificação de proteínas, o MASCOT, com posterior filtragem e validação dos peptídeos identificados pelo Scaffold, bem como a plataforma PATERNLAB, na qual foi utilizado o software COMET para identificação e posteriormente o software SePro para filtragem e validação dos dados. Os dados em formato (.RAW) foram exportados para o formato dos programas MASCOT (.MGF) e COMET (.MS2), respectivamente, utilizados para a identificação das proteínas. Os espectros foram analisados contra um banco de dados de culicídeos obtido do UNIPROT (download em maio de 2015), contendo, aproximadamente, 101,000 sequencias. As proteínas identificadas foram classificadas de acordo com as categorias de ontologia gênica (Gene Ontology, GO). As principais categorias do GO são: (1) Componente celular; (2) Processo Biológico e (3) Função Molecular. Dentro de cada uma destas categorias, há uma estrutura de sub-categorias mais específicas. As proteínas foram adicionalmente categorizadas pelo Gene Ontology Explorer – GOEx, que é um módulo do PatternLab for Proteomics (<http://pcarvalho.com/patternlab/>) e pela ferramenta BioMart do VectorBase.

Análises de alinhamento múltiplo e bioinformática

As sequências FASTA das serina peptidases do tipo tripsina identificadas por espectrometria de massas foram inteiramente retiradas do banco de dados do VectorBase (<http://biomart.vectorbase.org>). Os alinhamentos múltiplos foram realizados pelo CLUSTAL Omega na configuração padrão. As sequencias FASTA das tripsinas aqui identificadas foram alinhadas contra sequencias bem anotadas no UNIPROT como o quimiotripsinogenio bovino (CTRA_BOVIN), tripsinogênio bovino (TRY1_BOVIN), Tripsina 3A1 de *Ae. aegypti* (TRY3_AEDAE) e Tripsina-6 de *An. gambiae* (TRY6_ANOGA), as quais auxiliam na observação dos resíduos no sítio ativo, resíduos de cisteína envolvidas nas pontes dissulfeto, e demais regiões analisadas.

As sequências de aminoácidos de cada tripsina identificada foram analisadas em busca das características que as qualificam como tripsinas digestivas de invertebrados. Com auxílio da ferramenta PROSCAN disponível no PROSITE (<http://prosite.expasy.org>), foram analisados os domínios e motivos conservados destas enzimas, os resíduos do sítio ativo (His, Asp, Ser), o peptídeo sinal, os resíduos de cisteína das pontes dissulfeto e o tamanho das proteínas inativas e maduras. O peptídeo sinal foi também verificado pelo SignalP 4.0 (<http://cbs.dtu.dk/services/SignalP>). Os sítios de glicosilações do tipo N- e O-glicosilações nas sequências de aminoácidos foram preditas pelo NetNGlyc1.0 (<http://cbs.dtu.dk/services/NetNGlyc>) e NetOGlyc (<http://cbs.dtu.dk/services/NetOGlyc>), respectivamente. Para identificar peptídeos espécies específicos foi utilizado o Skyline software (<http://proteome.gs.washington.edu/software/skyline>), com o mesmo banco de dados de Culicidae utilizado para a identificação das proteínas por espectrometria de massas.

4. RESULTADOS

Atividade proteolítica em intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar

Curso temporal

Para verificar a influência do tempo de reação sobre as atividades proteolíticas de intestino de fêmeas de *Cx. quinquefasciatus*, os géis enzimográficos copolimerizados com gelatina foram incubados por 2, 4, 6 e 12 horas em tampão de reação pH 7,5 a 37 °C (Figura 8). A intensidade das atividades proteolíticas aumentou progressivamente até 12 horas. A partir desse tempo a intensa atividade enzimática levou a uma sobreposição das bandas não sendo possível resolver o perfil. Em 12 horas de reação, foi observado um perfil proteolítico composto de 9 bandas principais variando entre 17 e 150 kDa, aproximadamente. Assim, o tempo de reação enzimática em 12 horas foi escolhido para os ensaios subsequentes das amostras de intestino de *Cx. quinquefasciatus* fêmeas.

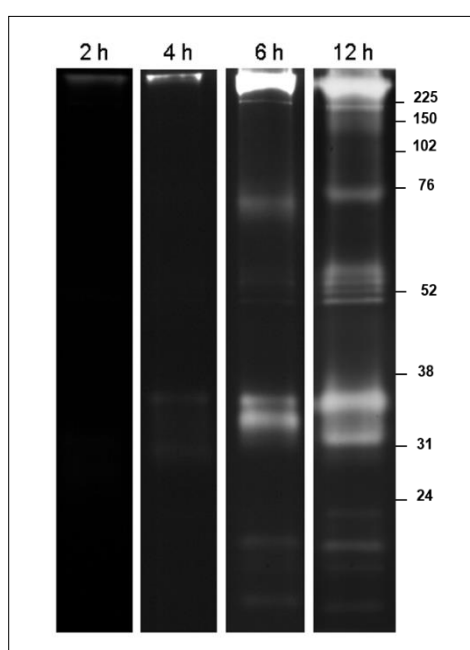


Figura 8. Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar. Atividades proteolíticas foram avaliadas após 2, 4, 6 e 12 horas de incubação em tampão de Tris-HCl 0.1 M, pH de 7,5. Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Influência do pH

Para determinar a influência exercida pelo pH sobre as atividades proteolíticas foram utilizadas as técnicas de enzimografia em gel copolimerizado com gelatina e ensaios em solução com substrato fluorogênico Z-phe-arg-MCA. Os géis enzimográficos foram incubados por 12 horas em tampões de reação de acetato de sódio 0,1M (pH 3,5; 5,5) ou Tris-HCl (pH 7,5 e 10). As atividades enzimáticas foram detectadas em todos os pHs testados, aumentando progressivamente até pH 10 (Figura 9. A). Em pH 3,5 e 5,5, um número reduzido de bandas proteolíticas foram observadas, comparadas com os perfís obtidos a pH 7,5 e 10. Em pH 10 a atividade foi mais intensa que em pH 7,5, sendo observado um perfil de bandas migrando entre 17 e 150 kDa. Entretanto, não foram observadas diferenças quantitativas entre pH 7,5 e 10. Dessa forma, a condição experimental de pH 7,5 foi a utilizada nos ensaios com inibidores.

Utilizando o substrato fluorogênico Z-phe-arg-MCA observou-se que a atividade enzimática aumenta progressivamente em função do pH (Figura 9. B). Amostras equivalentes a 1/10 de intestino foram ensaiadas em tampão acetato de sódio (pH 3,5 e 5,5) e Tris-HCl (pH 7,5 e 10), contendo 100 μ M do substrato. Em pH 10 a cinética de hidrólise do substrato atingiu o platô em menos de 15 minutos. Em pH 7,5 a atividade atingiu o platô com aproximadamente 30 min. de reação.

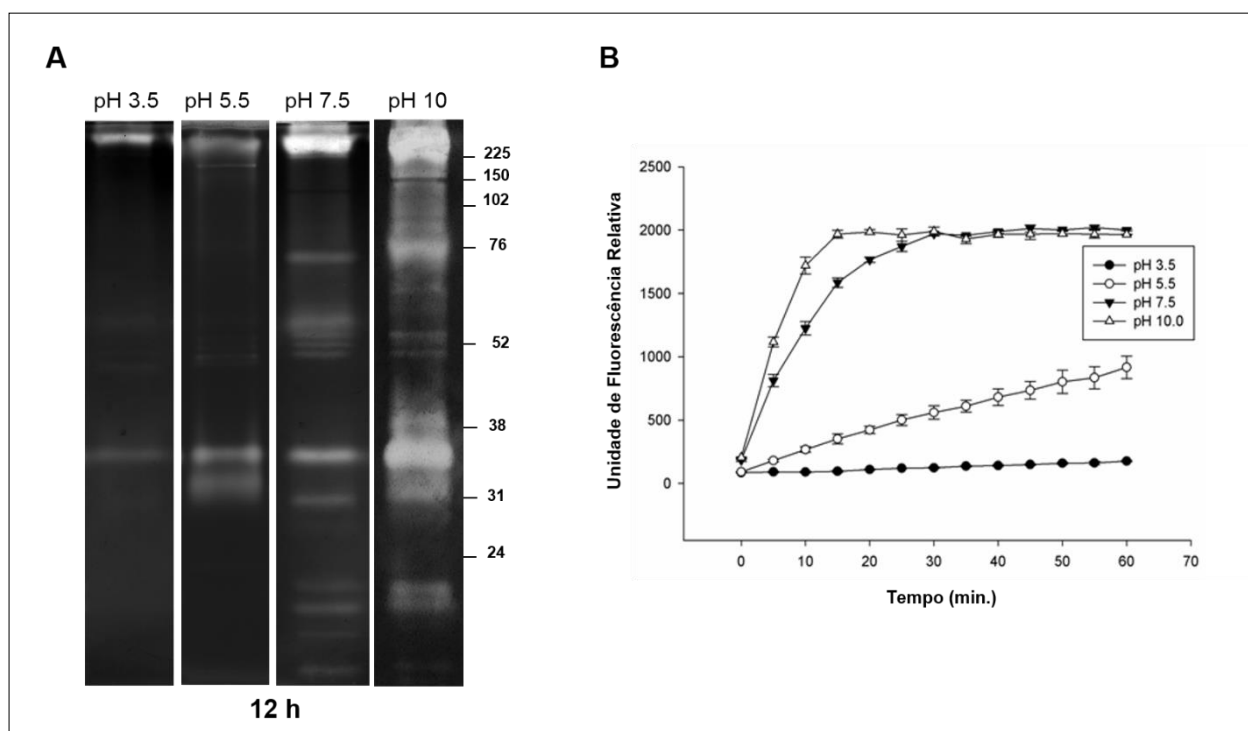


Figura 9. Efeito do pH sobre as atividades proteolíticas de extratos de intestinos de fêmeas de *Cx. quinquesfasciatus* alimentadas com açúcar. **(A).** A influência do pH foi avaliada através da incubação dos géis enzimográficos à 37°C por 12 horas em tampão de acetato de sódio (pH 3,5 e 5,5) ou com tampão Tris-HCl (pH 7,5 e 10,0). Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B).** Os ensaios em solução foram realizados com 100 μ M do substrato fluorogênico Z-Phe-Arg-MCA em tampão de acetato de sódio (pH 3,5 e 5,5) ou em tampão Tris-HCl (pH 7,5 e 10,0).

Efeito dos inibidores

A atividade enzimática exibida pelo intestino de fêmeas foi fortemente inibida por 1 mM PMSF ou 100 μ M TLCK, tanto na enzimografia em gel quanto em solução. Amostras equivalentes a 1/10 de intestino foram aplicados nos géis e incubados por 12 horas em tampão de reação Tris-HCl pH 7,5 (Figura 10). No entanto, as atividades proteolíticas não foram afetadas por 100 μ M TPCK. A atividade também não foi afetada pelos inibidores 10 μ M E-64, 10 μ M pepstatina A, e 10 mM 1,10-fenantrolina (dados não apresentados).

Para os ensaios em solução a amostra foi incubada em tampão Tris-HCl pH 7,5 contendo 100 μ M do substrato, na ausência (controle) ou presença de inibidores 1 mM PMSF, 100 μ M TLCK, 100 μ M TPCK, 10 μ M E64. Observou-se que tanto o PMSF

quanto o TLCK inibem a atividade das peptidases presentes no intestino, corroborando o observado nos ensaios de enzimografia em gel. Em conjunto, os ensaios indicam que o perfil proteolítico é composto por serina peptidases do tipo tripsina.

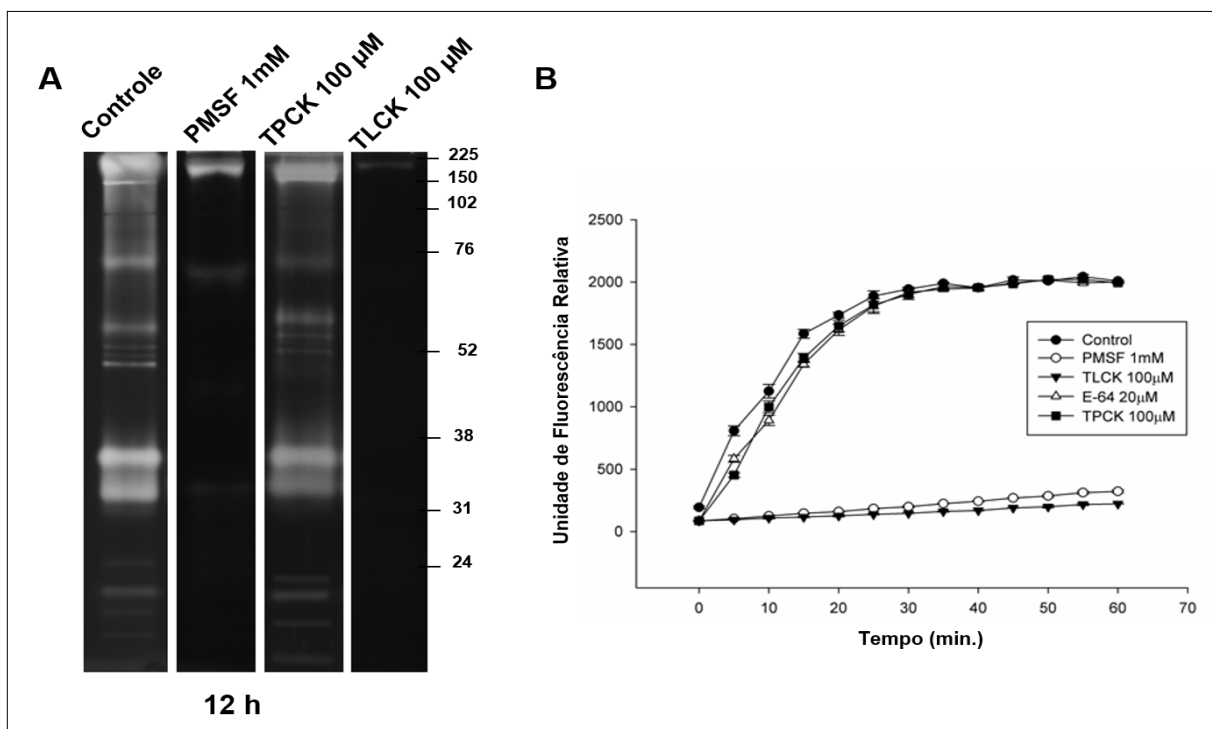


Figura 10. Efeito dos inibidores de peptidases no perfil proteolítico de extrato do intestino médio de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar. **(A).** Os géis enzimográficos foram pré-incubados por 30 min na presença de cada um dos inibidores: 1 mM de PMSF; 100 µM de TPCK e 100 µM de TLCK. As atividades proteolíticas foram detectadas após incubação dos géis por 12 horas à 37°C em tampão Tris HCl (pH 7,5). O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B).** Os ensaios em solução foram realizados com 100 µM substrato fluorogênico Z-Phe-Arg-MCA em tampão de Tris-HCl 100 mM (pH 7,5) na ausência (controle) ou na presença de 1 mM de PMSF, 100 µM de TLCK, 20 µM E-64 ou 100 µM de TPCK.

Atividades proteolíticas em intestinos de machos de *Cx. quinquefasciatus* alimentados com açúcar

Os mesmos procedimentos de caracterização bioquímica utilizados para as análises enzimáticas de amostras extraídas de intestinos de fêmeas alimentadas com açúcar foram empregados para as análises de amostras extraídas de machos. Amostras equivalentes a 1/10 de intestino foram aplicados nos géis, os quais foram incubados por 24 horas em tampão de reação pH 7,5. As atividades proteolíticas das amostras extraídas dos machos foram consideravelmente menos intensas que das fêmeas, necessitando de um tempo de 24 horas de incubação dos géis para que o perfil enzimático fosse revelado (Figura 11). Todas as bandas de atividade observadas no macho foram detectadas nas fêmeas. Entretanto o perfil proteolítico de machos foi menos complexo em número de bandas. A atividade proteolítica foi mais alta em pH alcalino e foi fortemente inibida por 1 mM PMSF e 100 μ M TLCK. No entanto, as atividades proteolíticas não foram afetadas por 100 μ M TPCK. O efeito dos inibidores sobre a atividade proteolítica indica que essas enzimas são do tipo tripsinas.

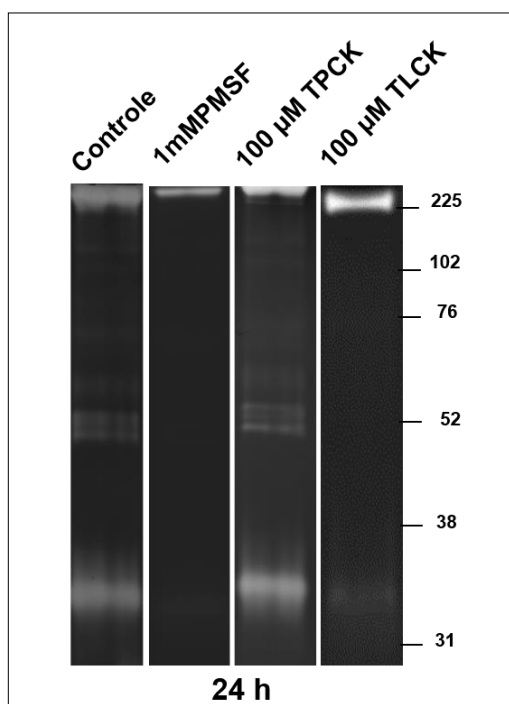


Figura 11. Efeito dos inibidores de peptidases no perfil proteolítico de intestino médio de machos de *Cx. quinquefasciatus* alimentados com açúcar. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μ M de TPCK e 100 μ M de TLCK. As atividades proteolíticas foram detectadas após incubação dos géis por 24 horas à 37°C em tampão Tris HCl com pH 7,5. O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Peptidases identificadas por espectrometria de massas

Com o intuito de identificar as peptidases expressas no intestino de fêmeas alimentadas com açúcar, após a coloração do gel de SDS-PAGE e visualização das bandas de proteínas, o gel foi dividido em 24 partes, as quais foram digeridas com tripsina e analisadas em espectrômetro de massas. Sete serina peptidases do tipo tripsina foram identificadas no intestino de fêmeas alimentadas com açúcar utilizando duas engenharias de buscas independentes: Mascot (seguida da validação pelo Scaffold) e o COMET (seguida de validação pelo SEPro) (Tabela 1). Adicionalmente, uma tripsina peptidase foi exclusivamente identificada pelo Mascot (seguida da validação pelo Scaffold) com um peptídeo único e um espectro (Tabela 1, B0WW44, em cinza). Interessantemente, as bandas de SDS-PAGE, onde as peptidases foram identificadas por MS/MS, coincide, em parte, com as regiões onde observamos atividade nos géis de enzimografia (Figura 12). Apesar da maioria das peptidases identificadas estares na região de 25 a 40 kDa, duas tripsinas, a Trypsin5 e Trypsin7 foram as únicas identificadas por LC-MS/MS nas regiões de alta massa molecular (Figura 12).

Tabela 1. Serina peptidases do tipo tripsina, identificadas por LC-MS/MS, em intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar

Proteínas Identificadas	Identificador Uniprot	Massa Molecular	Mascot Peptídeos Únicos	Mascot Espectro Total	Mascot Cobertura %	Peptídeos Identificados pelo MASCOT	Mascot Ion score	Sequências identificadas pelo COMET	COMET Peptídeos Únicos	COMET Espectro Total	COMET Cobertura %
Trypsin 4 OS=Culex quinquefasciatus GN=CpipJ_CPIJ017414	B0XCW2_CULQU	28 kDa	3	13	18	(R)VGSSYDYQGGTVIDVAGMTIHPR(Y) (K)DFDFALLR(L) (K)GCAQPDYGVYADVEK(A)	35.21 52.94 39.12	(R)VGSSYDYQGGTVIDVAGMTIHPR(Y) (K)DFDFALLR(L) (K)GCAQPDYGVYADVEK(A) (K)NMLCAGYDEGLR(D) (R)LSWIGVR(V) (R)ENYAESR(L)	6	29	28
Trypsin 7 OS=Culex quinquefasciatus GN=CpipJ_CPIJ017964	B0XES8_CULQU	27 kDa	4	14	18	(R)GGQLIIVTR(K) (R)DYALLNLAK(S) (R)AVDVPIADHDR(C) (K)DACLGDSGGPLTCSGK(V)	53.31 50.34 24.69 49.46	(R)GGQLIIVTR(K) (R)DYALLNLAK(S) (R)AVDVPIADHDR(C) (K)DACLGDSGGPLTCSGK(V) (F)M*LCAGYDAGGK(D)	5	25	18
Trypsin-5 OS=Culex quinquefasciatus GN=CpipJ_CPIJ015103	B0X667_CULQU	30 kDa	3	7	16	(K)IIGGFPAEQGDTLHQVSIR(F) (K)GCGLAAYPGIYSDVAYYR(G) (R)GWIDSCLAGK(C)	35.64 29.61 31.7	(K)IIGGFPAEQGDTLHQVSIR(F) (K)GCGLAAYPGIYSDVAYYR(G) (R)GWIDSCLAGK(C)	3	14	16
Trypsin-1 OS=Culex quinquefasciatus GN=CpipJ_CPIJ007079	B0WIS4_CULQU	29 kDa	3	6	18	(R)IVGGFEISIADAPHQVLSQSR(G) (K)HASGGSVISIK(R) (R)AAYVPAYNQNCNSAYAR(Y) (K)DACQGGDSGGPLVADGK(L)	51.51 26.21 29.15 33.42	(R)IVGGFEISIADAPHQVLSQSR(G) (K)HASGGSVISIK(R) (R)AAYVPAYNQNCNSAYAR(Y) (R)NTIDYDYSLELK(S) (R)GSHICGGSII SPK(W) (K)WILTAHCTDGASVSNLR(I)	6	10	34
Trypsin 2 OS=Culex quinquefasciatus GN=CpipJ_CPIJ005273	B0WE94_CULQU	28 kDa	2	2	12	(R)LEFGHAVQPVDLVR(D) (R)DEPADESQSLVSGWGDTR(S)	19.14 27.7	(R)LEFGHAVQPVDLVR(D) (R)DEPADESQSLVSGWGDTR(S) (R)WVLTAAHCTENTDAGIYSVR(V) (R)GVLVPLVNR(E) (K)LGMPVTESMICAGFAK(E)	5	13	30
Serine protease1/2 OS=Culex quinquefasciatus GN=CpipJ_CPIJ003826	B0W9S9_CULQU	30 kDa	2	3	18	(R)TGTFVDNQATVSGFGR(T) (R)TVDGGPVSPTK(N)	35.91 35.12	(R)TGTFVDNQATVSGFGR(T) (R)TVDGGPVSPTK(N)	2	3	10
Serine protease SP24D OS=Culex quinquefasciatus GN=CpipJ_CPIJ0151	B0X870_CULQU	27 kDa	1	1	10	(K)LGESIEDELSQPIALYEGDDLPK(D)	34.98	(K)LGESIEDELSQPIALYEGDDLPK(D)	1	3	10
Cationic trypsin OS=Culex quinquefasciatus GN=CpipJ_CPIJ011378	B0WW44_CULQU	26 kDa	1	1	8	(R)IIVHPQYAEGLANDIAVIR(V)	32.92				

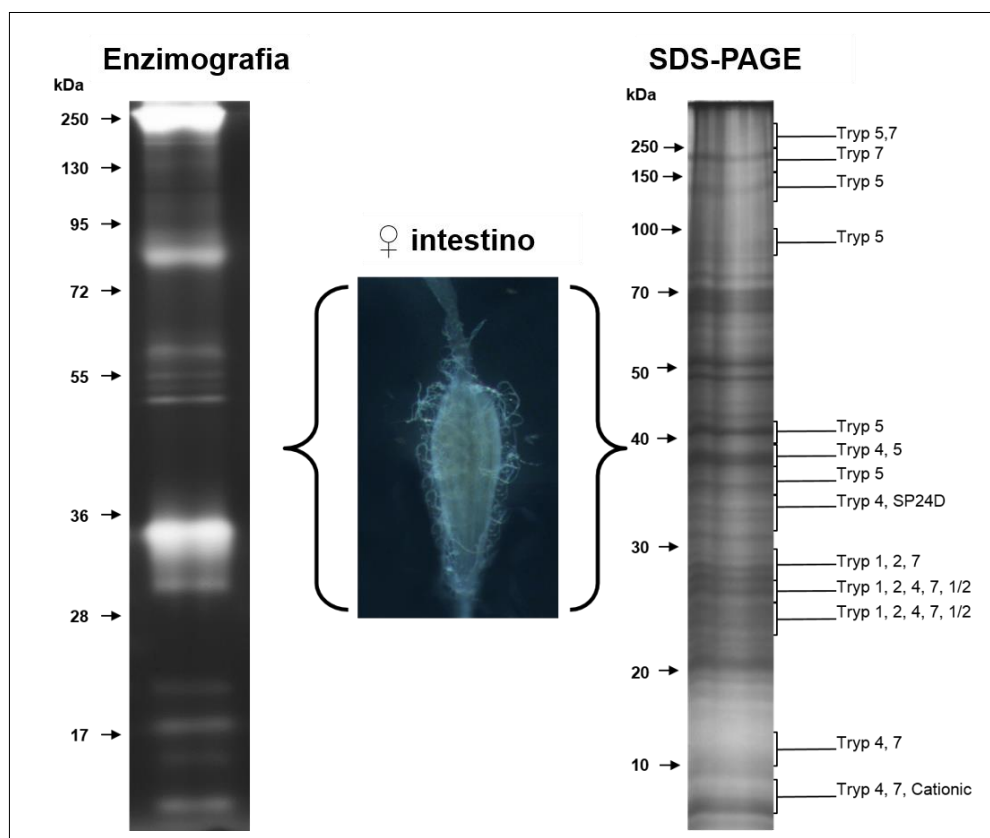


Figura 12. Perfil enzimográfico representativo e SDS-PAGE das proteínas totais solúveis de intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar. A figura mostra dois diferentes sistemas eletroforéticos utilizados para caracterizar e identificar, respectivamente, as serina peptidases do tipo tripsina: na enzimografia as amostras são separadas em condições não redutoras para posterior detecção das atividades proteolíticas, já na separação por SDS-PAGE as amostras são reduzidas. As bandas de proteínas do gel SDS-PAGE foram retiradas para a identificação por LC-MS/MS. Os números a esquerda dos dois géis indicam o padrão da massa molecular utilizado (kDa). A imagem do intestino foi obtida em um microscópio óptico com interferência de contraste (DIC) e com aumento de 40x.

Análise de bioinformática e alinhamento das tripsinas identificadas

As sequências FASTA das tripsinas identificadas dos extratos de intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar foram analisadas por alinhamentos múltiplos, a fim de melhor explorar as características das sequências destas enzimas. Para localizar os resíduos do sítio ativo, pontes dissulfeto, resíduos de especificidade ao substrato e demais motivos que caracterizam as tripsinas

digestivas, sequências FASTA de serina peptidases (TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA), que estão bem caracterizadas e anotadas nos bancos de dados foram utilizadas nos alinhamentos múltiplos. Não obstante as enzimas identificadas tenham apresentado similaridades em algumas regiões das sequências, observou-se que elas apresentam sequências motivos únicos, isto é, algumas das tripsinas peptidases identificadas apresentaram características distintivas entre elas. Por exemplo, o motivo auto catalítico da Trypsin4 possui um resíduo de histidina no lugar dos resíduos R/K; os motivos de ativação na Trypsin5, IIGG, e da Cationic trypsin, VVGG, diferem por um aminoácido do clássico motivo IVGG. Adicionalmente, foi observado nas peptidases identificadas características típicas de peptidases digestivas de invertebrados como: (i) os resíduos conservados de histidina, ácido aspártico e serina formando a tríade catalítica; (ii) resíduos de cisteínas responsáveis por formar as pontes dissulfeto; (iii) presença do peptídeo sinal; (iv) os putativos motivos de ativação auto catalítica imediatamente após um resíduo de arginina ou lisina (R/K-IVGG); (v) os motivos característicos de peptidases ativas LTHAAC, DIAL e GDSGGP (Tabela 2; Figura 13).

Tabela 2. Características gerais das serina peptidases do tipo tripsina identificadas em *Cx. quinquefasciatus* Δ

Identificador Uniprot	Proteína	Posição dos Resíduos do Sítio Ativo	Posição dos Resíduos dos Pares de Cisteína	Posição dos Resíduos de Especificidade ao Substrato	Tamanho da Proteína (aa)		Sítio de Ativação	Regiões conservadas		
					Precursor	Maturo		LTAAHC	DIAL	GDSGGP
BOWIS4	Trypsin 1	His88, Asp133, Ser229	73-89, 198-214, 225-249	Asp223, Gly246, Gly256	274	226	YR^IVGG	LTAAHC	DYSL	GDSGGP
BOWE94	Trypsin 2	His75, Asp120, Ser216	60-76, 183-200, 212-236	Asp210, Gly234, Gly244	261	226	GK^IVGG	LTAAHC	DFCL	GDSGGP
BOXCW2	Trypsin 4	His70, Asp116, Ser213	151-219, 181-198, 209-233	Asp207, Gly230, Gly240	258	233	FH^IVNG	LTA AHL	DFAL	GDSGGP
BOX667	Trypsin 5	His73, Asp127, Ser221	58-74, 160-227, 192-208, 217-241	Asp215, Gly238, Gly248	293	268	PK^IIGG	LTA AHC	DIAL	GDSGGP
BOXES8	Trypsin 7	His67, Asp112, Ser207	52-68, 146-213, 176-192, 203-227	Asp201, Gly224, Gly234	252	229	SR^IVNG	LTA GHC	DYAL	GDSGGP
BOX870	SP24D	His63, Asp109, Ser195	48-64, 172-181, 191-216	?	240	217	RR^IFGG	LTA AHC	DIAL	GDSGGP
BOW9S9	Serine protease 1/2	His92, Asp135, Ser233	77-93, 202-217, 229-259	?	283	235	SR^IVNG	LTA AHC	DIGL	GDSGGP
B0WW44	Cationic trypsin	His69, Asp114, Ser202	54-70	?	244	216	GR^VVGG	LTA GHC	DI AV	YDGGSP

Δ Extraídos após o alinhamento com CLUSTAL Omega.

(aa) número total de resíduos de aminoácidos.

? = outros resíduos diferentes dos DGG.

Table with 4 columns: Identifier, Line Number, Sequence, and Score. Rows include BOWIS4_CULQU, BOWE94_CULQU, BOXCW2_CULQU, BOX667_CULQU, BOXES8_CULQU, BOX870_CULQU, BOWW44_CULQU, BOW9S9_CULQU, TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA.

o *o

Table with 4 columns: Identifier, Line Number, Sequence, and Score. Rows include BOWIS4_CULQU, BOWE94_CULQU, BOXCW2_CULQU, BOX667_CULQU, BOXES8_CULQU, BOX870_CULQU, BOWW44_CULQU, BOW9S9_CULQU, TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA.

*

Table with 4 columns: Identifier, Line Number, Sequence, and Score. Rows include BOWIS4_CULQU, BOWE94_CULQU, BOXCW2_CULQU, BOX667_CULQU, BOXES8_CULQU, BOX870_CULQU, BOWW44_CULQU, BOW9S9_CULQU, TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA.

o

Table with 4 columns: Identifier, Line Number, Sequence, and Score. Rows include BOWIS4_CULQU, BOWE94_CULQU, BOXCW2_CULQU, BOX667_CULQU, BOXES8_CULQU, BOX870_CULQU, BOWW44_CULQU, BOW9S9_CULQU, TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA.

o \$ o #* \$ o \$

Table with 4 columns: Identifier, Line Number, Sequence, and Score. Rows include BOWIS4_CULQU, BOWE94_CULQU, BOXCW2_CULQU, BOX667_CULQU, BOXES8_CULQU, BOX870_CULQU, BOWW44_CULQU, BOW9S9_CULQU, TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA.

Figura 13. Alinhamento de sequências de tripsina em fêmeas de *Cx. quinquefasciatus* identificadas por LC-MS/MS com sequências tripsina e quimotripsina bem anotados (TRY1_BOVIN, CTRA_BOVIN, TRY3_AEDAE, TRY6_ANOGA). Regiões de importância são representados da seguinte forma: péptido sinal (Cinza); resíduos N-terminais da enzima ativa (itálico e negrito); pontes dissulfeto conservado de cisteína (O); tríade catalítica conservada (*); resíduos acessórios catalíticos (§); Asp 194 (baseado em α -quimotripsinogénio bovina) altamente conservada (#); locais de glicosilação (sublinhado e negrito).

O alinhamento das tripsinas e análise pelo InterproScan, evidenciou que cinco delas, Trypsin1, Trypsin2, Trypsin4, Serine protease SP24D e a Serina protease ½, possuem três pontes dissulfeto, ao passo que a Trypsin5 e Trypsin7 possuem quatro pontes dissulfeto e a Cationic trypsin apenas uma. Análises de predição da localização celular utilizando o Target P indicam que todas as tripsinas identificadas são secretadas devido a presença do peptídeo sinal característico. A análise de predição de hélices trans- membranas, utilizando o software TMHMM, mostrou que a Trypsin4 e Trypsin5 possuem um domínio trans-membrana (Tabela 3). Cinco das oito tripsinas identificadas apresentaram sítios preditos para O- e N- glicosilações (Figura 13, Tabela 3).

O VectorBase foi usado para analisar a estrutura dos genes que codificam para as serina peptidases do tipo tripsina. Observou-se que o número de exons das tripsinas variam de um a três. Com exceção da Serine protease ½, que possui um intron com 298 nucleotídeos, as demais peptidases apresentam introns com 25 a 71 nucleotídeos (Tabela 3). Além disso, foi verificado que essas enzimas possuem um número de parálogos que variam de 13 a 38 (Tabela 3), sendo os genes da Trypsin1, 2, 4, 7, SP24D e Cationic trypsin, parálogos entre si. Outra observação, quando exploramos as informações disponíveis no VectorBase, foi a de que vários genes de tripsinas se encontram em clusters. Por exemplo, vizinhos ao gene que codifica para a Trypsin1, encontram-se cinco outros genes que codificam para tripsinas (Figura 14).

O software SKYLINE foi usado para confrontar as sequências das tripsinas identificadas contra as sequências de todas as proteínas disponíveis no banco de dados de cuicídeos, a fim de identificar peptídeos proteotípicos nas tripsinas. Esta abordagem foi utilizada com intuito de padronizar a técnica de Monitoração de Reação Seleccionada (Selected Reaction Monitoring- SRM), para monitorar um peptídeo proteotípico alvo de

tripsinas em amostras complexas de proteínas durante análise por espectrometria de massas. Apesar da existência de diversos motivos conservados nas tripsinas peptidases, a análise no SKYLINE revelou que cinco das oito tripsinas identificadas por espectrometria de massas possuem diferenças em suas sequencias que permitem identificação de peptídeos únicos. Estes peptídeos únicos foram também os mesmos identificados por MS/MS (Tabela 4).

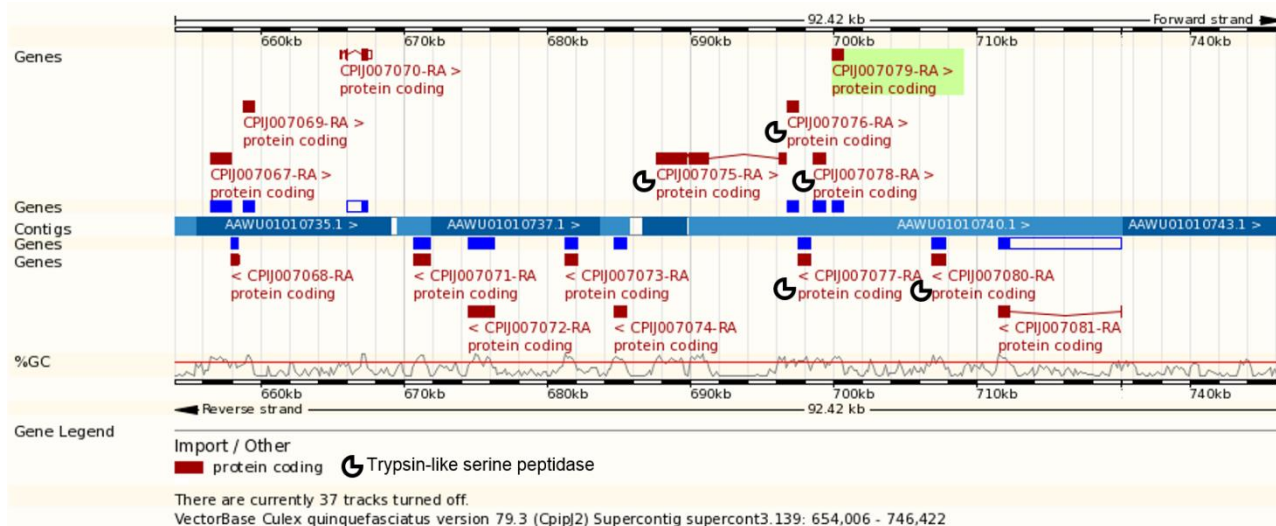


Figura 14. Localização do gene da Tripsina1 de *Cx. quinquefasciatus* em seu supercontig. Genes de outras tripsinas são encontrados no mesmo supercontig (), próximos ao gene da Tripsina1. Dados extraídos utilizando a ferramenta BioMart do *VectorBase*.

Tabela 3. Caracterização *in silico*, das serina peptidases do tipo tripsina identificadas no intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar

Uniprot Identificador	Proteína	Target P ①	Signal P ②	TMHMM ③	N-Glycosylation ④	O-Glycosylation ⑤	Exon number ⑥	Paralogues number ⑥	Supercontig ⑥
BOWIS4	Trypsin 1	S (0.901)	23^24 (0.761)	No	158-NETV (0.7243)	36-T (0.627995) / 40-S (0.689668)	1	36	3,139
BOWE94	Trypsin 2	S (0.910)	18^19 (0.818)	No	No	No	2	36	3,94
BOXCW2	Trypsin 4	S (0.973)	22^23 (0.935)	inside: 1-6 / Tmhelix:7-26 / outside: 27-258	27-NGTQ (0.8040)	No	2	36	3,911
BOX667	Trypsin 5	S (0.952)	17^18 (0.855)	inside: 291-293 / Tmhelix: 268- 290 / outside: 1-267	65-NRTV (0.6702) / 183-NVTV (0.8306)	151-T (0.653105) / 159-S (0.523482)	3	13	3,592
BOXES8	Trypsin 7	S (0.832)	21^22 (0.665)	No	106-NVTF (0.6360)	No	3	36	3,1053
BOX870	SP24D	S (0.891)	20^21 (0.801)	No	69-NGSV (0.6998) / 75-NLSV (0.6183)	No	2	24	3,664
B0W9S9	Serine protease 1/2	S (0.960)	26^27 (0.750)	No	No	44-S (0.785129)	2	38	3,54
B0WW44	Cationic trypsin	S (0.926)	20^21 (0.794)	No	No	No	3	24	3,325

① TargetP 1.1 Server. Prediz a localização sub-celular das proteínas. S = secretada. () = Probabilidade Prediction of the subcellular location of trypsin. S = secreted. The number into the parenthesis indicates the probability.

② SignalP 4.0 Server. Prediz a presença e localização do sítio de clivagem de peptídeos sinais. () = Probabilidade.

③ TMHMM 2.0 Server. Prediz a hélices trans-membranas em proteínas. Tmhelix: Hélices trans-membrana.

④ NetNglyc 1.0 Server. Prediz sítios de N-Glicosilações baseados na presença de sequencias motivos Asn-Xaa-Ser/Thr . () = ptobabilidade.

⑤ NetOglyc 4.0 Server. Prediz sítios de O-Glicosilações.

⑥ Análises com a ferramenta BioMart do VectorBase.

Tabela 4. Peptídeos proteotípicos das tripsinas identificadas por LC-MS/MS, preditos pelo software SKYLINE

Uniprot Identificador	Proteína	Peptídeos identificados por MS/MS	Peptídeos espécie específicos	Peptídeos Únicos de Tripsinas	Outras proteínas com o mesmo peptídeo
BOWIS4	Trypsin 1	R.AAYVPAYNQNCNSAYAR.Y R.IVGGFEISIADAPHQVSLQSR.G R.GSHICGGSII SPK.W R.NTIDYDYSLLELK.S K.WILTAAHCTDGASVSNLR.I K.HASGGSVISIK.R	SIM SIM SIM SIM SIM -	NÃO SIM SIM SIM SIM -	Q1KWX6 (Trypsin-like fragment) - <i>Cx. quinquefasciatus</i> - - - - Not predicted by Skyline
BOWE94	Trypsin 2	R.DEPADESQSLVSGWGDTR.S R.LEFGHAVQPVDLVR.D R.GVLVPLVNR.E	NÃO NÃO NÃO	NÃO NÃO NÃO	Q962G7 / Q56GM3 (Trypsin) - <i>Culex pipiens</i> Q962G7 / Q56GM3 (Trypsin) - <i>Culex pipiens</i> Q962G7 / Q56GM3 (Trypsin) - <i>Culex pipiens</i>
BOXCW2	Trypsin 4	K.GCAQPDYGVYADVEK.A K.DDFALLR.L R.VGSSYDYQGQTVIDVAGMTIHPR.Y	SIM SIM -	SIM SIM -	- - Not predicted by Skyline
BOX667	Trypsin 5	K.IIGGFPAEQGDTLHQVSIR.F K.GCGLAAYPGIYSDVAYYR.G R.GWIDSLAGK.C	SIM SIM SIM	SIM SIM SIM	- - -
BOXES8	Trypsin 7	K.DACLGDSSGGPLTCSGK.V R.DYALLNLAK.S R.AVDVPIADHDR.C R.GGQLIAVTR.K	SIM SIM SIM -	SIM SIM SIM -	- - - Not predicted by Skyline
BOX870	SP24D	K.LGESIEYDELSQPIALYEGDDLPK.D	-	-	Not predicted by Skyline
BOW9S9	Serine protease 1/2	R.TGETFVDNQATVSGFGR.T R.TVDGGPVSPTK.N	SIM SIM	NÃO NÃO	<i>Cx. quinquefasciatus</i> - Q23731 (Serine protease) <i>Cx. quinquefasciatus</i> - Q23731 (Serine protease)
BOWW44	Cationic trypsin	R.IVVHPQYAEGNLANDIAVIR.V	SIM	SIM	-

Identificação e classificação funcional do perfil proteico do intestino fêmeas *Cx. quinquefasciatus* alimentadas com açúcar

Além das peptidases identificadas por espectrometria de massas a partir das frações de géis SDS-PAGE, foi possível identificar centenas de outras proteínas. Utilizando o software Mascot como ferramenta de busca e identificação das proteínas por PSM (validado estatisticamente pelo Scaffold), foi possível identificar 1397 proteínas distribuídas em 1090 grupos (Suplementar I). Utilizando o software COMET (seguida da validação pelo SEPro), logramos identificar 1183 proteínas sendo 942 proteínas diferentes distribuídas em 1034 grupos (Suplementar II). Os dados são provenientes de dois experimentos independentes provenientes de três replicatas biológicas.

As proteínas identificadas pelo Mascot foram analisadas utilizando a ferramenta do BioMart disponível no VectorBase. Foi possível acessar a ontologia genica das proteínas identificadas de acordo com os termos anotados para cada proteína depositadas no UNIPROT. Foram observadas 201, 313 e 81 diferentes termos de processos biológicos, funções moleculares e componentes celulares, respectivamente. Com relação aos componentes celulares assinados, o termo intracelular (47) foi o mais observado, seguido de ribossomo (41), citoplasma (37), membrana (32), componente integral de membrana (18), núcleo (16), complexo do proteosoma (12), nucleosoma (10) (Figura 15 A). Os processos biológicos mais comuns desempenhados pelas proteínas identificadas no intestino são: processo de oxidação-redução (123), processos metabólicos (91), tradução (43), proteólises (43), processos metabólicos de carboidratos (24), dobramento de proteínas (24), homeostase do redox celular (14), transporte intracelular de proteínas (13), proteólises em processos catabólicos de proteínas (12), processo glicolítico (10), montagem de nucleossomo (10), transporte (10), transporte de prótons acoplado a síntese de ATP (10) e transporte de prótons acoplado a hidrólise de ATP (10) (Figura 15 B).

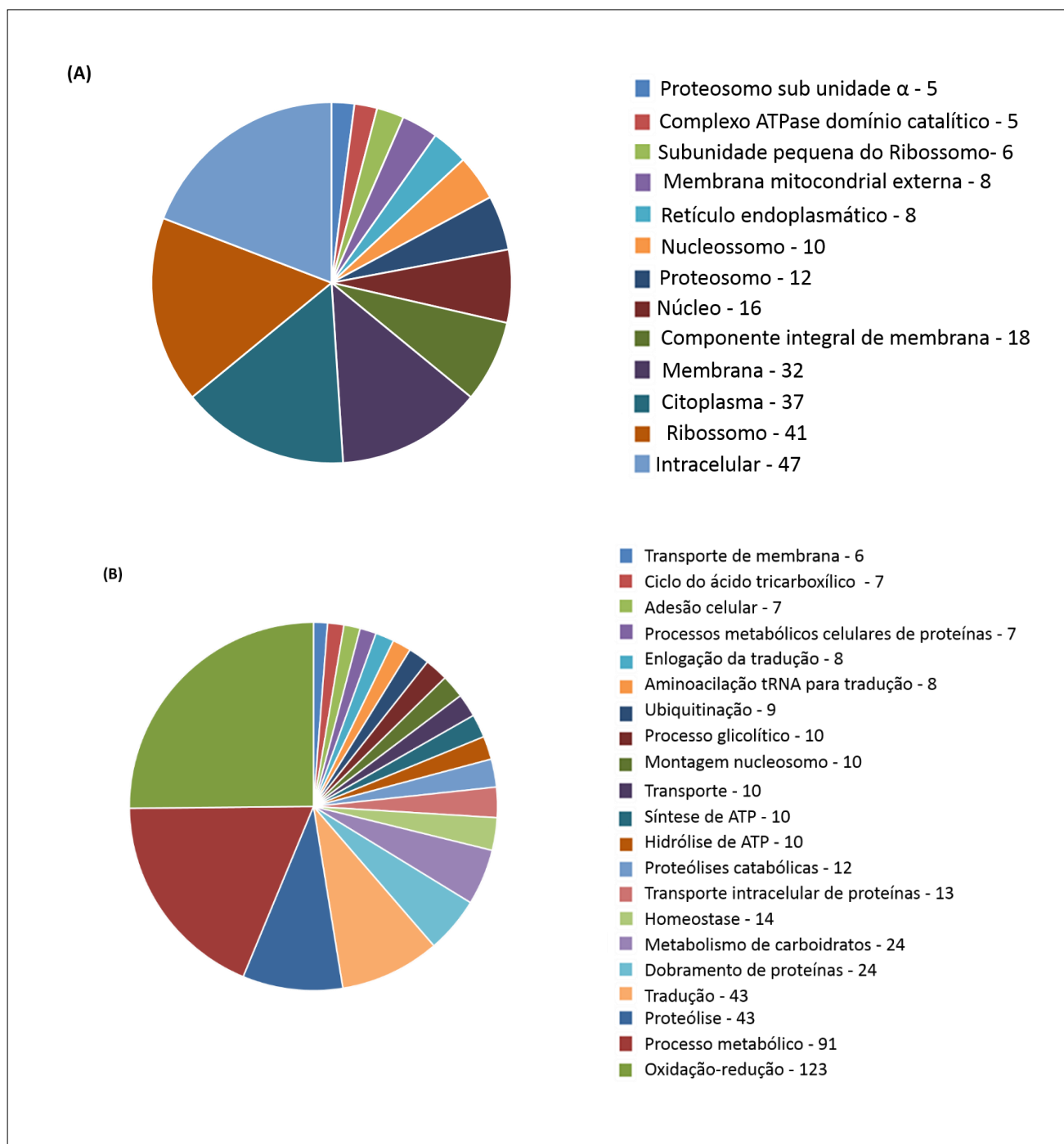


Figura 15. Ontologia gênica das proteínas totais de intestino de *Cx. quinquefasciatus* identificadas por espectrometria de massas. As ontologias foram acessadas através da ferramenta BioMart do VectorBase e o hits correspondentes para cada proteína foram contados. **(A).** Componente celular. **(B).** Processo celular.

Em relação às funções moleculares, as mais encontradas foram: atividade catalítica (112), ligação em proteínas (94), atividade de oxireductase (86), ligação em

ATP (59), constituintes estruturais dos ribossomos (45), ligação ao ácido nucléico (32), ligação ao nucleotídeo (31), ligação a ions de ferro (27), ligação a ions de cálcio (27), ligação a ions de zinco (27), ligação a ion metálico (23), atividade de hidrolase (22), ligação ao RNA (21), ligação (20) (Figura 16 A). Por fim, 168 proteínas não possuíram informações anotadas a respeito de suas ontologias genicas.

A exploração da ontologia gênica de todas as proteínas identificadas ao mesmo tempo representa uma visão holística dos papéis desempenhado pelas proteínas nos processos fisiológicos que ocorrem no intestino. Entretanto, quando analisamos essas proteínas após uma pré-seleção quanto às características específicas como, por exemplo, presença domínios de peptídeos sinais, podemos entender com mais clareza o papel das proteínas identificadas (Figura16 B).

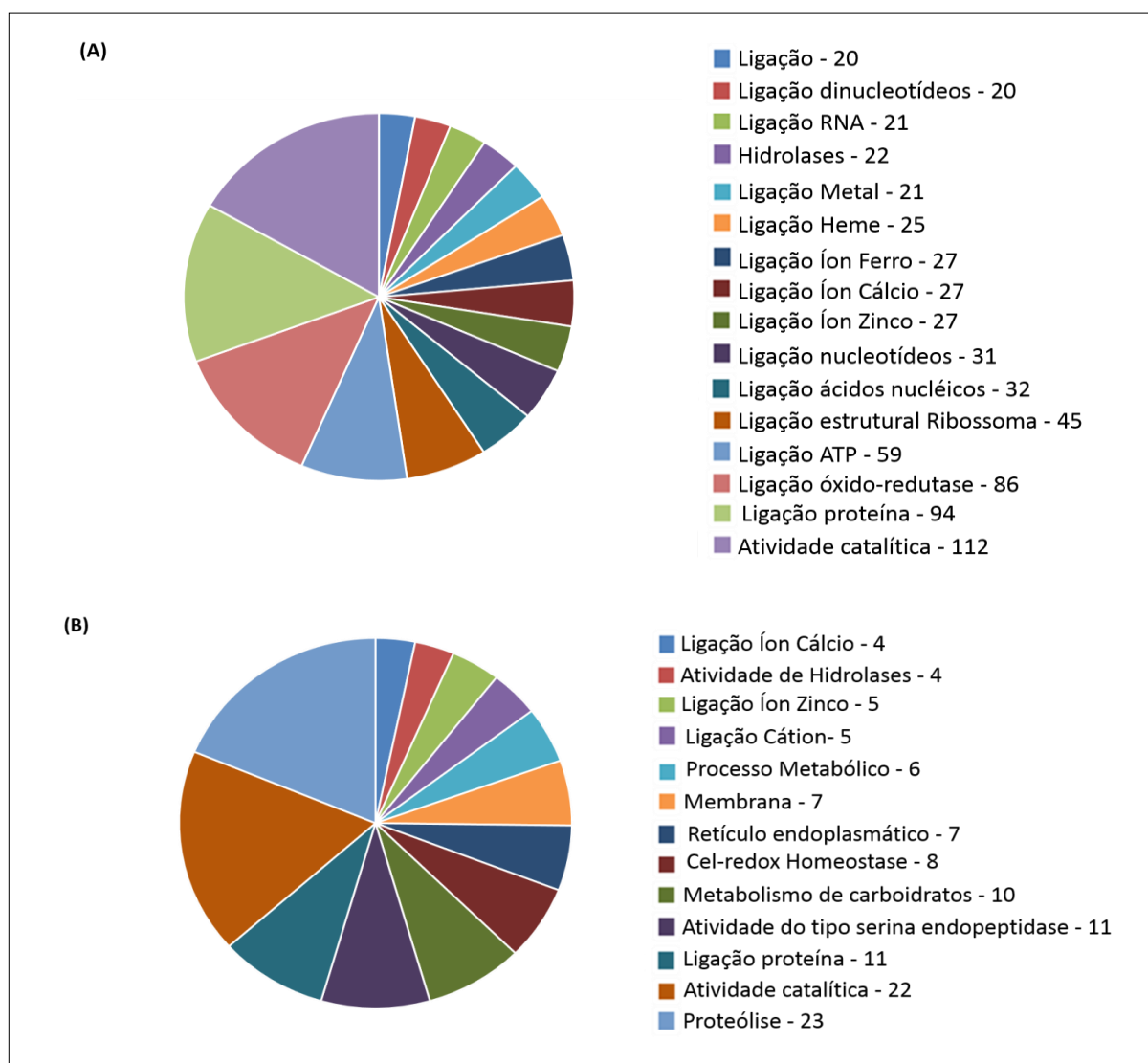


Figura 16. Ontologia gênica das proteínas totais de intestino de *Cx. quinquefasciatus* identificadas por espectrometria de massas. As ontologias foram acessadas através da ferramenta BioMart do VectorBase e o hits correspondentes para cada proteína foram contados. **(A).** Função molecular. **(B).** As proteínas foram previamente separadas de acordo com a predição de peptídeo sinal e a função molecular das proteínas preditas como secretadas foram assinadas.

Grupo de proteínas com funções fisiológicas importantes

Além das análises de ontologia gênica das proteínas totais, as análises foram direcionadas para alguns grupos de proteínas historicamente importantes em

mosquitos. As proteínas foram selecionadas com base em estudos relacionados não só com parálogos em *Cx. quinquefasciatus*, mas também por ortólogos, em outras espécies de mosquitos. Dentre todas as proteínas identificadas, louvou-se identificar outras enzimas relacionadas com proteólises e inibidores de SPs, enzimas relacionadas com a digestão de carboidratos, proteínas envolvidas no processo de detoxificação e resistência a inseticidas, proteínas relacionadas com resposta imune e de ligação à toxina do *Bacillus thuringiensis*, um inseticida amplamente empregado no controle de pré-imaginais de mosquitos (Anexo I).

Enzimas proteolíticas

Além das tripsinas descritas anteriormente, outras enzimas proteolíticas foram identificadas nesse estudo. Identificou-se três diferentes quimiotripsinas (BOW653, BOW174, BOWH74), cinco aminopeptidases (BOWS64, B0XAAQ5, BOW6N8, BOWH45, B0X730), cinco carboxipeptidases (BOWS12, BOW4H6, B0XCT6, BOWTW4, B0WNR6), duas dipeptidases (BOWDL4, B0X281), quatro metalo peptidases (BOW1Y3, BOW1Y6, BOWH17, BOW1Y2), duas catepsinas (BOWI10, BOWOV3), e uma Lipase (BOWDK6).

Inibidores de serina peptidases

Foram identificados dois inibidores de seriana peptidases. Uma serpina (BOWQ11) e uma Alaserpina (B0X5Z9).

Glicosidases

As enzimas responsáveis pela digestão de carbotidatos são também importantes para garantir um bom fitness aos mosquitos. Com relação às glicosidases foram identificadas: quatro alpha-glicosidases (B0XJR7, B0XAA1, B0XBN3, BOWS02), uma Glicosidase 2 subunidades beta (BOWS45), uma neutral alfa-

glicosidase AB (B0WQR9), uma oligo-1,6-glicosidase (B0WFP7) e duas alfa-amilases (B0XGH0, B0XJR6).

Detoxificação e resistência a inseticidas

Neste trabalho foi identificado por espectrometria de massas 17 proteínas da família do Citocromo P450 (B0XJT6, B0WRU2, B0VZI7, B0XCA1, B0WTS8, B0WTS9, B0W673, B0X581, B0W6Y, B0WTQ1, B0WTP9, B0XJY3, B0WTS7, B0WFB9, B0WFC1, B0WTS5, B0WTS3).

Foram identificadas proteínas da família das Glutathione s-transferases, conhecidas por causarem resistência em mosquitos. Foram identificadas quatro glutathione-s-transferase theta (B0XGK3, B0W6C9, B0XGJ5, B0X3C8), três glutathione s-transferase (B0WFX0, B0VZ90, B0W6C9), duas glutathione S-transferase 1 (B0XAJ0, B0W6C3), duas glutathione S-transferase 1 microsomal (B0XG85, B0XG85), três glutathione transferases (C4B4V7, C4B4V8, B0WFX0), uma glutathione S-transferase 1-6 (B0W6B0) e uma glutathione sintetase (B0WSC1).

Foram identificadas também enzimas candidatas a participarem de processos de detoxificação como as carboxilesterase-6 (B0WIH9), proteína palmitoil tioesterase 1 (B0W1X8), esterase A11 (A5Y5K2), ubiquitina thioesterase OTUB1 (B0W5S7), esterase A4-B4 (R9RIG8) e esterase B1 (EST1).

Interação parasito/vetor

Dentre as proteínas que participam da interação parasito/vetor foram identificadas várias famílias de proteínas conhecidas como reconhecedoras de padrões. Dentre elas, foram identificadas: reconhecedoras de padrões do tipo lectina (CTLs, C-Type Lectins), PRGP (proteína reconhecedora de peptideoglicanos), GGBP (proteína de ligação gram negativa), GALE (lectinas ligadoras a galactosídeos), LRRD (domínios repetidos ricos em leucina) e FBN (imunolectinas com domínio tipo fibrinogênio).

Identificamos neste trabalho moléculas do tipo lectina, lectina salivar tipo C (B0WID9, B0WIF2, B0WIE3, B0WID7), lectina ligadora a manose ER (B0XG12) e galectinas (B0XEF4, B0W0J6, B0X1F3).

Arginine kinase (B0WLS0) promove o desenvolvimento F-actin capping protein (B0WTE9), NADPH--cytochrome P450 reductase (B0X829), Glutathione-s-transferase theta (B0XGJ5, B0XGK3, B0X3C8), Enolase B0W1N4, Leucine-rich transmembrane proein (B0WPC7, B0X0G6), Signal recognition particle subunit SRP68 (B0WMF9).

Atividade proteolítica em intestino de fêmeas de *An. albitarsis* s.s alimentadas com açúcar

Curso temporal

Para verificar a influência do tempo de reação sobre as atividades proteolíticas de intestino de fêmeas de *An. albitarsis* alimentadas com açúcar, os géis foram incubados por 0.5, 1, 2 e 4 horas em tampão de reação Tris-HCl a pH 7,5 e 37 °C (Figura 17). A intensidade das atividades proteolíticas aumentou progressivamente até 4 horas. Entretanto, no tempo de 4 horas a atividade foi demasiada e degradou o substrato de tal forma que impediu a visualização precisa dos halos de atividade. Em 2 horas de reação, foi observado um perfil proteolítico com bandas principais migrando entre 24 e 100 kDa, aproximadamente. Assim, o tempo de reação enzimática de 2 horas foi escolhido para os ensaios subsequentes das amostras de intestino de *An. albitarsis* fêmeas.

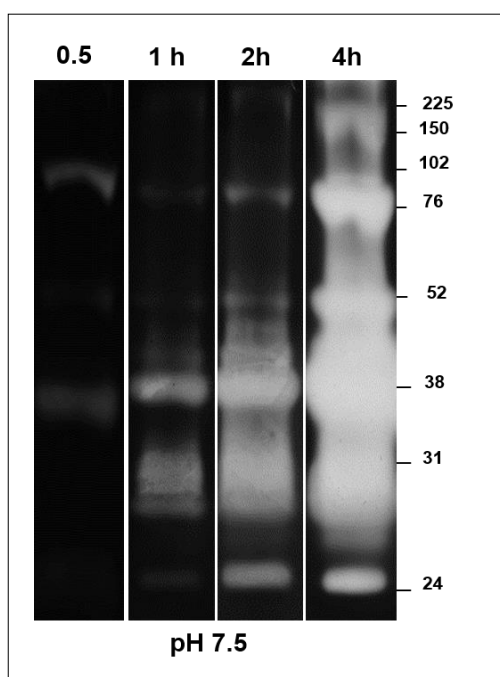


Figura 17. Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *An. albitarsis* alimentadas com açúcar. Atividades proteolíticas foram avaliadas após 0.5, 1, 2 e 4 horas de incubação em tampão de Tris-HCl 0.1 M com pH 7,5. Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Influência do pH

Semelhante ao que foi feito com as amostras de fêmeas de *Cx. quinquefasciatus*, as atividades proteolíticas de fêmeas de *An. albitarsis* foram caracterizadas utilizando enzimografia em gel copolimerizado com gelatina e substrato sintético fluorogênico em solução. Para determinar a influência exercida pelo pH sobre as atividades proteolíticas, os géis foram incubados por 2 horas em tampões de reação de pH 3,5; 5,5; 7,5 e 10. As atividades enzimáticas foram mais intensas em pH 10 (Figura 18 A). Porém, a alta atividade enzimática em pH 10 dificultou a visualização das bandas de atividade com nitidez. Em pH 3,5 e 5,5 nenhuma banda pode ser observada com clareza. O perfil proteolítico foi melhor resolvido em pH 7,5, no qual visualizou-se bandas migrando entre 24 e 100 kDa, aproximadamente.

A atividade de intestino de fêmeas foi analisada utilizando o substrato fluorogênico Z-phe-arg-MCA (Figura 18 B). Tendo em vista que a alta atividade proteolítica de intestino de *An. albitarsis*, o extrato proteico foi diluído dez vezes para os ensaios em solução, diferentemente das outras amostras de *Cx. quinquefasciatus* e *Ae. aegypti* alimentadas com açúcar. Assim, o equivalente 1/100 de intestino foi incubado em tampão acetato de sódio (pH 3,5 e 5,5) e Tris-HCl (pH 7,5 e 10), contendo 100 μ M do substrato. Em pH 3,5 e 5,5 a atividade foi reduzida e aumentou progressivamente até pH 10. Em pH 10 a cinética de hidrólise do substrato atingiu o platô em menos de 15 minutos e se manteve até os 45 minutos de reação. Em pH 7,5 a atividade atingiu o platô com aproximadamente 40 min. de reação.

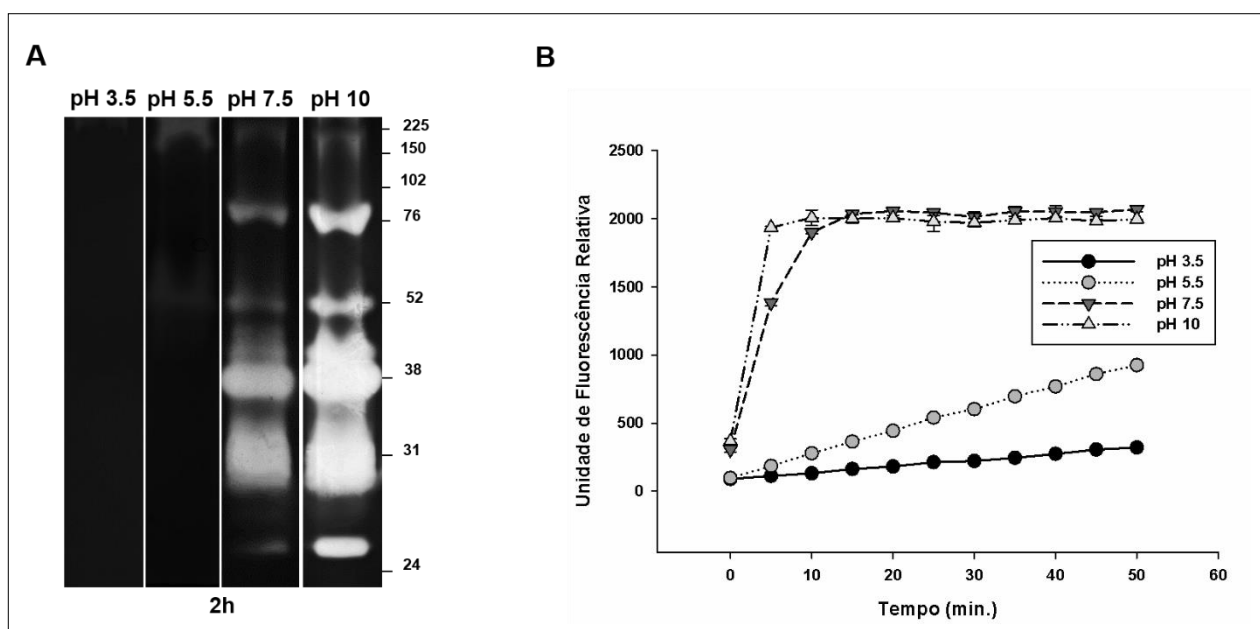


Figura 18. Efeito do pH sobre as atividades proteolíticas de extratos proteicos de intestinos de fêmeas de *An. albitarsis* alimentadas com açúcar. **(A)** A influência do pH foi avaliada através da incubação dos géis de enzimografia à 37°C por 2 horas em tampão de acetato de sódio (pH 3,5 e 5,5) e em tampão Tris-HCl (pH 7,5 e 10,0). Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B)** Os ensaios em solução foram realizados com substrato fluorogênico Z-Phe-Arg-MCA em tampão acetato de sódio (pH 3,5 e 5,5) e tampão Tris-HCl (pH 7,5 e 10,0).

Efeito dos inibidores

O perfil enzimático exibido pelo intestino de fêmeas foi fortemente inibido por 1 mM PMSF e 100 μ M TLCK em géis incubados por 12 horas em tampão de reação Tris-HCl pH 7,5 (Figura 8). No entanto, as atividades proteolíticas não foram afetadas por 100 μ M TPCK. O perfil proteolítico do intestino não foi afetado por: de 20 μ M E-64; 10 μ M pepstatina A; e 10 mM 1,10-fenantrolina (dados não apresentados).

A atividade de intestino de fêmeas foi analisada utilizando o substrato fluorogênico Z-phe-arg-MCA (Figura 19). A amostra foi incubada em tampão Tris-HCl pH 7,5 contendo 100 μ M do substrato, na ausência (controle) ou presença de inibidores 1 mM PMSF, 100 μ M TLCK, 100 μ M TPCK, 20 μ M E64. Foi possível verificar que tanto o PMSF quanto o TLCK inibem a atividade das peptidases presentes no intestino, corroborando o observado no ensaio enzimográfico em gel e indicando que o perfil proteolítico dessa espécie é composto por serina peptidases do tipo tripsina.

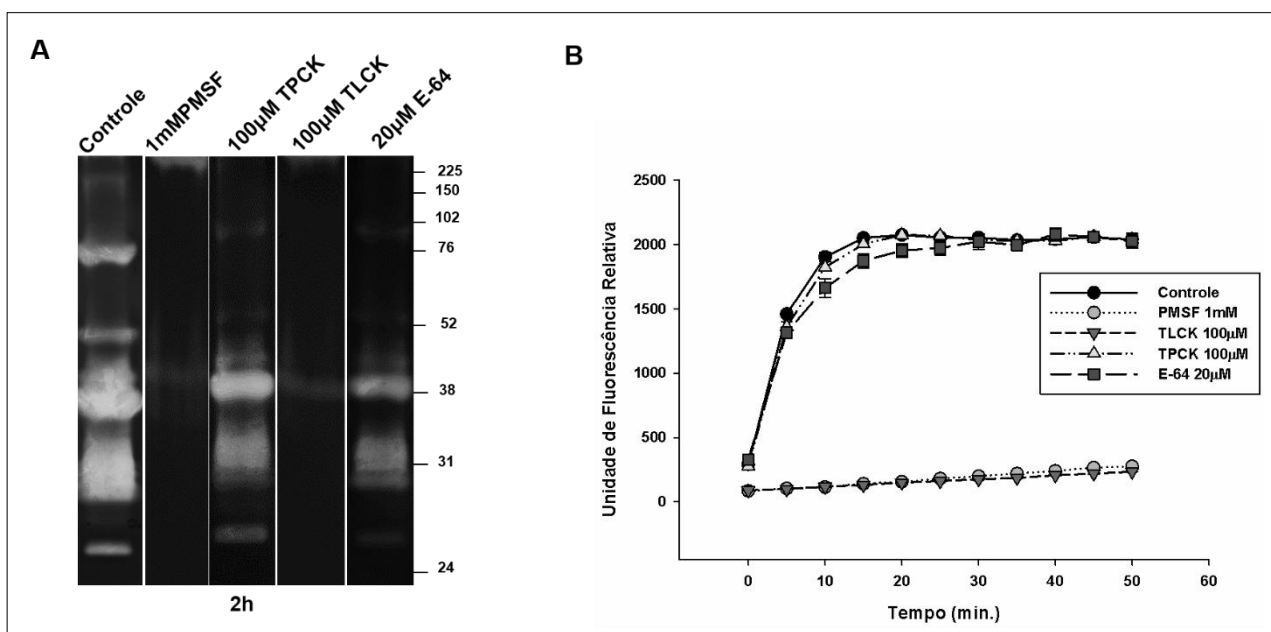


Figura 19. Efeito dos inibidores de peptidases no perfil proteolítico de extrato do intestino médio de fêmeas de *An. albitalis* alimentadas com açúcar. **(A)**. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μ M de TPCK e 100 μ M de TLCK. As atividades proteolíticas foram detectadas após incubação dos géis por 2 horas à 37°C em tampão Tris HCl com pH 7,5. O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B)**. Os ensaios em solução foram realizados com substrato fluorogênico Z-Phe-Arg-MCA em tampão de Tris-HCl 100 mM com pH 7,5, na ausência (controle) ou na presença de 1 mM de PMSF, 100 μ M de TLCK, 20 μ M E-64 ou 100 μ M de TPCK.

Atividade proteolítica em intestinos de machos de *An. albitarsis* s.s alimentados com açúcar

Assim como ocorrido para as amostras de intestino das fêmeas o perfil enzimático exibido pelo intestino de machos foi fortemente inibido por 1 mM PMSF e 100 μ M TLCK em géis incubados por 12 horas em tampão de reação pH 7,5 (Figura 20). No entanto, as atividades proteolíticas não foram afetadas por 100 μ M TPCK, 20 μ M E-64. O perfil proteolítico não foi afetado por 10 μ M pepstatina A e 10 mM 1,10-fenantrolina (dados não apresentados).

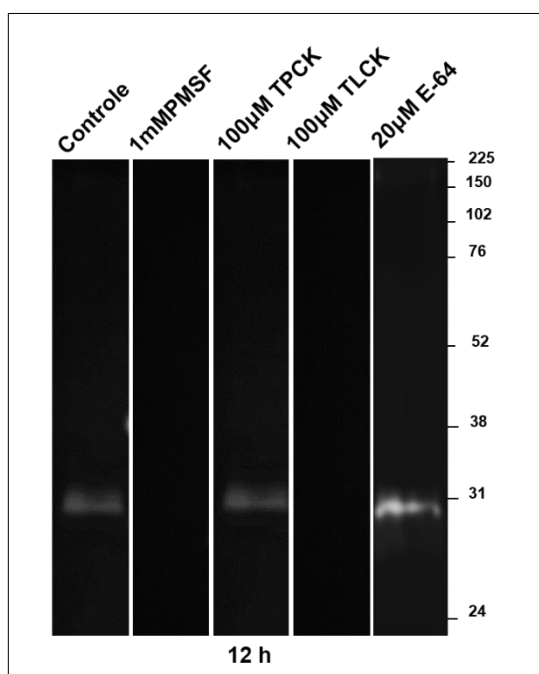


Figura 20. Efeito dos inibidores de peptidases sobre o perfil proteolítico de intestino de machos de *An. albitarsis* alimentados com açúcar. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μ M de TPCK e 100 μ M de TLCK. As atividades proteolíticas foram detectadas após incubação dos géis por 12 horas à 37°C em tampão Tris-HCl (pH 7,5). O controle foi processado sob as mesmas condições, mas na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Peptidases identificadas por espectrometria de massas

Assim como realizado com as amostras proteicas de intestino de *Cx. quinquefasciatus*, os extratos de intestinos de *An. albitarsis* fo aplicados em gel de SDS-PAGE e após a coloração, o gel foi dividido em 24 frações e as proteínas digeridas por tripsinas e análizadas em espectrômetro. Utilizando dois motores de busca independentes: Mascot (seguida da validação pelo Scaffold) e o COMET (seguida de validação pelo SEPro), foram identificadas sete serina peptidases do tipo tripsina constitutivamente expressas no intestino de fêmeas alimentadas com açúcar (Tabela 5). Adicionalmente, uma proteína desconhecida, que possui em sua sequência características de serina peptidases, foi exclusivamente identificada pelo Mascot com cinco diferentes peptídeos, dentre eles, um considerado peptídeo único (Tabela 5, E3XCU3, em cinza). Além disso, duas outras serina peptidases foram identificadas exclusivamente pelo COMET. A SPtrip T1E9K9 foi identificada com três peptídeos únicos. A Sptrip W5JUG9 foi identificada com quatro peptídeos, sendo um deles considerado único (Tabela 5).

Identificação e classificação funcional do perfil proteico do intestino fêmeas *An. albitarsis* s.s alimentadas com açúcar

Em *An. albitarsis* foram identificadas 916 proteínas distribuídas em 748 grupos sendo 554 proteínas únicas, utilizando o COMET como ferramenta de busca. Através da ferramenta GO explorer disponível no PATTERNLAB, as proteínas foram classificadas de acordo com seu componente celular, processo biológico e função molecular, segundo os termos de ontologias gênicas depositados para cada proteína na base de dados UNIPROT (Suplementar 3). (Figuras 21 e 22).

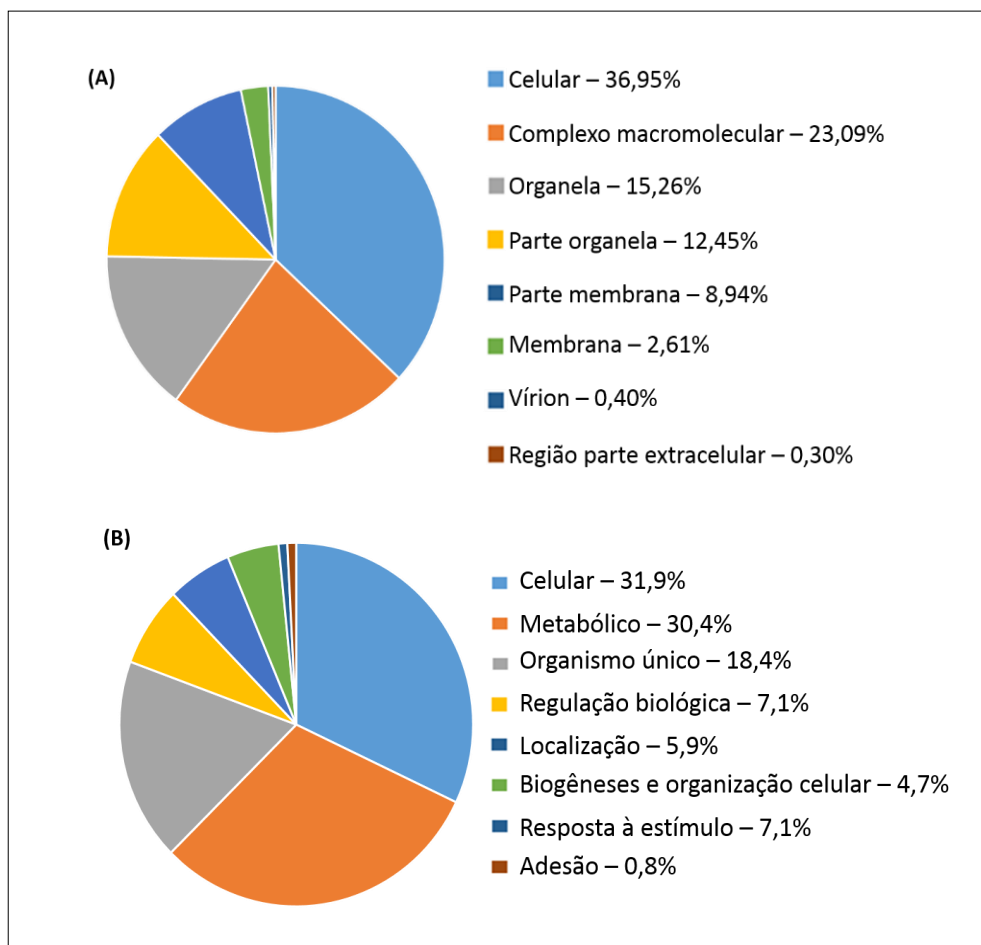


Figura 21. Classificação funcional das proteínas identificadas em intestino de fêmeas de *An. albitalarsis* segundo o Gene Ontology. **(A)**. Componente celular. **(B)**. Processo biológico. As ontologias foram acessadas utilizando a ferramenta de GO explorer do PATTERNLAB.

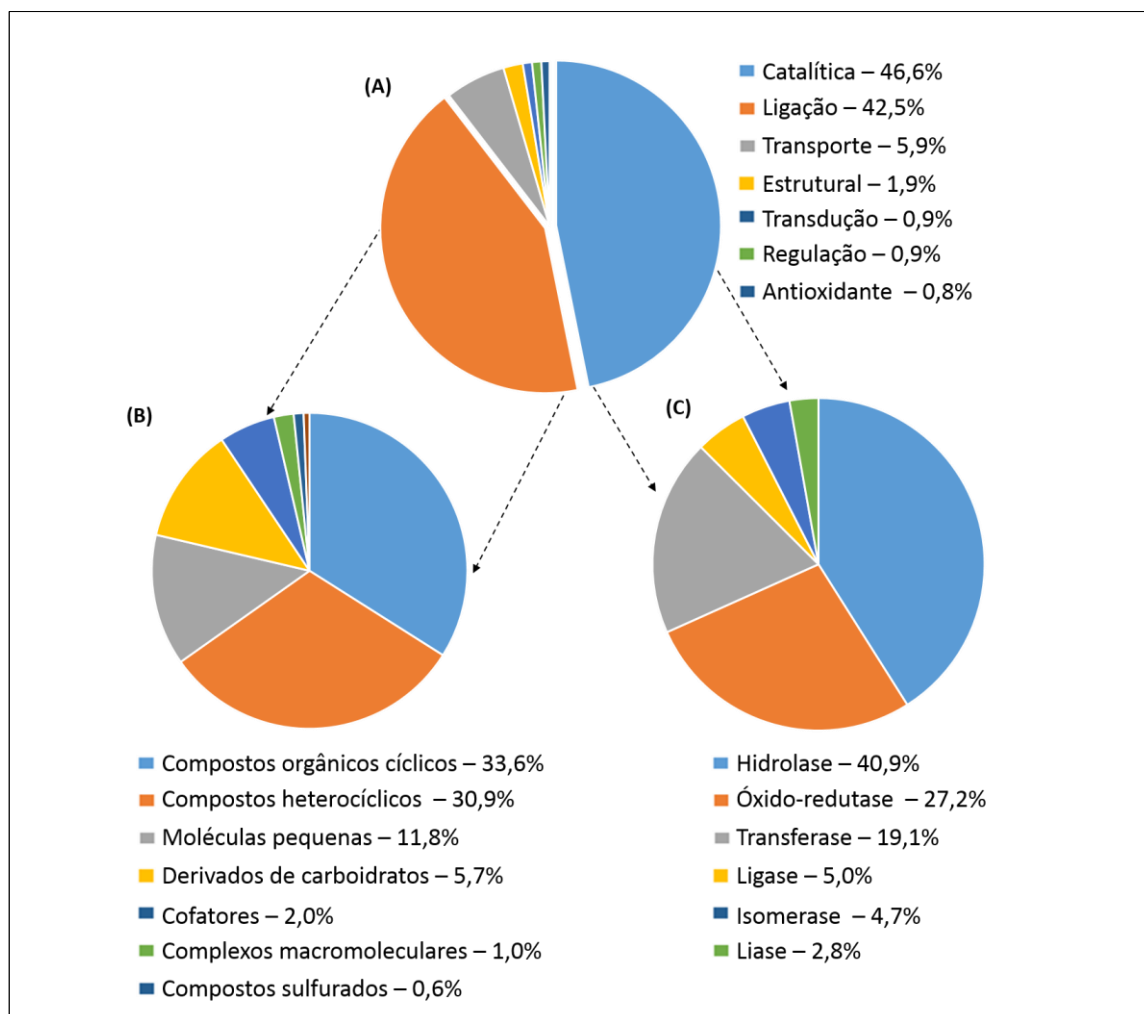


Figura 22. Classificação funcional das proteínas identificadas em intestino de fêmeas de *An. albitarsis* segundo o Gene Ontology. (A). Função molecular. As proteínas envolvidas na função molecular de ligação e catalítica foram separadas em (B). Ligação e (C). Catalítica. As ontologias foram acessadas utilizando a ferramenta de GO explorer do PATTERNLAB.

Tabela 5. Serina peptidases do tipo tripsina, identificadas por LC-MS/MS, em intestino de fêmeas de *An. albitalarsis* s.s alimentadas com açúcar

Identificador Uniprot	Massa Molecular	Mascot Peptídeo Único	Mascot Espectros Totais	Mascot Cobertura %	Peptídeo Identificado pelo MASCOT	Mascot Ion score	Peptídeo Identificado pelo COMET	COMET Peptídeo Único	COMET Espectros Totais	COMET Cobertura %
T1DN08_ANOAQ	29,5 kDa	7	68	40	(K)AQSVDLPEQNEPVEDATLcK(V) (K)DAcQGDSGGPLYANQK(L) (K)VSGWGNTQNWSESTK(L) (K)WVLTAAHcTAGASPSSLAVR(Y) (R)IVGGFAIDISEAPYQVSLQHR(G) (R)GSHNcGGSIASK(W) (R)YVSGGNR(I)	25,8 69,52 52,4 48,07 93,34 78,33 76,02	(K)AQSVDLPEQNEPVEDATLCK(V) (K)DACQGDSSGGPLYANQK(L) (K)VSGWGNTQNWSESTK(L) (K)WVLTAAHCTAGASPSSLAVR(Y) (R)IVGGFAIDISEAPYQVSLQHR(G) (R)GSHNCGGSIASK(W) (K)YDGSTISYDFSLLEEEAIHFSDK(A)	7	104	46
T1E7R5_ANOAQ	27,4 kDa	1	30	17	(R)EVVLSGWGLQK(D) (R)VVLVSER(F) (K)YNPPIFVR(N) (R)LTIHLPGGK(R) (K)LGQFPYQAR(L)	72.76 44.82 30.04 30.3 50.48	(R)EVVLSGWGLQK(D) (R)VVLVSER(F) (K)YNPPIFVR(N) (R)LTIHLPGGK(R)	1	22	13
T1DJF4_ANOAQ	33,7 kDa	4	65	22	(R)AIELQHEDDTGPEEDEQAmVSGWGATK(D) (R)NRGDDIHR(S) (R)FDcGGSIISPNIWLTAAHcTHGIPSR(E) (K)DRYTSNR(V)	73,71 38,62 14,99 21,11	(R)AIELQHEDDTGPEEDEQAmVSGWGATK(D) (R)NRGDDIHR(S) (R)AIELQHEDDTGPEEDEQAmVSGWGATK(D)	2	85	11
T1DRW4_ANOAQ	30,7 kDa	2	10	10	(R)SDHSDFAGLQGIASGFGR(T) (R)LPPIAVFNDR(V)	50.3 47.8	(R)SDHSDFAGLQGIASGFGR(T) (R)LPPIAVFNDR(V) (R)VQPIALPAR(S)	3	11	13
T1DEY9_ANOAQ	30,1 kDa	2	6	13	(R)IVGGFAIDISEAPYQVSLQR(F) (K)YNGNTISYDFSLQLK(Q)	48.04 82.57	(R)IVGGFAIDISEAPYQVSLQR(F) (K)YNGNTISYDFSLQLK(Q)	2	12	13
T1DEQ6_ANOAQ	30,9 kDa	2	18	10	(R)GHICGGSLIGAK(L) (K)DScQGDSGGPLVcEGR(L)	48.55 60.96	(R)GHICGGSLIGAK(L) (K)DSCQGDSSGGPLVCEGR(L)	2	18	10
T1DSH5_ANOAQ	33,3 kDa	1	12	5	(R)TAEVPAGSEVIISGWGR(I)	89.8	(R)TAEVPAGSEVIISGWGR(I)	1	13	5
T1E9K9_ANOAQ	28,8 kDa	-	-	-	- - -	- - -	(R)VSSFIDWINDK(L) (R)AIVEFIVHEGYSGGIGPNDAVFR(V) (R)VDRPFHLNR(N)	3	91	16
W5JUG9_ANODA	27,4 kDa	-	-	-	- - - -	- - - -	(R)LTIHLPGGK(R) (K)VISNSECAK(T) (K)YNPPIFVR(N) (R)VVLVSER(F)	1	20	12
E3XCU3_ANODA	27,4 kDa	1	27	16	(K)LGQFPYQAR(L) (R)LTIHLPGGK(R) (R)VVLVSER(F) (K)YNPPIFVR(N) (K)VISNSECAK(T)	50.5 30.3 44.8 30.0 38.5	- - - -	-	-	-

Caracterização e comparação do perfil proteolítico do intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar e sangue

Curso temporal de fêmeas alimentadas com açúcar

Para verificar a influência do tempo de reação sobre as atividades proteolíticas de intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar, os géis foram incubados por 2, 4, 6, 12 e 24 horas em tampão de reação pH 7,5 a 37 °C (Figura 23). A intensidade das atividades proteolíticas aumentou progressivamente até 24 horas. Em 24 horas de reação a intensa atividade enzimática levou a uma sobreposição das bandas, não sendo possível resolver o perfil. Em 12 horas de reação, foi observado um perfil proteolítico composto de 8 bandas principais variando entre 24 e 76 kDa. Tendo em vista que o tempo de reação enzimática de 12 horas foi o que melhor resolveu o perfil proteolítico, este foi então escolhido para os ensaios subsequentes das amostras de intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar.

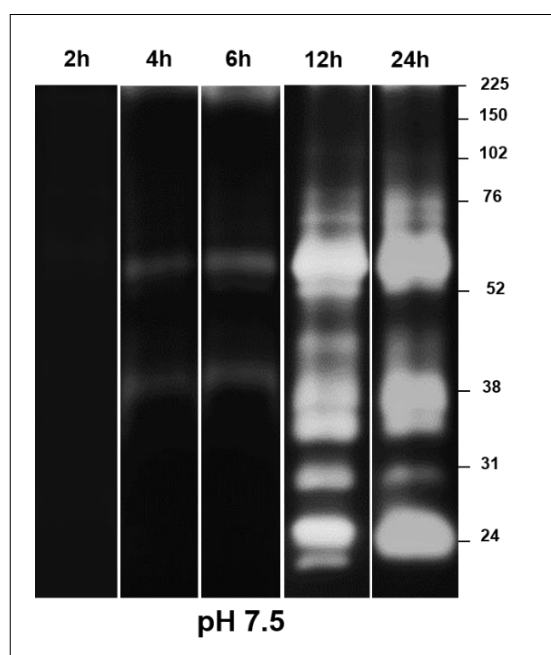


Figura 23. Curso temporal das atividades proteolíticas exibidas pelo intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar. Atividades proteolíticas foram avaliadas após 2, 4, 6, 12 e 24 horas de incubação em tampão de Tris-HCl 0.1 M (pH 7,5). Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Influência do pH na atividade proteolítica de fêmeas alimentadas com açúcar

A influência do pH na atividade enzimática foi avaliada em enzimografia em gel e em solução. Os géis foram incubados por 12 horas em tampões de reação acetato de sódio 0,1M (pH 3,5; 5,5) e Tris-HCl 0,1M (7,5 e 10). As atividades enzimáticas foram fracamente detectadas em pH 3,5 e 5,5, aumentando progressivamente até pH 10 (Figura 24 A). Em pH 3,5 não se observou atividade significativa e algumas bandas de atividade proteolítica observadas em pH 5,5 foram menos intensas quando comparadas com as atividades obtidas em pH 7,5 e 10. Em pH 10, a intensa atividade dos halos proteolíticos dificultou a visualização do perfil. Dessa forma, o perfil proteolítico foi melhor resolvido em pH 7,5, no qual visualiza-se bandas migrando entre 17 e 76 kDa, aproximadamente.

A atividade de intestino de fêmeas foi analisada utilizando o substrato fluorogênico Z-phe-arg-MCA (Figura 24 B). Extrato proteico equivalentes a 1/10 de intestino foi incubado em tampão acetato de sódio (pH 3,5 e 5,5) e Tris-HCl (pH 7,5 e 10), contendo 100 μ M do substrato. A atividade de intestino de fêmeas aumentou progressivamente até pH 10. Em pH 10, a cinética de hidrólise do substrato atingiu o platô em menos de 15 minutos e se manteve até os 45 minutos de reação. Em pH 7,5 a atividade atingiu o platô com aproximadamente 40 minutos de reação.

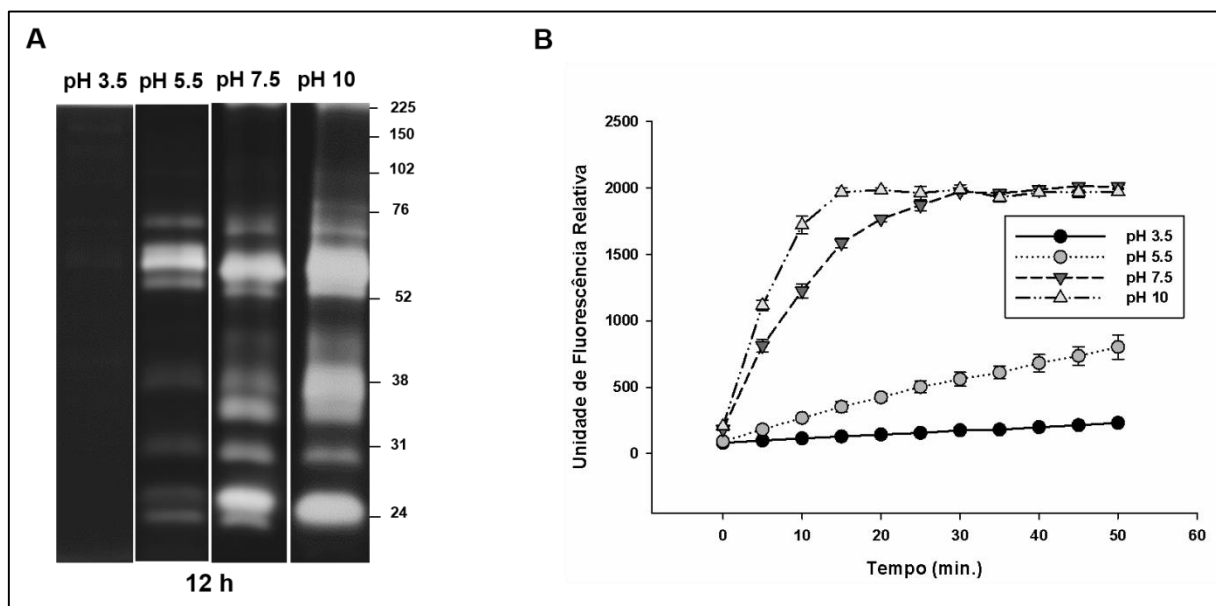


Figura 24. Efeito do pH sobre as atividades proteolíticas de extratos de intestinos de fêmeas de *Ae. aegypti* alimentadas com açúcar. **(A).** A influência do pH foi avaliada através da incubação dos extratos de proteína à 37°C por 12 horas em tampão de acetato de sódio (pH 3,5 e 5,5) e tampão Tris-HCl (pH 7,5 e 10,0). Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B).** Os ensaios em solução foram realizados com substrato fluorogênico Z-Phe-Arg-MCA em tampão de acetato de sódio com pH 3,5 e 5,5 e com tampão Tris-HCl com pH 7,5 e 10,0.

Efeito dos inibidores sobre a atividade proteolítica de intestino de fêmeas alimentadas com açúcar

Semelhante ao ocorrido com as duas outras espécies de mosquitos analisadas neste trabalho, o perfil proteolítico exibido pelo intestino de fêmeas *Ae. aegypti* foram fortemente inibidos por 1 mM PMSF e 100 µM TLCK em géis incubados por 12 horas em tampão de reação Tris-HCl pH 7,5 (Figura 25). No entanto, as atividades proteolíticas não foram afetadas por 100 µM TPCK e 20 µM E-64. As atividades enzimáticas não foram afetadas por 10 µM pepstatina A, e 10 mM 1,10-fenantrolina (dados não apresentados).

Em relação aos ensaios em solução, os extratos de intestino foram incubados em tampão Tris-HCl pH 7.5 contendo 100 µM do substrato fluorogênico Z-phe-arg-MCA (Figura 25 B), na ausência (controle) ou presença de inibidores 1 mM PMSF,

100 μ M TLCK, 100 μ M TPCK, 20 μ M E-64. Foi possível verificar que tanto o PMSF quanto o TLCK inibem a atividade das peptidases presentes no intestino, corroborando os ensaios de enzimografia em gel e indicando que as atividades proteolíticas são de serina proteases do tipo tripsina.

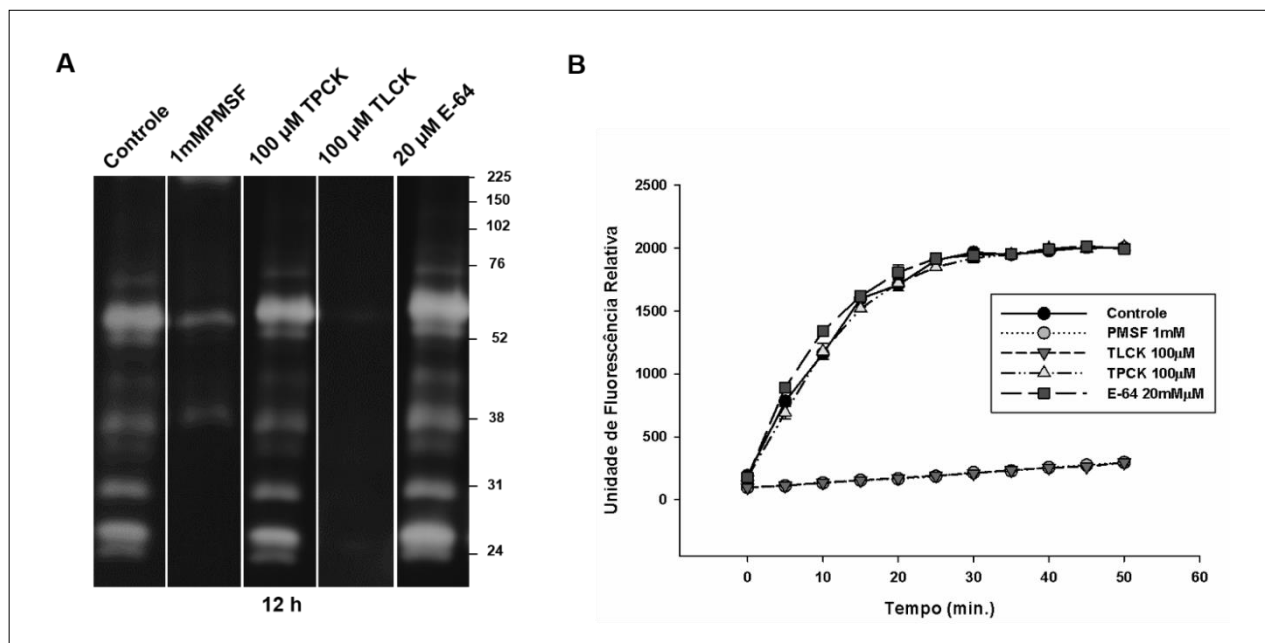


Figura 25. Efeito dos inibidores de peptidases no perfil proteolítico de intestino médio de fêmeas de *Ae. aegypti* alimentados com açúcar. **(A)**. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μ M de TPCK, 100 μ M de TLCK e 20 μ M de E-64. As atividades proteolíticas foram detectadas após incubação dos géis por 12 horas à 37°C em tampão Tris HCl com pH 7,5. O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B)**. Os ensaios em solução foram realizados com substrato fluorogênico Z-Phe-Arg-MCA em tampão de Tris-HCl 100 mM com pH 7,5, na ausência (controle) ou na presença de 1 mM de PMSF, 100 μ M de TLCK, 20 μ M E-64 ou 100 μ M de TPCK.

Atividades proteolíticas em intestinos de machos de *Ae. aegypti*

O perfil enzimático exibido pelo intestino de machos de *Ae. aegypti* foram fortemente inibidos por 1 mM PMSF e 100 μ M TLCK em géis incubados por 24 horas em tampão de reação Tris-HCl pH 7,5 (Figura 26). No entanto, as atividades

proteolíticas não foram afetadas por 100 μM TPCK ou 20 μM E-64. A atividade proteolítica também não foi afetada por 10 μM pepstatina A, e 10 mM 1,10-fenantrolina (dados não apresentados).

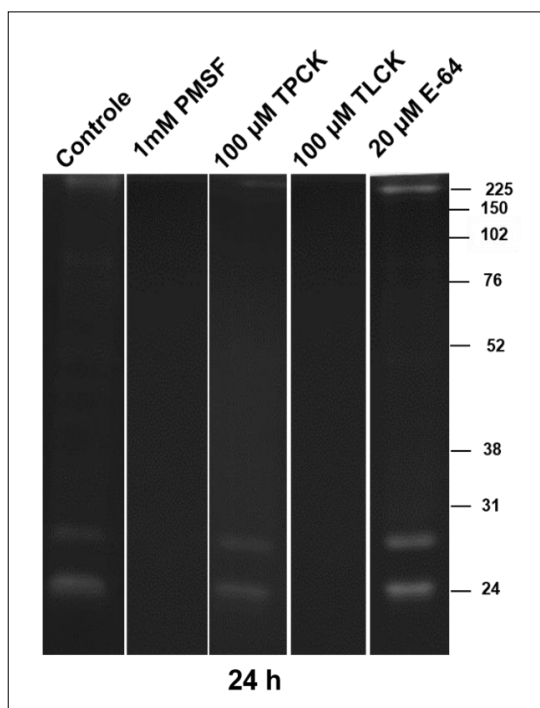


Figura 26. Efeito dos inibidores de peptidases no perfil proteolítico de intestino de machos de *Ae. aegypti* alimentados com açúcar. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μM de TPCK, 100 μM de TLCK e 20 μM de E-64. As atividades proteolíticas foram detectadas após incubação dos géis por 12 horas à 37°C em tampão Tris HCl com pH 7,5. O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Curso temporal de fêmeas de *Ae. aegypti* alimentadas com sangue

Neste ensaio foram comparadas as atividades proteolíticas de intestinos de fêmeas alimentadas com açúcar e com sangue. As fêmeas dos mosquitos foram alimentadas com sangue e os intestinos foram extraídos em diferentes tempos após a alimentação (0,5; 4; 9; 14; 21; 24; 36; 48; 60 e 72 horas). Para verificar a influência do tempo de reação sobre as atividades proteolíticas de intestino de fêmeas de *Ae. aegypti* alimentadas com sangue os géis foram incubados por 2 e 12 horas em tampão de reação Tris-HCl pH 7,5 a 37 °C (Figura 27). Foi usado como controle o extrato de intestino de fêmeas alimentadas com açúcar, cujo perfil proteolítico só pode ser resolvido com 12 horas de reação (Figura 27 A). Com relação ao tempo de reação, observou-se que o tempo de 12 horas (antes requerido para a resolução ideal do perfil de fêmeas alimentadas com açúcar) foi excessivamente alto, e as enzimas consumiram todo o substrato do gel (Dados não apresentados). Desta forma, o curso

temporal foi refeito e estabeleceu-se o tempo de 2 horas para os ensaios com as amostras de intestino de fêmeas alimentadas com sangue. Entretanto, com o tempo de reação enzimática de 2 horas, não foi possível detectar atividade enzimática nas amostras de intestino extraídos nos tempos iniciais da digestão (0,5, 4 e 9 horas PAS). Com o intuito de verificar se a alimentação com sangue induz a expressão de enzimas logo nas primeiras horas do repasto sanguíneo, as amostras de 0,5, 4 e 9 horas foram incubadas em um tempo de reação de 12 horas (Figura 27 A). Desta forma, foi possível observar nesses tempos iniciais da digestão sanguínea uma tênue atividade proteolítica.

No tempo de 2 horas reação, foi possível observar o surgimento de alguns halos de atividades, ainda que sutis, no tempo de 14 horas PAS, os quais aumentaram até 24 horas PAS, com uma ligeira queda no tempo de 36 horas PAS (Figura 27 A). Em seguida, ocorreu uma considerável diminuição de atividade em 48 horas PAS e nenhuma atividade foi observada nos tempos de 60 e 72 horas PAS. Interessantemente, foi possível observar o surgimento de halos de atividades, na região de 17 kDa, não observados nos experimentos realizados com amostras de insetos alimentados exclusivamente com açúcar. Ainda mais, apesar de ligeira semelhanças dos halos de atividade entre algumas regiões do gel dos extratos de intestino de fêmeas alimentadas com açúcar e sangue, como por exemplo, a presença das bandas migrando na região de 24, 31, 38 e 51 kDa, o perfil de atividade proteolítica foi diferente quando o inseto se alimentou com sangue. Ocorreu uma intensificação das bandas na região de 31 a 38 kDa nos extratos de fêmeas PAS, e alteração da intensidade e padrão das bandas visualizadas na região de 52 kDa. Assim, o perfil proteolítico de fêmeas com 24 horas PAS e com duas horas de reação enzimática foi composto de 10 bandas principais variando entre 17 e 76 kDa (Figura 27 A). Finalmente, em pH 10 a atividade proteolítica de intestino de fêmeas PAS, foi intensificada, mas o perfil enzimático se manteve (Figura 27 B).

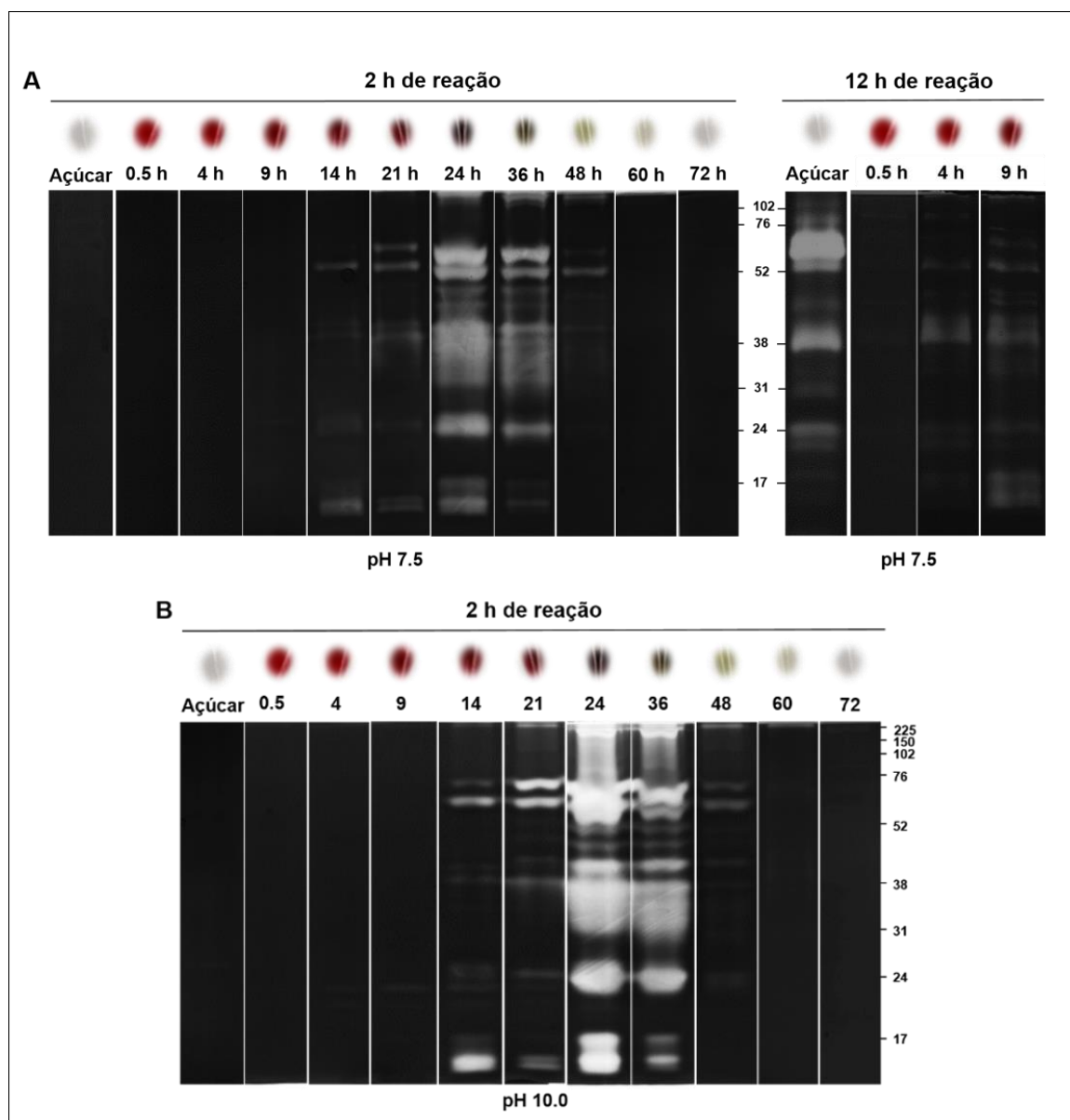


Figura 27. Perfil proteolítico de extrato de intestino de fêmeas de *Ae. aegypti* alimentadas com sangue e açúcar. **(A)**. Perfil da atividade proteolítica em diferentes tempos após alimentação com sangue (PAS). Os géis foram incubados à 37°C por 2 horas ou por 12 horas em tampão de Tris-HCl com pH 7,5. Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa. **(B)**. Efeito do pH sobre o perfil da atividade proteolítica em diferentes tempos após alimentação com sangue. Os géis foram incubados à 37°C por 2 horas em tampão Tris-HCl pH 10,0. Os números à direita indicam as massas moleculares aparentes das peptidases, expressas em kDa.

Efeito dos inibidores em fêmeas alimentadas com sangue

A fim de verificar o efeito dos inibidores sobre as atividades proteolíticas, as amostras foram incubadas na ausência (controle) ou presença dos inibidores. Amostras equivalentes a 1/10 de intestino foram utilizadas nos experimentos de enzimografia em gel. As atividades proteolíticas foram fortemente inibidas por 1 mM PMSF e 100 μ M TLCK em géis incubados por 2 horas em tampão de reação Tris-Cl pH 7,5. No entanto, as atividades proteolíticas não foram afetadas por 100 μ M TPCK; ou por 20 μ M E-64 (Figura 28). A atividade proteolítica também não foi afetada por 10 μ M pepstatina A, e 10 mM 1,10-fenantrolina (dados não apresentados).

Em relação aos ensaios com substrato sintético fluorogenico em solução, amostras equivalentes a 1/100 de intestino foram incubadas em tampão Tris-HCl pH 7.5 contendo 100 μ M do substrato Z-phe-arg-MCA (Figura 29), na ausência (controle) ou presença de inibidores 1 mM PMSF, 100 μ M TLCK, 100 μ M TPCK, 20 μ M E-64. Verificou-se que tanto o PMSF quanto o TLCK inibem a atividade das peptidases presentes no intestino PAS, indicando que são serina proteases do tipo tripsina.

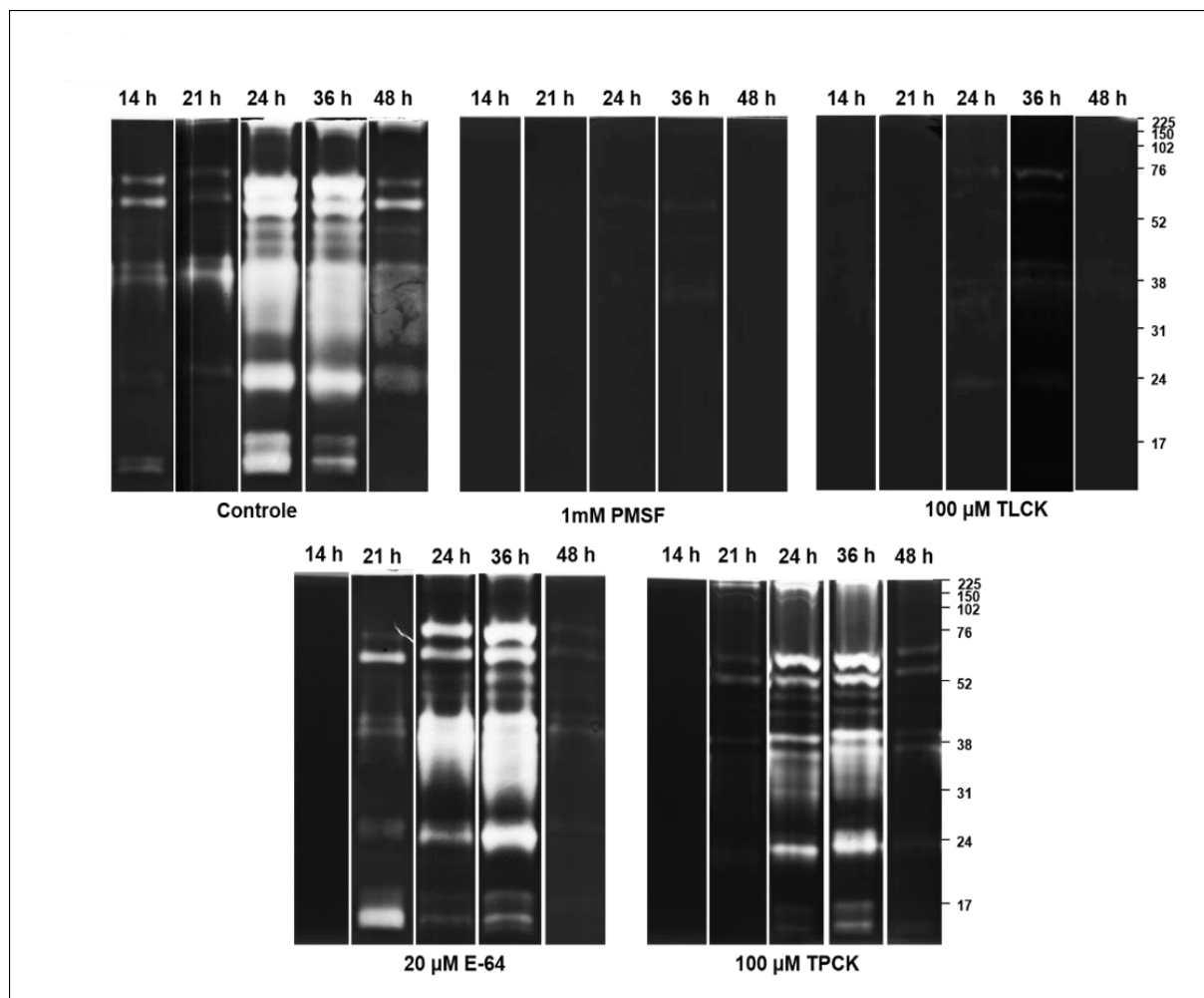


Figura 28. Efeito dos inibidores de peptidases sobre o perfil proteolítico de intestino de fêmeas de *Ae. aegypti* pós alimentados com sangue (PAS). Intestinos de fêmeas alimentadas com sangue foram isolados em diferentes tempos após a alimentação -14, 21, 24, 36 e 48 horas. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 μ M de TLCK, 20 μ M de E-64 e 100 μ M de TPCK. As atividades proteolíticas foram detectadas após a incubação do gel de enzimografia por 2 hora à 37°C em tampão Tris HCl com pH 7,5. O controle foi processado sob as mesmas condições na ausência dos inibidores. Os números à esquerda indicam as massas moleculares aparentes das peptidases, expressas em kDa.

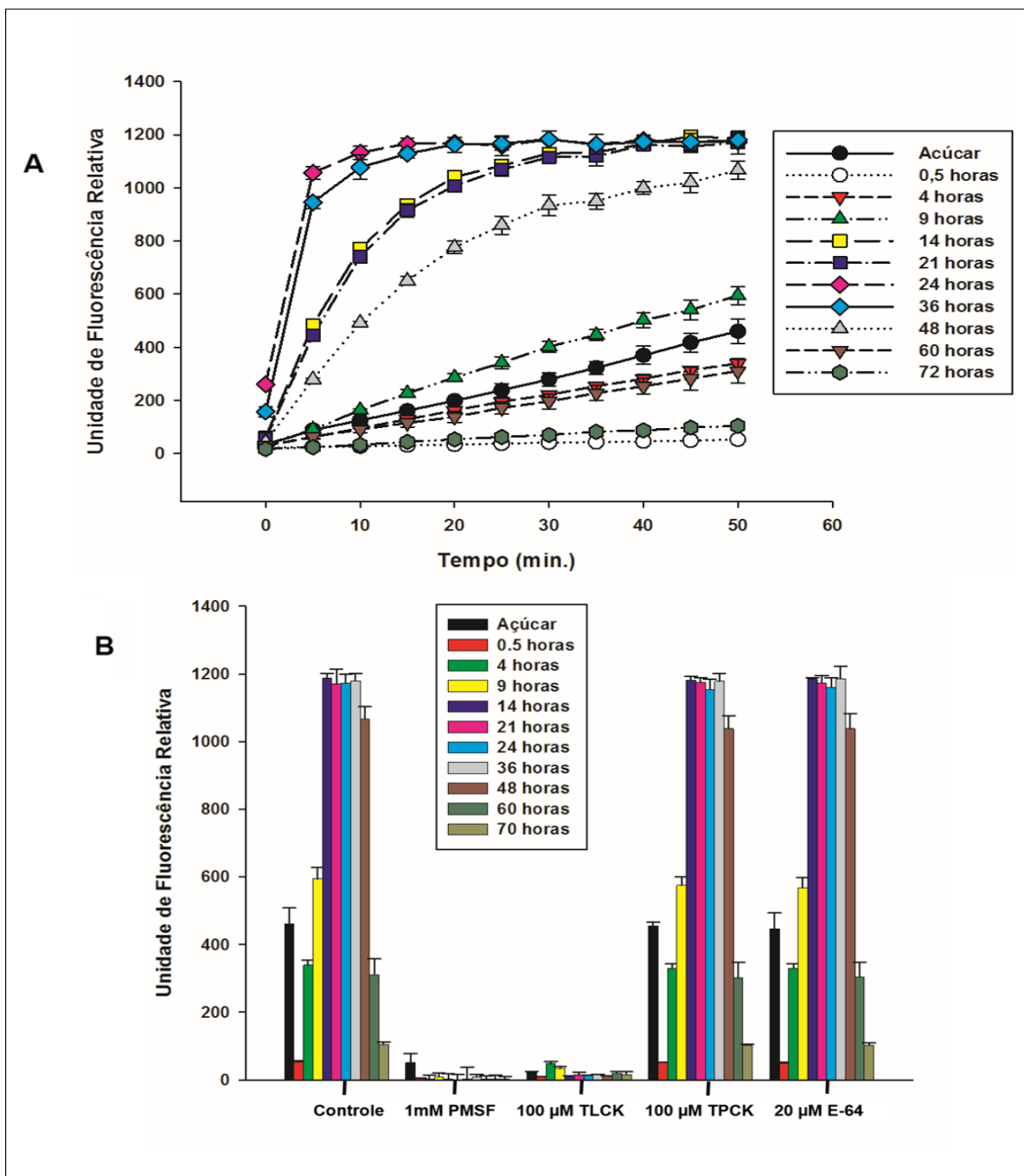


Figura 29. Análise da atividade proteolítica de intestino de fêmeas de *Aedes aegypti* alimentadas com sangue. Ensaio em solução **(A)**. Cinética da atividade proteolítica de intestino utilizando o substrato fluorogênico Z-Phe-Arg-MCA em tampão de Tris-HCl com pH 7,5. **(B)**. Ensaio de inibição da atividade em solução, utilizando o substrato fluorogênico Z-phe-arg- MCA. As amostras foram pré-incubadas por 30 min na presença de cada um dos seguintes inibidores: 1 mM de PMSF; 100 µM de TLCK, 20µM de E-64 e 100 µM de TPCK. A atividade foi iniciada com a adição do substrato e interrompida após 40 min. de reação com a adição solução de ácido acético 30%.

5. DISCUSSÃO

O estudo bioquímico e funcional das peptidases digestivas de invertebrados, como por exemplo, a caracterização bioquímica das atividades proteolíticas e compreensão da sequência de aminoácidos e estruturas tridimensionais, têm sido atravancadas principalmente por: (i) baixo rendimento das purificações das enzimas devido ao tamanho reduzido dos insetos; (ii) baixa transcrição, tradução e conseqüentemente abundância de algumas enzimas, dificultando a purificação; (iii) dificuldade em clonar e expressar enzimas ativas utilizando bactérias. Em mosquitos, o recente sequenciamento genômico de algumas espécies revelou a existência de um rico repertório de serina peptidases do tipo tripsina (SPtrips), corroborando as análises enzimáticas que indicavam as SPtrips como as principais responsáveis pela atividade proteolítica em intestino de mosquitos (Graf e Briegel, 1989; Kalhok et al., 1993; Barrillas-Mury e Wells, 1993; Müller et al., 1995; Barrillas-Mury et al., 1995). Entretanto, a demonstração de quais SPtrips são traduzidas em enzimas ativas é difícil. O aprimoramento das técnicas de biologia molecular, como o PCR e RNAi, tem possibilitado abordagens que conduzem a uma melhor compreensão do complexo processo de expressão e síntese dessas enzimas (Lu et al., 2006; Brackney et al., 2008; Wu et al., 2009; Isoe et al., 2009; Ráscon, et al., 2011). Em mosquitos, foi estabelecido que a expressão e síntese das SPtrips digestivas é bifásica e tem-se procurado entender como as variadas isoformas de tripsinas atuam no intestino. Contudo, os trabalhos recentes realizados com *Aedes aegypti*, lançando mão dessas novas abordagens, mostraram que a relação entre as isoformas de tripsinas digestivas parece não ser como postulado (Lu et al., 2006; Brackney et al., 2008; Isoe et al., 2009; Wu et al., 2009; Ráscon, et al., 2011). Ainda mais, uma tripsina descrita como a principal e mais abundante do processo digestivo de *Ae. aegypti*, parece ser uma serina colagenase e não uma Strip clássica. Este cenário confuso mesmo quando se trata do *Aedes aegypti*, mosquito cujas peptidases digestivas estão mais bem caracterizadas, mostra a necessidade de estudos que contemplem as peptidases. O estudo em diferentes espécies de mosquito também pode ajudar no entendimento acerca da natureza evolutiva destas enzimas. Assim, a utilização de abordagens de alta sensibilidade, como a enzimografia e a espectrometria de massas, possibilitou a

caracterização de SPtrips ativas, bem como a identificação de suas sequências e, indiretamente, dos genes que as codificam. Foi possível também identificar por espectrometria de massas centenas de proteínas solúveis presentes no intestino dos mosquitos na condição de alimentação com açúcar, gerando um perfil proteico que pode servir como base de comparação para diferentes condições experimentais das espécies aqui estudadas.

Neste trabalho, as análises enzimográficas do perfil proteolítico de intestino de fêmeas de *Cx. quinquefasciatus*, *An. albitalis* e *Ae. aegypti* alimentadas exclusivamente com solução 10% açucarada, revelaram a existência de um perfil enzimático reprodutível e composto por várias bandas de peptidases ativas. Entretanto, o número de bandas apresentado e a velocidade de degradação do substrato, foi diferente entre as espécies estudadas. O perfil proteolítico de *An. albitalis* foi completamente resolvido com 2 horas de reação e o das duas espécies de Culicíneos demorou 12 horas de reação para revelar o perfil mais completo.

Os ensaios de enzimografia e em solução para verificar a influência do pH do meio de reação sobre as atividades proteolíticas, mostraram baixas atividades proteolíticas em meios ácidos (pHs 3,5 e 5,5), enquanto em pH alcalino (pHs 7,5 e 10) as atividades foram maiores (Figuras 9, 17 e 23). Este alto desempenho proteolítico em meio alcalino, quando comparado ao meio ácido, condiz com a atividade destas enzimas no lúmen do intestino médio posterior dos mosquitos. É no intestino médio posterior que a maioria das enzimas digestivas são secretadas, e foi observado que em adultos de *Ae. aegypti* e *An. Gambiae*, o pH do intestino médio posterior varia de 8,5-9,5 e 8,0-9,5, respectivamente (Corena et al., 2005; Okuda, 2005).

Com respeito à sensibilidade a inibidores, a atividade proteolítica dos extratos de intestino das três espécies foi fortemente inibida por PMSF, um inibidor de tripsina e quimiotripsina, bem como por TLCK, um inibidor específico de tripsinas. Estes resultados indicam uma presença predominante de serina peptidases do tipo tripsina (SPtrip) no intestino de mosquitos, os quais estão de acordo com descrições prévias que relatam (usando outros métodos bioquímicos) a ocorrência de SPtrips em diversos mosquitos. As principais peptidases detectadas em outras espécies da ordem Diptera também pertencem à família das serina peptidases (Tabouret et al., 2003; Fazito-do-Vale et al., 2007; Pires et al., 2007, Cuervo et al., 2009). A presença dessas enzimas em intestino de fêmeas alimentadas com açúcar sugere que algumas

isoformas de peptidases são expressas constitutivamente, isto é, não é necessária a indução por um repasto sanguíneo. A atividade de tripsina observada por enzimografia, foi confirmada por ensaios em solução utilizando substrato fluorogênico Z-Phe-Arg-MCA, na presença ou na ausência de inibidores. Apesar de não se ter purificado e isolado as diferentes isoformas de tripsinas, o que permitiria ensaios de cinética enzimática mais robustos, o ensaio em solução corroborou os resultados enzimográficos em gel, evidenciando a existência de tripsinas ativas e possibilitou verificar as diferenças na velocidade de degradação do substrato fluorogênico entre as espécies estudadas. O uso em conjunto dessas duas metodologias permitiu uma análise quantitativa e qualitativa das atividades enzimáticas.

A expressão de peptidases em intestino de mosquitos alimentadas com açúcar pode estar relacionada com a digestão do meconium larval/pupal, ou terem sido induzidas pela presença de uma pequena quantidade de bactérias comensais que permanecem no intestino recém-formado após a metamorfose (Moll et al., 2001). Adicionalmente, a presença destas enzimas em condições de alimentação com açúcar pode ser explicada pelo fato dos mosquitos anautógenos necessitarem de uma alimentação com sangue para completar seu ciclo gonotrófico. Assim, a expressão constitutiva dessas enzimas seria uma preparação do intestino para uma eficiente digestão. De fato, genes que codificam para tripsinas são regulados negativamente em fêmeas anautógenas destinadas a diapausa. No fim do período de diapausa (após 2-3 meses a uma temperatura de 18 °C), a expressão de serina peptidases digestivas aumentam, preparando as fêmeas para a alimentação sanguínea (Robich e Denlinger, 2005). Dessa forma, a expressão constitutiva dos genes de tripsinas poderia garantir uma eficiente metamorfose e digestão.

As atividades proteolíticas de intestino de machos e fêmeas das três espécies estudadas apresentaram diferenças quantitativas e qualitativas intraespecíficas e interespecíficas. Os resultados mostraram que as fêmeas possuem um perfil com bandas de atividade mais intensas do que os machos. O tempo de reação para resolver o perfil proteolítico foi maior para os machos quando comparado com as fêmeas de uma mesma espécie. Os resultados de enzimografia indicaram que apesar do mesmo genoma, a fêmea expressa constitutivamente níveis mais altos e um perfil mais complexo de tripsinas do que o macho. Estas observações podem indicar que algumas peptidases envolvidas no processamento de proteínas do sangue sejam

expressas constitutivamente de forma basal em fêmeas e como os machos não fazem repasto sanguíneo, não necessitariam sintetizar um vasto repertório de SPtrips preparando o intestino para a digestão de uma grande carga de proteínas.

O tempo de reação de 12 horas, necessário para resolver o perfil das duas espécies de Culicíneos, se mostrou consideravelmente maior do que o tempo de reação de 2 horas necessários para que o perfil de atividade fosse completamente resolvido em *An. albitarsis*. O tempo de reação enzimática de 2 horas foi também observado para o *An. aquasalis* (Dias-Lopes et al., 2015). Por outro lado, trabalhos anteriores desenvolvidos pelo nosso grupo, têm revelado que o perfil proteolítico dos imagos de distintas espécies de culicídeos apresentam diferentes velocidades de degradação do substrato (Borges-Veloso et al., 2013; Saboia-Vahia et al., 2014; Dias-Lopes et al., 2015). No caso de *An. aquasalis*, as atividades são claramente resolvidas no tempo de 2 horas, ao passo que em *Ae. albopictus*, outro culicíneo, o tempo requerido é de 12 horas, similar ao observado neste estudo com as duas outras espécies de culicíneos.

Apesar da diferença na expressão de serina peptidases ser notável entre as espécies de mosquitos avaliadas neste e em outros estudos, a cinética da digestão do sangue tem se mostrado bastante similar. Foi observado que em *An. albitarsis*, *An. aquasalis*, *An. bellator* e *An. homunculus* apresentam um pico de atividade 12-24h após a alimentação com sangue. No tempo de 60 h após o repasto sanguíneo, nenhuma atividade proteolítica foi observada nessas espécies (Chege et al., 1996). Apesar de *Anopheles albitarsis* ter a maior atividade de tripsina, seguido pelo *An. aquasalis*, *An. bellator*, e *An. homunculus*, o desenvolvimento ovariano nessas espécies não foi afetado pelas diferenças nos padrões de atividade proteolítica, hemólise dos eritrócitos e desenvolvimento da membrana peritrófica (Chege et al., 1996). As diferenças observadas na intensidade das bandas proteolíticas dentro de uma mesma família (no presente trabalho e por outros autores), indicam que a regulação e síntese das tripsinas nesses mosquitos é diferente, mas o tempo necessário para realizar a digestão é bastante similar. Assim, os resultados aqui apresentados mostraram que o perfil proteolítico é espécie específico, sugerindo que o processo de expressão e síntese de peptidases é diferente entre as espécies das duas sub-famílias estudadas até o presente. Além de diferenças no tempo de reação, o perfil proteolítico revelado foi diferente entre as três espécies. Este resultado mostra

uma diferença qualitativa e quantitativa no perfil de atividade dessas peptidases entre as duas subfamílias de Culicídeos. A diferença na atividade proteolítica entre as amostras extraídas de mosquitos da subfamília Culicinae e Anophelinae pode ser dar devido principalmente a: (i) distintas cinéticas de atividade das isoformas de enzimas expressas pelas diferentes espécies, ou (ii) diferenças na quantidade de enzimas transcritas e traduzidas. A existência de complexos mecanismos de regulação da expressão das tripsinas em insetos poderia suportar estes dados (Barrillas-Mury et al., 1991; Kalhok et al., 1993; Noriega et al., 1996; Venâncio et al., 2009). Por exemplo, em fêmeas de *An. gambiae* da subfamília *Anophelinae*, quantidades consideráveis de RNAm da tripsina 1 (Antryp1) e do zimogênio ficam estocadas em vesículas de secreção nas células do epitélio intestinal, mesmo antes da alimentação com sangue. Essas vesículas, são liberadas no lúmen do intestino médio em resposta à distensão do intestino após a ingestão do alimento, independentemente de sua composição. Em um segundo momento, os produtos da digestão proteolítica (possivelmente peptídeos ou aminoácidos livres) atuam como sinal para a transcrição dos RNAm presentes e até mesmo transcrição de genes de outras enzimas digestivas importantes na digestão do sangue (Müller et al., 1993; Müller et al., 1995; Lemos et al., 1996). Já nas espécies de culicíneos, poucas vesículas secretoras são visualizadas nas condições de pré-alimentação e uma grande quantidade de RNAm é estocado (Hecker, 1977). Assim, após a alimentação, ocorre intensa atividade dos retículos endoplasmático rugoso devido a tradução dos RNAm que codificam as enzimas digestivas e as proteínas da membrana peritrófica. Apesar das diferentes estratégias de expressão de peptidases digestivas, a atividade dessas enzimas no intestino dos mosquitos de ambas subfamílias parece proporcionar uma eficiente digestão (Horler e Briegel, 1995).

A reprodutibilidade dos ensaios enzimográficos e a detecção da diferença na expressão de isoformas, em diferentes espécies de culicídeos, mostram que a enzimografia pode ser utilizada para a designação de espécies, algo análogo as técnicas de isoenzimas já utilizadas para diferenciar complexos de espécies como o *An. albitarsis* e *Culex pipiens* (Harbach et al., 2013). Desta forma, a partir da padronização de protocolos para análise enzimográfica e demonstração da estabilidade do perfil proteolítico, considerando uma mesma fase do ciclo de vida, poderia-se discriminar grupos taxonômicos indiferenciáveis por outras técnicas.

Entretanto, análises com outras espécies destes complexos, seus híbridos e demais espécies próximas de mosquitos são necessárias para a comprovação desta hipótese.

Apesar da expressão constitutiva de SPtrip ter sido verificada em mosquitos (Yang e Davies, 1971; Briegel e Lea, 1975; Müller et al., 1995) muitos trabalhos não observaram a presença de atividade proteolítica em intestino de mosquitos na condição de alimentação com açúcar, como o obtido no presente estudo (Berner et al., 1983; Billingsley e Hecker, 1991; Billingsley e Rudin, 1992). Esse fato se deve provavelmente à (i) baixa abundância das enzimas; (ii) limitações das técnicas disponíveis; (iii) idade dos mosquitos utilizados nos ensaios. De fato, com relação à idade dos mosquitos, a expressão de tripsinas parece aumentar após os adultos emergirem das pupas, atingindo os níveis mais altos após 5 dias (Lemos et al., 1996). No presente trabalho, os insetos adultos utilizados possuíam 4-5 dias de idade, o que pode ter possibilitado a identificação do perfil enzimático. O uso da enzimografia para a detecção de atividade proteolítica, como levado a cabo neste trabalho, tem se mostrado uma eficiente técnica para explorar o repertório de peptidases ativas. Apesar dos trabalhos para caracterizar a atividade de tripsinas em intestino de várias espécies de culicídeos, a enzimografia não foi utilizada para a caracterização bioquímica e estudo comparativo da expressão dessas peptidases em intestino de machos e fêmeas destas espécies. Através da enzimografia é possível: (i) avaliar a massa molecular aparente das peptidases ativas presentes; (ii) detectar presença de distintas enzimas e de suas isoformas; (iii) analisar a expressão diferencial de peptidases entre espécies, estágios evolutivos e em órgãos específicos; (iv) fazer análise quantitativa relativa do perfil proteolítico; (v) verificar a estabilidade do perfil proteolítico, i. e., verificar se a expressão do perfil de peptidases de uma espécie é constitutiva independente das condições ambientais; e (vi) observar a expressão de peptidases ativas, i. e., o que está sendo realmente expresso e provavelmente funcional (Zhao e Russell, 2003; Cuervo et al., 2008; De Jesus et al., 2009; Mesquita-Rodrigues et al., 2011).

A presença de um vasto repertório de SPtrip em mosquitos tem sido relacionada a necessidade da digestão de uma grande quantidade de proteínas presente no sangue. Vários dípteros hematófagos expressam uma série de tripsinas constitutivas e outras são expressas apenas após a alimentação (Tabouret et al., 2003; Fazito-do-Vale et al., 2007; Pires et al., 2007, Cuervo et al., 2009). Nos ensaios

realizados aqui com *Ae. aegypti* alimentados com sangue, foi verificado que o perfil proteolítico dessas fêmeas é diferente daquelas que se alimentaram exclusivamente com açúcar. O curso temporal de atividade enzimática e os ensaios em solução revelaram que a atividade enzimática foi extremamente alta após a alimentação com sangue. O tempo de 12 horas, antes utilizado para resolver o perfil proteolítico de fêmeas alimentadas com açúcar, foi muito alto para as amostras de intestino PAS, causando a degradação de toda gelatina do gel. Assim, o tempo de 2 horas de reação foi o suficiente para resolver o perfil e observou-se que algumas bandas de atividade proteolítica na região de 17 kDa foram as primeiras a surgirem após 14 horas PAS. Nos tempos de PAS seguintes (21, 24 e 36 horas), surgiram as bandas de alto peso molecular – na faixa de 31-38 kDa e as de 24 kDa. As bandas de atividade na faixa de 17 kDa não foram observadas em fêmeas alimentadas com açúcar, sugerindo um papel importante dessas isoformas na fase inicial da digestão sanguínea. De fato, tem sido sugerido que as enzimas que atuam nas fases iniciais da digestão, devem possuir um tamanho pequeno o suficiente para atravessar os poros da membrana peritrófica e atuarem na luz do intestino, dentro do bolo alimentar (Terra et al., 1994). Para os ensaios em solução, as amostras de intestino de fêmeas PAS foram diluídas dez vezes para a realização dos experimentos, visto que as atividades das amostras não diluídas eram tão altas que consumiam todo substrato antes que se iniciasse a leitura no espectrofluorímetro. O uso de inibidores nos ensaios enzimográficos em gel e em solução, confirmaram que as enzimas ativas são SPtrip. De fato, a regulação pós-transcricional e pós-traducional da expressão de peptidases tem sido bem descrita em fêmeas adultas de *Ae. aegypti* e *An. gambiae* após ingestão de sangue (Felix et al., 1991; Noriega et al., 1996; Noriega e Wells, 1999). A síntese e secreção de tripsina em adultos podem ser afetadas por fatores externos, como alimentação, ou por fatores internos, como liberação de hormônios (Borovsky, 2003). Foi demonstrado em *Ae. aegypti* que: (i) hormônios secretados pelos ovários aumentam (ecdisona) ou diminuem (TMOF) a secreção de peptidases; (ii) células neurosecretoras contêm fatores que aumentam a secreção de tripsinas (Briegel e Lea, 1975). A decapitação e/ou retirada dos ovários, logo após a metamorfose do adulto, reduz pela metade a síntese de tripsina, comprovando a influência dos hormônios sobre a taxa de síntese dessa enzima (Graf et al., 1998). Adicionalmente, foi observado que o hormônio juvenil (HJ) regula a síntese de uma tripsina expressa em adultos denominada *early trypsin*

(Noriega et al., 1996 a, Noriega et al., 1996 b; Noriega e Wells, 1999). Esta enzima é sintetizada em baixa quantidade, aparece no intestino após uma hora de ingestão de sangue e desaparece entre 6-8 horas após. Embora a expressão natural de “early trypsin” não tenha sido detectada em machos, pupas e larvas de *Ae. aegypti*, tem sido demonstrado que a expressão de *early trypsin* pode ser induzida em pupa, na presença de análogos do HJ, muito embora a pupa não se alimente (Noriega e Wells, 1999). Em relação às tripsinas expressas em intestinos, em *An. gambiae* foram descritos sete genes em um cluster no cromossomo 3R que codificam cinco proteínas funcionais (Müller et al., 1995). Destas, foi verificado que as tripsinas 1 e 2 são induzidas pelo repasto sanguíneo. Enquanto as tripsinas 3, 4 e 7 são expressas constitutivamente, sendo que a 3 e 7 são reprimidas após o repasto sanguíneo (Müller et al., 1993; Müller et al., 1995). Foi identificado um motivo de 12 nucleotídeos presentes em todos os promotores desses genes, designado elemento regulatório de tripsinas putativas (PTRE, putative trypsin regulatory element). No entanto, diferenças na região reguladora entre os genes expressos nas primeiras horas PAS e os que são expressos tardiamente, garantem uma expressão diferenciada dos genes (Giannoni et al., 2001). Em anofelinos neotropicais *An. aquasalis*, *An. darlingi* e *An. albitarsis* foi verificada a presença de duas tripsinas induzidas pela alimentação sanguínea (Caroci et al., 2003). Apesar das SPtrip já serem descritas como as enzimas majoritárias em intestino de algumas espécies de mosquitos, e as condições utilizadas neste estudo serem padronizadas para a detecção de atividade de SPtrip, não se pode excluir a possibilidade de outras classes de peptidases atuarem no intestino.

O genoma de *Cx. quinquefasciatus* codifica para 403 putativas serina peptidases do tipo tripsina e o genoma do *Ae. aegypti* codifica para 380 putativas tripsinas. O genoma do *An. albitarsis* ainda não foi sequenciado, mas o do *An. gambiae* possui 305 putativos genes (Wu et al., 2009). Apesar da existência de um variado repertório de genes que codificam para estas enzimas, não se sabe ao certo, a quais tripsinas estão presentes em um determinado tecido e/ou em que condições fisiológicas são ativas. Utilizando a espectrometria de massas, seguida de duas ferramentas de identificação, Mascot (seguido da validação pelo Scaffold) e COMET (seguido da validação pelo SEpro), foi possível identificar diferentes tripsinas nos extratos de intestino de *Cx. quinquefasciatus* e *An. albitarsis* alimentados com açúcar,

validando a observação de um complexo perfil proteolítico nos ensaios enzimográficos.

Nas amostras de *Cx. quinquefasciatus* foi possível identificar sete serina peptidases do tipo tripsina, utilizando o COMET e o Mascot. Adicionalmente, uma serina peptidase do tipo tripsina foi identificada exclusivamente pelo Mascot, totalizando oito serina peptidases identificadas. Todas as peptidases foram identificadas por homologia com peptidases da própria espécie *Cx. quinquefasciatus*, já que esta possui o genoma sequenciado. Já nas amostras de *An. albitarsis*, foi possível identificar sete SPtrip, utilizando ambas as ferramentas de identificação. Uma proteína não caracterizada, mas com sequência similar à de peptidases ativas, foi exclusivamente identificada pelo mascot; duas outras serina peptidases foram identificadas exclusivamente pelo COMET, totalizando 10 peptidases identificadas para esta espécie. Em *An. albitarsis*, cujo genoma ainda não é conhecido, os peptídeos identificados apresentaram homologia principalmente com as peptidases de anofelinos neotropicais. Oito foram identificadas por homologia com *An. aquasalis* e duas por homologia com *An. darlingi*. Interessantemente, as bandas de SDS-PAGE de onde as peptidases foram extraídas e identificadas por MS/MS, coincidem com as regiões dos halos de atividade dos ensaios de enzimografia. Por exemplo, em *Cx. quinquefasciatus* a maioria das peptidases identificadas por MS/MS em bandas eletroforéticas SDS-PAGE migrando de 25-40 kDa, com excessão da Tripsina5 e Tripsina7 que foram as únicas encontradas nas regiões de alta massa molecular (acima de 100 kDa). Estas duas enzimas exibem sítios preditos de N-glicosilações, e a Tripsina5 também apresenta sítios preditos para O-glicosilações e domínios transmembrana. Estas características poderiam alterar o padrão de migração das proteínas maduras, justificando o padrão de migração observado em gel. Entretanto, como a preparação das amostras para SDS-PAGE de proteínas totais é diferente da enzimografia em gel, a comparação da migração das bandas fica dificultada, ainda que ambos sirvam para o mapeamento das peptidases identificadas.

O alinhamento das sequências completas de aminoácidos das peptidases identificadas por MS/MS, permitiu analisar as sequencias para a caracterização das peptidases no intestino de mosquitos alimentados com açúcar. O alinhamento revelou diversas características estruturais típicas de tripsinas peptidases digestivas de invertebrados como: (i) os resíduos conservados de histidina, ácido aspártico e serina

formando a tríade catalítica; (ii) resíduos de cisteínas responsáveis por formar as pontes dissulfeto, (iii) presença do peptídeo sinal, (iv) os putativos motivos de ativação auto catalítica imediatamente após um resíduo de arginina ou lisina (R/K-IVGG); (v) os motivos característicos de peptidases ativas LTHAAC, DIAL e GDSGGP. Algumas tripsinas identificadas apresentaram características distintas, como por exemplo, a região auto catalítica da Tripsina4 possui um resíduo de His no lugar do R/K, o que sugere que essa enzima possui uma ativação diferenciada do peptídeo sinal. Adicionalmente, os motivos de ativação na Tripsina5 (IIGG), e da Tripsina Catiônica (VVGG), diferem por um aminoácido quando comparadas à clássica sequência (IVGG) das tripsinas de invertebrados (Muhlia-Almazán et al., 2008). Uma importante diferença entre as tripsinas de vertebrados e invertebrados é o número e o local das pontes dissulfeto. As tripsinas de vertebrados possuem, normalmente, seis pontes dissulfeto, ao passo que as tripsinas de invertebrados apresentam apenas três pontes dissulfeto conservadas, geralmente próximas aos resíduos do sítio ativo (Muhlia-Almazán et al., 2008). O alinhamento múltiplo mostrou que cinco das tripsinas de *Cx. quinquefasciatus*, Tripsina1, Tripsina2, Tripsina4, Serina peptidase SPD24D e Serina peptidase ½ possuem três pontes dissulfeto, a TripsinaBer5 e Tripsina7 possuem quatro e a Tripsina Cationica apenas uma ponte dissulfeto. Apesar do número de pontes dissulfeto ser diferente nas tripsinas identificadas, o papel da ponte dissulfeto é crucial para a estrutura tridimensional das enzimas, e conseqüentemente, para suas atividades (Kalhok et al., 1993; Muhlia-Almazán et al., 2008; Ráscon et al., 2011).

Com o objetivo de entender como as tripsinas e quimiotripsinas são ativadas, a região do sítio de ativação do zimogênio, composta pelas três posições imediatamente anteriores ao sítio de clivagem, foi analisada neste trabalho. Os tripsinogênios de vertebrados possuem, de uma forma geral, uma sequência no sítio de clivagem composta por um resíduo de K (carregado positivamente) em P1, seguido imediatamente por quatro resíduos de D (carregado negativamente). Em alguns organismos, este processo de ativação do tripsinogênio é finamente regulado por diferentes enzimas conhecidas como enteroquinases (Kitamoto et al., 1994; Chen et al., 2003). A regulação da ativação das tripsinas por enzimas especializadas, mesmo sendo mais lento na ativação, traz como benefício o decréscimo no risco de efeitos indesejáveis sobre alvos inespecíficos, devido ao excesso de tripsinas ativas (Kitamoto et al., 1994). Mutações nos tripsinogênios humanos na região do sítio de

clivagem D por G, em P2, e K por R, em P1, estão associadas com doenças relacionadas a auto ativação de tripsinas como, por exemplo, pancreatites (Ferec et al., 1999; Teich et al., 2000). Dentre as sequências de tripsinas de mosquito, analisadas neste trabalho, foi possível observar que cinco tripsinas possuem uma R em P1, duas possuem uma K e apenas uma apresenta um resíduo de H. Um achado interessante é que apenas resíduos carregados positivamente foram observados nessa região P1, o que é condizente com a auto ativação dessas enzimas por tripsinas, as quais clivam preferencialmente aminoácidos carregados positivamente, com preferência pela R, seguida de K e H. Na região do P2 e P3 não foram observados um consenso entre as tripsinas identificadas aqui, mas predominou-se resíduos de aminoácidos hidrofóbicos, carregados ou não, na região P2 e P3. Com exceção da tripsina BOXES8, que apresentou um resíduo de glutamato carregado negativamente, nenhum outro resíduo negativo foi encontrado nessa região de ativação do zimogênio de tripsinas de mosquitos, como ocorre em vertebrados. Nas tripsinas identificadas em *Cx. quinquefasciatus*, o motivo DDDDK conservado (típico de tripsinas de vertebrados ativadas por enteroquinases), não foi observado. Porém, a composição de resíduos observada na região de auto ativação nas tripsinas de *Cx. quinquefasciatus*, é semelhante às sequências motivo de tripsinas auto ativadas em humanos (Gly em P2 e Arg em P1). Estas observações sugerem que as SPtrip de mosquitos são auto ativadas por tripsinas (Ferec et al., 1999; Teich et al., 2000). Diferenças nas regiões de ativação do zimogênio, bem como de reconhecimento dos substratos, permitem variações entre as interações dessas regiões, levando a diferentes taxas de ativação entre elas. De um modo geral, exceto pela posição P1, a região de ativação consensual de tripsinas de culicídeos é variada (Ráscon et al., 2011). Assim, as análises do motivo de ativação do tripsinogênio, realizadas aqui, mostraram que as tripsinas de *Cx. quinquefasciatus* parecem ser mais propensas a apresentar auto ativação do que as enzimas de mamíferos (Ferec et al., 1999; Teich et al., 2000). O fato de nenhuma enzima semelhante a enteroquinase de mamíferos ter sido identificada em insetos, corrobora a hipótese da auto ativação de tripsinas de mosquitos. Além disso, o sítio de ativação do tripsinogênio de *Cx. quinquefasciatus* apresenta características semelhantes às aquelas observadas nos sítios de ligação ao substrato das tripsinas dessa espécie de mosquito, ressaltando a ideia de que a auto ativação deve ocorrer nessa família.

O presente estudo mostra que as tripsinas em mosquitos têm claramente seguido um caminho evolutivo peculiar, quando comparadas com os seus homólogos em mamíferos. Os resultados suportam a hipótese de que as tripsinas são enzimas chaves na adaptação de insetos aos seus distintos hábitos alimentares (Terra e Ferreira, 1994; Muhlia-Almazán et al., 2008). Estas enzimas estão envolvidas no processo de ativação de outras peptidases e delas próprias, proporcionando assim o rápido suprimento de enzimas funcionais. A adaptação a esses hábitos alimentares pode ter conduzido esses insetos a possuir um grande número de enzimas diversificadas, necessárias para o processo digestivo, as quais podem oferecer, por um lado, auto ativação, como uma estratégia simples para aumentar a velocidade de ativação e, por outro lado, oportunidades para escapar dos inibidores ou aumentar a eficiência frente a diferentes afinidades aos substratos (Terra e Ferreira, 1994; Muhlia-Almazán et al., 2008).

As SPtrip identificadas em *Cx. quinquefasciatus* foram também analisadas quanto a predição de localização celular, presença de hélices transmembrana e de sítios de glicosilações. A análise de predição celular, utilizando o Target P, indicou que todas as tripsinas identificadas são secretadas, condizendo com a típica função das peptidases digestivas que atuam no lúmen do intestino médio (Muhlia-Almazán et al., 2008; Venâncio et al., 2009; Wu et al., 2009). Entretanto, a predição de hélices transmembrana, utilizando o TMHMM, revelou que a Tripsina4 e Tripsina5 possuem um domínio transmembrana, sugerindo que algumas peptidases do tipo tripsina possam estar aderidas às membranas. Apesar de glicosilações não serem modificações pós- traducionais típicas de tripsinas, alguns motivos glicosilados têm sido observados em tripsinas de invertebrados (Dinglasan et al., 2009). Cinco, das oito tripsinas identificadas aqui, apresentaram sítios preditos para O- e N-glicosilações. A glicosilação destas enzimas pode estar relacionada com a associação destas com a membrana peritrófica, uma vez que tripsinas com esta característica foram encontradas em membranas peritróficas de *An. gambiae* (Dinglasan et al., 2009).

Utilizando o VectorBase, foi possível analisar os genes que codificam para as SPtrip em *Cx. quinquefasciatus*, e observar como eles estão organizados no genoma. Verificou-se que o número de exon dos genes que codificam para as tripsinas variou de um a três. Com exceção da Serina peptidase 1/2, que possui um intron com 298 nucleotídeos, as demais sequencias dos introns são menores do que as observadas

nos genes de tripsinas de vertebrados, variando de 25 a 71 nucleotídeos. As análises mostraram que as estruturas dos introns/exons não são conservadas nessas tripsinas, sugerindo que diversos eventos de perda e/ou ganho de introns pode ter ocorrido nesta espécie, o que vai ao encontro do observado para outras espécies (Craik et al., 1984; Roach et al., 1997; Klein et al., 1998; Wang et al., 1999; Bao et al., 2014). Estas análises revelaram também a existência de 13 a 38 parálogos destas enzimas distribuídos pelo genoma do inseto. Por exemplo, a Tripsina1, 2, 4, 7, SP24D e Tripsina Cationica, são parálogas entre elas, sugerindo que estas peptidases possam ter sido originadas por eventos de duplicação gênica (Müller et al., 1993; Wang et al., 1999; Ross et al., 2003). Além disso, verificou-se aqui que os genes que codificam para diferentes SPtrip em *Cx. quinquefasciatus*, são encontrados próximos uns dos outros no genoma, formando “*clusteres*”. Por exemplo, de acordo com o VectorBase, o gene da Tripsina1 possui em sua vizinhança cinco outros genes que codificam para tripsinas. Foi sugerido que o ancestral dos dípteros possuía apenas um gene de tripsinogênio e que cópias extras foram adquiridas por duplicação genica (Wu et al., 2009). Em Culicídeos, muitos genes que codificam para SPtrip formam “clusters”, em diferentes cromossomas, indicando que a duplicação genica desempenha um importante papel na expansão e diversificação desta família de genes (Müller et al., 1993; Wu et al., 2009). O genoma de *Cx. quinquefasciatus* possui um grande repertório de genes que codificam para SPtrip (Wu et al., 2009). Tal repertório, poderia estar relacionado com a habilidade deste inseto de processar sangue obtido de diferentes fontes alimentares quando imagos, e também para a habilidade de se alimentarem em criadouros com alta quantidade de microrganismos quando na fase de larvas. De fato, esta espécie possui uma alta plasticidade no seu comportamento alimentar, sendo capaz de realizar o repasto sanguíneo em humanos, cães, aves e gado (Farajollahi et al., 2011; Kilpatrick et al., 2012; Reddy et al., 2012; Simpson et al., 2012).

Apesar das ferramentas de biologia molecular permitirem um avanço no conhecimento dos vários genes que codificam para tripsinas, a confirmação da presença dessas enzimas, e em diferentes condições biológicas, tem sido pouco verificada. Isso pode se dever pela dificuldade de se obter quantidades dessas diferentes isoformas para a realização dos ensaios de caracterização bioquímica. Com o objetivo de padronizar condições experimentais para identificação de tripsinas

específicas de *Cx. quinquefasciatus* por SRM, utilizou-se o SKYLINE para verificar a ocorrência de peptídeos proteotípicos nas SPtrip identificadas por MS/MS. O SRM é uma metodologia poderosa utilizada em estudos proteômicos de hipótese-dirigida para monitorar peptídeos alvos em uma amostra complexa de proteína (Prakash et al., 2009; Gallien et al., 2015). Apesar da ocorrência de motivos conservados nas sequências de peptídeos identificadas em *Cx. quinquefasciatus*, o SKYLINE revelou que cinco das oito tripsinas identificadas em *Cx. quinquefasciatus* possuem diferenças em suas sequências de aminoácidos que permitem a detecção de peptídeos únicos. Além disso, esses peptídeos únicos foram os mesmos identificados pelo espectrômetro de massas em nosso estudo, indicando que a metodologia de SRM pode ser aplicada para a detecção de peptídios específicos de tripsinas em amostras complexas de proteínas. O banco de dados utilizado para a análise no SKYLINE contempla todas as outras espécies de culicídeos sequenciados até janeiro de 2015. Assim, essas identificações de peptídios proteotípicos feita pelo SKYLINE utilizando o banco de dados com várias espécies de culicídeos, indicam que tais peptídeos proteotípicos são peptidase específico e espécie específico. Apesar desses resultados não serem conclusivos, pois os genomas não estão bem anotados e muitas espécies de mosquitos ainda não possuem genoma sequenciado, a possibilidade de identificar as peptidases usando essa metodologia poderia auxiliar na identificação de peptidases de culicídeos, inclusive aquelas de baixa abundância, que por ventura estejam presentes em amostras complexas, como por exemplo, no extrato proteico de intestino de mosquitos alimentados com sangue.

Estudos anteriores realizados pelo nosso grupo, juntamente com os desenvolvidos neste trabalho, revelaram que o perfil proteolítico em gel copolimerizado com gelatina é singular para cada amostra de intestino das espécies até então verificadas. Ainda mais, este perfil proteolítico parece ser estágio, tecido e sexo específico (Borges-Veloso et al., 2012; Saboia-Vahia et al., 2014; Dias-Lopes et al., 2015). A caracterização do perfil proteolítico por eletroforese semi-desnaturante acoplada a identificação das proteínas por LC-MS/MS após separação por SDS-PAGE, mostrou ser uma eficiente estratégia para o mapeamento de peptidases ativas. As análises de bioinformática mostraram que apesar das tripsinas de *Cx. quinquefasciatus* identificadas por MS/MS possuírem características clássicas das tripsinas digestivas de invertebrados, a detecção de diferenças nas sequências de

aminoácidos entre essas enzimas sugere que devam possuir diferentes habilidades para substratos e inibidores. Esta abordagem (enzimografia para detecção de peptidases ativas e espectrometria para a identificação das enzimas) parece ser de alta sensibilidade e pode auxiliar no entendimento do complexo mecanismo de síntese e secreção de tripsinas em mosquitos, e resolvendo, com isso, as controversas existentes no conhecimento sobre essas enzimas. A espectrometria de massas é também importante para que se possa identificar qual proteína está de fato sendo traduzida em uma determinada condição biológica. A identificação da sequência de aminoácidos das peptidases é essencial para o conhecimento profundo acerca da fisiologia do sistema digestivo dos mosquitos. Por exemplo, recentemente, o grupo do Dr. Walter Terra vem tentando desenvolver inibidores para SPtrip digestivas que possam interromper a digestão dos mosquitos. Este grupo procura, dentre as diferentes tripsinas, aquelas que apresentam características singulares e possam desempenhar um papel diferenciado. Estes estudos baseiam-se em modelagem molecular 3D e dependem da anotação acuradas de sequências de aminoácidos de tripsinas. Assim, estudos como os realizados aqui podem auxiliar no desenvolvimento de metodologias de controle de vetores.

Infelizmente, não logramos até o momento, assinar a identidade das tripsinas presentes no intestino de fêmeas de *Ae. aegypti* alimentadas com açúcar ou sangue por MS/MS. Entretanto os resultados produzidos concordam com a literatura revisada no sentido de que várias isoformas de tripsinas devem ser expressas no intestino de fêmeas de mosquitos e que grande parte delas é induzida após a alimentação com sangue. Tanto as diferenças no tempo de reação dos ensaios de enzimografia como as cinéticas em solução mostram que após a alimentação com sangue a atividade enzimática é muito maior nesta espécie. A tomada do sangue pela fêmea acaba introduzindo no intestino uma grande quantidade de proteínas do sangue, não só aumentando a complexidade da amostra como também tronando as proteínas do intestino minoritárias neste sistema. Este modelo experimental dependente do repasto sanguíneo dificulta os experimentos de espectrometria de massas, ainda mais para as proteínas secretadas presentes no suco digestivo. Os experimentos de proteoma que procuraram explorar diferenças da composição proteica do intestino de mosquitos em diferentes condições biológicas, inclusive com sangues infectados, pecam ao lavar o conteúdo intestinal no procedimento de preparação das amostras. Apesar deste

procedimento ter como objetivo retirar a maior quantidade possível do sangue (e facilitar a identificação de proteínas do intestino), acaba também por remover grande parte das proteínas secretadas para o lúmen deste órgão, dentre elas, as peptidases que podem ter papel fundamental em diversos processo fisiológico importante.

Finalmente, além da caracterização bioquímica e dos estudos comparativos para identificação das peptidases nesses insetos, foi um objetivo deste trabalho, identificar as proteínas totais presentes no intestino de fêmeas adultas de *Cx. quinquefasciatus* e *An. albitarsis* alimentadas com açúcar. Para compreender o papel das proteínas envolvidas nos complexos processos biológicos que ocorrem no intestino, como por exemplo, as proteínas envolvidas na defesa contra os parasitos e resistência a inseticidas, é essencial caracterizar o perfil de proteínas expressas no intestino das fêmeas. Para as análises proteômicas do intestino foi utilizada uma base de dados de Culicidae obtida do UNIPROT. Na abordagem por PSM, é imprescindível um banco de dados, e as sequências do organismo de estudo são fundamentais para as identificações. Em *Cx. quinquefasciatus*, utilizando o Mascot como ferramenta de busca e identificação das proteínas por PSM, seguida do Scaffold para a validação estatística das identificações, foi possível identificar 1397 proteínas em 1090 grupos de proteínas. Já por intermédio do COMET, seguida da validação pelo SEPro, logramos identificar 1183 proteínas, sendo 942 proteínas únicas distribuídas em 1034 grupos. O *Cx. quinquefasciatus* possui um repertório de 18,968 genes que codificam para proteínas (Arensburger et al., 2010). Assim, nesse estudo, mostrou-se que o intestino contém pelo menos 7,5% das proteínas codificadas pelo genoma. A lista de identificações realizadas pelo Mascot foi então analisada utilizando a ferramenta do BioMart, disponível no VectorBase. Foi possível acessar a ontologia gênica das proteínas identificadas de acordo com os termos anotados para cada proteína e depositadas no UNIPROT. Foram observadas 201, 313 e 81 diferentes termos de processos biológicos, funções moleculares e componentes celulares, respectivamente. Proteínas com atividade oxirredutase, importantes na detoxificação de espécies reativas de oxigênio, também foram identificadas no intestino de *Cx. quinquefasciatus* alimentadas com açúcar. De fato, transcritos de diversas enzimas que participam do metabolismo redox e da detoxificação, após o repasto sanguíneo, foram encontradas no intestino de fêmeas de *Anopheles stephensi* alimentadas com açúcar. Essas enzimas são reguladas positivamente após a ingestão do sangue e

quando ocorre infecção (Patil et al, 2009). Durante o repasto sanguíneo, a lise de hemoglobinas gera grandes quantidades de heme, que são moléculas tóxicas e podem agir como pró-oxidante gerando radicais livres. Os mosquitos podem agregar o heme junto à matriz peritrófica diminuindo o estresse oxidativo (Devenport, 2006). Além disso, estes insetos podem equilibrar o desbalanço redox através de enzimas e moléculas antioxidantes (Graça-Souza, 2006).

O intestino é o primeiro órgão do inseto a entrar em contato com os patógenos durante sua jornada através do inseto até sua transmissão a um novo hospedeiro vertebrado. Acredita-se que no intestino os patógenos estão vulneráveis. A chegada a este novo ambiente envolve mudanças abruptas nas condições de pH e temperatura, que pode causar danos aos patógenos e impedir a disseminação para outros tecidos, visto que estão em pequeno número e ainda não se replicaram. Este novo ambiente rico em peptidases é degradativo, além de outras enzimas digestivas e moléculas que causam stress oxidativo (Billingsley e Lehane, 1996; Shahabuddin et al., 1996; Terra et al., 1996). A discriminação de proteínas constitutivas ou induzidas tem um importante papel no desenvolvimento de métodos para controle dos patógenos (Dinglasan et al., 2007; Ubaida et al., 2012). Durante décadas, esforços têm sido empenhados para identificar quais moléculas estariam envolvidas na interação mosquito/parasito. Os tecidos onde ocorrem o desenvolvimento dos parasitos são os principais alvos dos estudos proteômicos. Vários trabalhos de interação mostram como o patógeno pode alterar a expressão protéica do hospedeiro invertebrado, por exemplo: em intestinos de *An. albimanus* infectados com *Plasmodium berghei* (Serrano-Pinto et al., 2010); em glândula salivar, intestino e células C6/36 de *Ae. albopictus* infectados com o vírus da dengue (DENV-2) (Zhang et al., 2013); em intestino de *Ae. aegypti* infectado com o vírus da dengue (DENV-2) e com chikungunya (CHIKV) (Tchankouo-Nguetcheu et al., 2010); em larvas de *A. aegypti* após infecção com o microsporídeo parasita, *Vavraia culicis* (Biron et al., 2005) e após infecções ou coinfeções com *Edhazardia aedis* e *Vavraia culicis* (Duncan et al., 2012). No presente estudo, procuramos verificar, dentre as proteínas indentificadas por espectrometria, a presença de proteínas relacionadas com a interação entre parasitos e vetores. Identificamos a presença no intestino de *Cx. quinquefasciatus* as famílias de receptores de reconhecimento de padrões como lectinas do tipo C (B0WID9, B0WIF2, B0WIE3, B0WID7), lectinas ligadoras a manose

(B0XG12) e galactosamida (B0XEF4, B0W0J6, B0X1F3). Identificamos também proteínas de membrana, ricas em leucina (B0WPC7, B0X0G6), e uma SRP68 (B0WWMF9). As lectinas, encontradas em todos os organismos, são proteínas que se ligam a carboidratos com alta especificidade, mas a afinidade com que estas proteínas interagem com carboidratos é variada. Elas participam de vários processos de reconhecimento celular, sinalização e adesão (Liu et al., 2014). Em anofelinos, as lectinas se ligam a carboidratos específicos dos oocinetos, os quais seriam utilizados como receptores, inibindo a adesão dos oocinetos às células epiteliais do intestino. Além das proteínas reconhecedoras de padrões, identificamos também outras proteínas como a arginine quinase (B0WLS0), proteína F-actina (B0WTE9), NADPH-citocromo P450 redutase (B0X829), glutathione-s-transferase theta (B0XGJ5, B0XGK3, B0X3C8) e enolase (B0W1N4), as quais já foram descritas participando da interação parasito vetor (Sreenivasamurthy et al., 2013; Liu et al., 2014).

Neste trabalho, apesar de não termos detectado atividade de outras enzimas proteolíticas nas análises bioquímicas, foram identificadas por MS/MS outras peptidases. Dentro da família S1 das serina peptidases nós identificamos três diferentes quimiotripsinas (BOW653, BOW174, BOWH74). A presença de quimiotripsinas em *Cx. quinquefasciatus* foi descrita como tendo uma atividade máxima há 36 horas após a alimentação sanguínea (Okuda et al., 2002). Em *An. darlingi* e *An. aquasalis*, foram descritos dois genes de quimiotripsinas relacionadas com a digestão do sangue (Caroci et al., 2003). Em *An. aquasalis*, a atividade dessas enzimas foi observada 24h após a ingestão do sangue (de Almeida et al., 2003). Em *An. gambiae*, são 3 genes descritos: Anchym1 e Anchym2 que têm sua transcrição regulada positivamente após alimentação sanguínea (Vizioli et al., 2001), enquanto os níveis de RNAm de AgChyL decai entre 12 e 24 horas após a ingestão do sangue e após 48 horas volta a ser abundante novamente (Shen et al., 2000). Em *Ae. Aegypti*, um gene de uma quimiotripsina específica de fêmeas é transcrito após 24 horas da metamorfose dos adultos, mas sua tradução só ocorre após o repasto sanguíneo (Jiang et al., 1997). O fato de identificarmos estas enzimas por espectrometria e não observarmos atividade nos ensaios bioquímicos, pode ser explicado pelo fato destas enzimas serem traduzidas, mas não serem ativadas. Os zimogênios de quimiotripsinas de *An. gambiae*, Anchym1 e Anchym2, necessitam de ativação trípica para ativação,

porém, *in vitro*, sua exposição prolongada à ação das tripsinas pode degradá-las, persistindo apenas na forma de zimógenos (Müller et al., 1995; Vizioli et al., 2001).

Apesar das serina peptidases serem historicamente as endopeptidase mais abundantes descritas em mosquitos, foi identificado em *Cx. quinquefasciatus* quatro metalo peptidases (B0W1Y3, B0W1Y6, B0WH17, B0W1Y2), duas catepsinas (B0WI10, B0W0V3) e uma Lipase (B0WDK6). Dentre as exopeptidases, foi identificado por espectrometria de massas cinco aminopeptidases (B0WS64, B0XAQ5, B0W6N8, B0WH45, B0X730) e cinco carboxipeptidases (B0WS12, B0W4H6, B0XCT6, B0WTW4, B0WNR6), além de duas dipeptidases (B0WDL4, B0X281). Apesar da identificação não garantir que estas enzimas estejam ativas, a presença de exopeptidases, conhecidas por atuarem no final da digestão, corrobora a hipótese, discutida anteriormente, de que existe uma pequena demanda digestiva e/ou o inseto está preparando o tecido para realizar uma eficiente e rápida digestão. Em *Cx. quinquefasciatus*, aminopeptidases também apresentaram o pico de atividade em 36 h após o repasto sanguíneo, como aconteceu para tripsinas e quimiotripsinas (Okuda et al., 2002). Entretanto, as identidades dessas enzimas eram desconhecidas até então, passando esse estudo, a primeira descrição destas enzimas no intestino de fêmeas alimentadas com açúcar, o que, de forma indireta, nos permite explorar quais genes estão sendo exigidos em determinada condição fisiológica.

No intestino de fêmeas de *Cx. quinquefasciatus* alimentadas com açúcar foram identificadas, por espectrometria, uma variedade de enzimas que hidrolisam carboidratos. Essas enzimas possuem massa molecular aproximada de 70 kDa e, dentre elas, foram identificadas quatro alfa-glicosidase (B0XJR7, B0XAA1, B0XBN3, B0WS02), glicosidase 2 subunidade beta (B0WS45), oligo-1,6-glicosidase (B0WFP7), alfa-amilase (B0XGH0, B0XJR6) e maltase 1 (B0X220). Identificamos também uma proteína que participa do processamento das glicosidases inativas, a neutral alfa-glicosidase AB (B0WQR9). Apesar de não termos realizado ensaios de atividade para caracterização destas enzimas, a análise de bioinformáticas das sequencias primárias de aminoácidos das quatro alfa-glicosidases identificadas aqui, mostrou a presença de assinaturas típicas que possibilitam uma atividade destas enzimas. Os mosquitos adultos, machos e fêmeas alimentam-se constantemente de soluções açucaradas, o que constitui parte essencial da dieta destes insetos (Grimstad e DeFoliart, 1974; Van Handel, 1984; Foster, 1995). A ingestão de açúcares serve não apenas como

combustível inicial para a produção de ovos (Magnarelli, 1978), mas também supre a energia necessária para o voo (Van Handel, 1984), permitindo a busca por uma alimentação sanguínea e por reservatórios para a oviposição (Jamnback, 1961; Hunter, 1977). Além disso, estes insetos podem armazenar o excesso de carboidratos na forma de glicogênio, e, uma vez atingido o limiar de seu armazenamento, os açúcares ingeridos são utilizados para a síntese de lipídeos (Van Handel, 1984). As α -glicosidases (EC 3.2.1.20, α -glicosídeo glicohidrolase) compõem um grupo de exo-glicosídeo hidrolases que catalisam a hidrólise de resíduos α -glicosil das extremidades não reduzidas de substratos α -ligados, liberando glicose (Chiba, 1997). Sacarases, termo geralmente utilizado como sinônimo de α -glicosidases, foram primeiramente identificadas em mosquitos, e foi demonstrado que as glândulas salivares de *Culex tarsalis* contêm uma enzima capaz de clivar sacarose (Schaefer e Miura 1972). Em seguida àquela observação, foi detectada uma α -glicosidase secretada pelas glândulas salivares de *Ae. aegypti*, a qual foi caracterizada mais tarde como sendo uma proteína monomérica secretada pelos lóbulos laterais das glândulas salivares em resposta à ingestão de açúcares ou sangue (Marinotti e James, 1990). Nestes trabalhos também foi sugerido que esta enzima poderia ser ingerida junto com a alimentação e ajudaria na digestão dos carboidratos no divertículo. Atividades α -glicolíticas também foram detectadas em glândulas salivares dos mosquitos *Aedes albopictus* (Marinotti et al., 1996) e *Anopheles darlingi* (Moreira-Ferro et al., 1999), e em intestinos de *Anopheles stephensi*, antes e após uma refeição sanguínea (Billingsley e Hecker, 1991). O receptor da toxina binária de *Bacillus sphaericus* localizado na membrana intestinal de larvas de *Culex pipiens* foi identificado como uma α -glicosidase de 60 kDa (Darboux et al., 2001). Além disso, três genes codificando proteínas do tipo maltase (maltase-like proteins) foram seqüenciadas em *Anopheles gambiae* (Holt et al., 2002). Dois deles correspondendo aos genes intestinais Agm1 e Agm2, seqüenciados e caracterizados por Zheng et al. (1995), e o terceiro, um fragmento (ENSANGG00000012889), que por algum motivo não foi identificado e caracterizado por estes autores. Atividades α -glicolíticas foram detectadas nos intestinos médios dos flebotomíneos *Phlebotomus langeroni* (Dillon e El Kordy, 1997), *Lutzomyia longipalpis* (Gontijo et al., 1998) e *Phlebotomus papatasi* (Jacobson e Schlein, 2001), sendo correlacionadas com a digestão de sacarose nestes insetos. Entretanto, em mosquitos, a importância dos açúcares na dieta, e, até mesmo a origem das α -

glicosidases e o local em que ocorre a digestão dos carboidratos ingeridos, ainda permanece uma discussão aberta. Em todos estes relatos anteriores (James et al., 1989; Marinotti e James, 1990; Marinotti et al., 1996; Moreira-Ferro et al., 1999), a produção das α -glicosidases estava sempre associada às glândulas salivares e à digestão de sacarose no divertículo daqueles mosquitos. Sendo assim, a importância do açúcar na dieta do mosquito bem como o papel das glândulas salivares, divertículo e intestino no metabolismo deste nutriente precisam ser reavaliados. Há vertentes que apoiam a teoria de que múltiplas alimentações sanguíneas possam substituir o papel de uma dieta contendo açúcares (Gary Jr e Foster, 2001). Entretanto, tem sido observado que uma refeição rica em carboidratos é essencial para manutenção das principais funções fisiológicas dos mosquitos (Grimstad e DeFoliart, 1974; Van Handel, 1984). A dieta de carboidratos prolonga de forma notável a longevidade dos mosquitos (Nayar e Sauerma, 1971). Foi demonstrado que fêmeas que possuíam soluções açucaradas em sua dieta, apresentavam maior longevidade, quando comparadas àquelas que se alimentavam exclusivamente de sangue, as quais morriam 6 dias após emergirem das pupas (Souza-Neto et. al., 2007). Neste mesmo trabalho, foi detectada a atividade de três isoformas de alfa-glicosidases (alphaGlu1, alphaGlu2, alphaGlu3), com aproximadamente 70 kDa, mas com consideráveis diferenças nos valores de Km. Além disso, observou-se que as alfa-glucosidases eram sintetizadas no mesmo local onde as enzimas eram digeridas e absorvidas, ou seja, no intestino médio (Souza-Neto et. al., 2007). Assim, as identificações destas enzimas em extratos do intestino de *Cx. quinquefasciatus* alimentados com açúcar não excluem a hipótese dessas enzimas serem produzidas nas glândulas salivares ou divertículo, e terem sido ingeridas junto com o bolo alimentar. Porém, a identificação de diferentes carboidrases mostra que esses insetos conservam diferentes cópias de genes que codificam para estas enzimas e corrobora a hipótese sustentada por outros autores (Souza-Neto et. al. 2007) sobre a importância do açúcar na dieta de fêmeas. Esses achados apontam as glicosidases como alvos potenciais de drogas que possam inibir a atividade dessas enzimas e levar ao controle dos vetores.

A caracterização molecular da resistência a inseticidas começou a ser uma área explorada desde o sequenciamento genômico da primeira espécie de artópode vetor, o *An. gambiae*, em 2002. Com a publicação do genoma do *Ae. aegypti* e mais recentemente do *Cx. quinquefasciatus*, surgiram novas oportunidades, através de

análises genômicas comparativas, para caracterizar os genes envolvidos na resistência a inseticidas. A resistência a inseticidas é um processo que pode ocorrer por diferentes mecanismos, dentre eles, destaca-se a síntese de enzimas que degradam os compostos xenobióticos tóxicos dos inseticidas. Algumas famílias de genes detoxificantes têm sido reportadas em *Cx. quinquefasciatus*, por exemplo: citocromo P450 (166 genes), glutaciona-S-transferases (40 genes) e esterases (62 genes) (Arensburger et al., 2010). Análises comparativas mostraram que o número de genes de *Cx. quinquefasciatus* é consideravelmente alto em relação a outros mosquitos (Arensburger et al., 2010). No presente trabalho, foram identificadas, no intestino de *Cx. quinquefasciatus*, várias proteínas relacionadas à resistência à inseticidas, como: citocromo P450 (B0XJT6, B0WRU2, B0VZI7, B0XCA1, B0WTS8, B0WTS9, B0W673, B0X581, B0W6Y, B0WTQ1, B0WTP9, B0XJY3, B0WTS7, B0WFB9, B0WFC1, B0WTS5, B0WTS3); enzimas do tipo carboxilesterase-6 (B0WIH9); palmitoil tioesterase 1 (B0W1X8); esterase A11 (A5Y5K2); ubiquitina tioesterase OTUB1 (B0W5S7); esterase A4-B4 (R9RIG8) e esterase B1 (EST1); glutaciona-s-transferase theta (B0XGK3, B0W6C9, B0XGJ5, B0X3C8); glutaciona s-transferase (B0WFX0, B0VZ90, B0W6C9); glutaciona s-transferase 1 (B0XAJ0, B0W6C3); microsomal glutaciona S-transferase 1 (B0XG85, B0XG85); glutaciona transferase (C4B4V7, C4B4V8, B0WFX0); glutaciona S-transferase 1-6 (B0W6B0); e uma proteína envolvida na síntese de glutaciona, a glutaciona sintetase (B0WSC1). A expressão de uma grande quantidade destas enzimas no intestino de *Cx. quinquefasciatus* indica que outros tecidos, além dos tecidos mais externos do inseto, podem produzir essas enzimas.

Com relação ao *An. albitalis*, verificamos que o número de proteínas identificadas por MS/MS no intestino foi menor que o de *Cx. quinquefasciatus*, provavelmente, por: (i) como não possui o genoma sequenciado, a identificação das proteínas por PSM é com base em sequências homologas de espécies próximas evolutivamente. Assim, muitas proteínas que apresentam algumas diferenças nas sequências peptídicas, possuindo um espectro de massas diferente daqueles depositados, não são reconhecidas pelo programa de identificação; (ii) o número de genes codificantes para proteínas é menor neste inseto. Por exemplo, o número de genes codificantes para proteínas é 52% menor no *An. gambiae*, quando comparados ao *Cx. quinquefasciatus* (Arensburger et al., 2010). Para o estudo de organismos de

genoma desconhecido, uma outra abordagem utilizada, que não o PSM, é o sequenciamento “*de novo*”, a qual pode ser realizada, por exemplo, pelo PepExplorer. Neste caso, não é necessário um banco de dados para a identificação das proteínas por homologia. Assim os estudos das proteínas totais de intestino de *An. albicans* é dificultada pela ausência do genoma sequenciado e necessitará de abordagens complementares, além do já obtido neste trabalho.

A condição de alimentação com açúcar é uma condição controle para diversos experimentos, visto que nela ocorre níveis de expressão gênica basal. Assim o mapeamento do perfil proteico nessas condições pode servir como comparação para vários experimentos de alteração da expressão genica durante diferentes estímulos. Os avanços nas técnicas de biologia molecular têm permitido a identificação de alguns genes que codificam para tripsinas. Entretanto, a confirmação da codificação desses genes em enzimas ativas, bem como a diferenciação dos papéis desempenhadas por estas enzimas nos complexos processos fisiológicos que envolvem o ciclo de vida de um mosquito, tem sido um desafio. O tamanho reduzido dos mosquitos e a baixa abundancia de algumas isoformas dessas enzimas dificultam os trabalhos de purificação e ensaios enzimáticos clássicos com enzimas purificadas. Em contraste à abundante informação disponível sobre as sequencias de cDNAs que codificam para tripsinas, poucos estudos trataram da purificação e sequenciamento dessas enzimas (Graf et al., 1985; Barillas-Mury e Wells, 1993; Kalhok et al., 1993; Noriega et al., 1996 b, Caroci et al 2003; Isoe et al., 2009; Ráscon et al., 2011). Além dos problemas intrínsecos à técnica de clonagem e expressão, o sequenciamento genômico de diferentes espécies de culicídeos vetores tem revelado que esta família de insetos possuem um grande repertório de genes que codificam para serina peptidases do tipo tripsinas, o que também contribui para o entrave das pesquisas de clonagem e expressão, visto que seria necessário a expressão de centenas de genes.

O sequenciamento genômico de diversas espécies de mosquitos tem oferecido uma oportunidade para explorar novas estratégias de controle dos vetores. Milhares de genes podem ser investigados como potenciais alvos pra intervenções. Entretanto, o completo sequenciamento genômico das espécies não se dá por terminado até que a maioria dos “scaffolds” sequenciados sejam agrupados e organizados em cromossomos. Sequencias fragmentadas e não mapeadas criam sérios problemas para as análises genômicas, pois saltos não identificados ou sequenciamentos mal

interpretados, provocam anotações incorretas ou incompletas das sequências genéticas. Além disso, a mera deposição automática das sequências, sem a validação manual, provoca um acúmulo de genes mal sequenciados e fragmentados. Assim, experimentos de proteoma ajudam na anotação funcional do genoma.

Apesar de não termos finalizado os experimentos comparativos entre as condições de alimentadas com açúcar e sangue, nosso estudo traz informações importantes a respeito da fisiologia do intestino das três espécies estudadas e nos atenta para o vasto repertório de tripsinas sintetizados pelo intestino de fêmeas alimentadas com açúcar, observação bastante surpreendente, pois os estudos sempre apontavam para a expressão destas enzimas em intestino de fêmeas PAS. Este estudo também traz o primeiro mapeamento e descrição das proteínas solúveis presentes no intestino de *Cx. quinquefasciatus* e *An. albitalarsis* alimentadas com açúcar. Estudos de sub-proteoma, como realizados aqui, podem ser importantes para apontar quais proteínas são de fato expressadas e sintetizadas no intestino desses mosquitos, o que, juntamente com a exploração das ferramentas de bioinformática, ajuda a inferir as funções do repertório proteico desse órgão.

6. CONCLUSÕES

- Pelo menos oito e dez diferentes SPtrip são sintetizadas constitutivamente pelo intestino de *Cx. quinquefasciatus* e *An. albitalarsis*, respectivamente, as quais foram identificadas por espectrometria de massas;
- O intestino de fêmeas de mosquitos alimentados com açúcar possui um complexo repertório de SPtrip;
- As SPtrip identificadas por MS/MS em *Cx. quinquefasciatus* possuem características típicas de tripsinas digestivas de invertebrados e apesar do sítio ativo conservado, possuem diferenças nas sequencias de aminoácidos em regiões como a de especificidade ao substrato, ativação ao zimogênio;
- Tripsinas de *Cx. quinquefasciatus* possuem peptídeos proteotípicos que podem possibilitar a identificação de genes expressos sob condições distintas em amostras complexas.

7. REFERÊNCIAS BIBLIOGRÁFICAS

1. Alves LC, Almeida PC, Franzoni L, Juliano L, Juliano MA. 1996. Synthesis of N α -protected aminoacyl 7-amino-4-methyl-coumarin amide by phosphorous oxychloride and preparation of specific fluorogenic substrates for papain. *Pept Res.* 9:92-96.
2. Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B, et al. 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science.* 330(6000):86-88.
3. Ahid SMM, Vasconcelos PSS, Lourenço-de-Oliveira R. 2000. Vector competence of *Culex quinquefasciatus* Say from different regions of Brazil to *Dirofilaria immitis*. *Mem Inst Oswaldo Cruz.* 95: 769–775.
4. Amenya DA, Chou W, Li J, Yan G, Gershon PD, James AA, Marinotti O. 2010. Proteomics reveals novel components of the *Anopheles gambiae* eggshell. *J Insect Physiol.* 56(10): 1414-1419.
5. Arnon R, Neurath H. 1969. An immunological approach to the study of evolution of trypsins. *Proc Natl Acad Sci USA.* 64(4):1323–1328.
6. Bao YY, Qin X, Yu B, Chen LB, Wang ZC, Zhang CX. 2014. Genomic insights into the serine protease gene family and expression profile analysis in the planthopper, *Nilaparvata lugens*. *BMC Genomics.* 15:507.
7. Barrillas-Mury C, Graf R, Hagedorn HH, Wells MA. 1991. cDNA and deduced amino acid sequence of the blood meal-induced trypsin from the mosquito, *Aedes aegypti*. *Insect Biochem.* 21, 825-831.
8. Barrillas-Mury C, Wells MA. 1993. Cloning and sequencing of the blood meal-induced late trypsin gene from the mosquito, *Aedes aegypti* and the characterization of the upstream regulatory region. *Insect Mol Biol.* 2(1): 7-12.
9. Barrillas-Mury C, Noriega FG, Wells MA. 1995. Early trypsin activity is part of the signal transduction system that activates transcription of the late trypsin gene in the midgut of the mosquito, *Aedes aegypti*. *Insect Biochem Mol Biol.* 25(2), 241-246.
10. Barrett AJ, Rawlings ND, Woessner JF. (eds). 2004. Handbook of Proteolytic Enzymes, 2nd edition, Academic Press, London.

11. Bartholomay LC, Fuchs JF, Cheng LL, Beck ET, Vizioli J, Lowenberger C, Christensen BM. Reassessing the role of defensin in the innate immune response of the mosquito, *Aedes aegypti*. 2004. *Insect Mol. Biol.* 13: 125-132.
12. Bartholomay LC, Waterhouse RM, Mayhew GF, Campbell CL, Michel K, Zou Z, et al. 2010. Pathogenomics of *Culex quinquefasciatus* and meta-analysis of infection responses to diverse pathogens. *Science*. 330(6000):88-90
13. Baptista AM, Jonson PH, Hough E, Petersen SB. 1998. The origin of trypsin: evidence for multiple gene duplications in trypsins. *J Mol Evol.* 47(3): 353-362.
14. Beerntsen BT, James AA, Christensen BM. 2000. Genetics of mosquito vector competence. *Microbiol Mol Biol. Rev.*64:115-137.
15. Berner R, Rudin W, Hecker H. 1983. Peritrophic membranes and protease activity in the midgut of the malaria mosquito, *Anopheles stephensi* (Liston) (Insecta: Diptera) under normal and experimental conditions. *J Ultrastruct Res.* 83(2): 195 204.
16. Beyenbach KW, Baumgart S, Lau K, Piermarini PM, Zhang S: 2009. Signaling to the apical membrane and to the paracellular pathway: changes in the cytosolic proteome of *Aedes* Malpighian tubules. *J Exp Biol.* 212(Pt 3): 329–340.
17. Billingsley PF. 1990. Blood digestion in the mosquito, *Anopheles stephensi* Liston (Diptera: Culicidae): partial characterization and post-feeding activity of midgut aminopeptidases. *Arch Insect Biochem Physiology* 15(3): 149–163.
18. Billingsley PF, Hecker H. 1991. Blood digestion in the mosquito, *Anopheles stephensi* Liston (Diptera:Culicidae): activity and distribution of trypsin, aminopeptidases and α glycosidase in the midgut. *J. Med. Entomol.* 28(6), 865-871.
19. Billingsley PF, Rudin W. 1992. The role of the mosquito peritrophic membrane in bloodmeal digestion and infectivity of *plasmodium* species. *J Parasitol.* 78(3):430-440.
20. Billingsley PF, Lehane MJ. 1996. Structure and ultrastructure of the insect midgut. In: Lehane M J, Billingsley P F (eds). *Biology of the insect midgut*. 1st edition. London: Chapman & Hall, 86-114.

21. Biron DG, Agnew P, Marche L, Renault L, Sidobre C, Michalakis Y. 2005. Proteome of *Aedes aegypti* larvae in response to infection by the intracellular parasite *Vavraia culicis*. *Int J Parasitol.* 35(13): 1385-1397.
22. Biessmann H, Nguyen QK, Le D, Walter MF. 2005. Microarray-based survey of a subset of putative olfactory genes in the mosquito *Anopheles gambiae*. *Insect Mol Biol.* 14:575–589.
23. Bode W, Meyer E Jr, Powers JC. 1989. Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism based inhibitors: review. *Biochemistry.* 28(5): 1951-1963.
24. Borges-Veloso A, Saboia-Vahia L, Cuervo P, Pires RC, Britto C, Fernandes N, d'Avila-Levy CM, De Jesus JB. 2012. Proteolytic profiling and comparative analyses of active trypsin-like serine peptidases in preimaginal stages of *Culex quinquefasciatus*. *Parasit Vectors.* 20;5:123.
25. Borovsky D. 1986. Proteolytic enzymes and blood digestion in the mosquito, *Culex nigripalpus*. *Arch Insect Biochem Physiol.* 3(2):147-160.
26. Borovsky D. 2003. Biosynthesis and control of mosquito gut proteases. *IUBMB Life.* 55(8):435–441.
27. Botos I, Meyer E, Nguyen M, Swanson SM, Koomen JM, Russell DH, et al. 2000. The structure of an insect chymotrypsin. *J Mol Biol.* 298(5):895-901.
28. Brackney DE, Foy BD, Olson KE. 2008. The effects of midgut serine proteases on dengue virus type 2 infectivity of *Aedes aegypti*. *Am J Trop Med Hyg.* 79(2):267-274
29. Brackney DE, Isoe J, Zamora J, Foy BD, Miesfelde RL, Oslon KE: 2010. Expression profiling and comparative analyses of seven midgut serine proteases from the yellow fever mosquito, *Aedes aegypti*. *J Insect Physiol.* 56(7):736-744.
30. Briegel H, Lea AO. 1975. Relationship between protein and proteolytic activity in the midgut of mosquitoes. *J Insect Physiol.* 21(9):1597-1604.
31. Bruce C, Stone K, Gulcicek E, Williams K. 2013. Proteomics and the analyses of proteomic data: 2013 overview of current protein-profiling technologies. *Curr Protoc Bioinformatics.* Chapter 13: Unit 13.21

32. Calvo E, Dao A, Pham VM, Ribeiro JM. 2007. An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. *Insect Biochem Mol Biol.* 37:164–175.
33. Caroci AS, Calvo E, Ribolla PE, De Biachi AG, Marinotti O. 2003. Two digestive trypsins occur in three species of neotropical anophelines. *J. Med. Entomol.* 40(6): 991-995.
34. Chait B. 2006. Mass spectrometry: bottom-up or topdown? *Science.* 314(5796):65-66.
35. Chege GMM, Pumpuni CB, Beier JC. 1996. Proteolytic enzyme activity and *Plasmodium falciparum* sporogonic development in three species of *Anopheles* mosquitoes. *J Parasitol.* 82(1): 11-6.
36. Chen JM, Kukor Z, Le Marechal U, Toth M, Tsakiris L, Raguene O, et al. 2003. Evolution of trypsinogen activation peptides. *Mol Biol Evol.* 20(11): 1767-1777.
37. Chiba S. 1997. Molecular mechanism in α -glucosidase and glucoamylase. *Biosci Biotechnol Biochem.* 61(8):1233-1239.
38. Choumet V., Carmi-Leroy A, Laurent C, Lenormand P, Rousselle JC, Namane A, Roth C, Brey PT. 2007. The salivary glands and saliva of *Anopheles gambiae* as an essential step in the *Plasmodium* life cycle: a global proteomic study. *Proteomics.* 7(18):3384-3394.
39. Christophides GK, Zdobnov E, Barillas C, Blandin S, Blass C, Brey PT, et al. 2002. Immunity related genes and gene families in *Anopheles gambiae*: A comparative genomic analysis. *Science.* 298: 159-165.
40. Clements AN. 2000. The biology of mosquitoes. Volume 1, Development, nutrition and reproduction. Wallingford, UK: CABI Publishing.
41. Corena M del P, VanEkeris L, Salazar MI, Bowers D, Fiedler MM, Silverman D, et al. 2005. Carbonic anhydrase in the adult mosquito midgut. *J Exp Biol.* 208(Pt 17):3263-73.
42. Cox J, Mann M. 2011. Quantitative, high resolution proteomics for data driven. *Annu Rev Biochem.* 80:273-299.
43. Craik CS, Choo QL, Swift GH, Quinto C, MacDonald RJ, Rutter WJ. 1984. Structure of two related rat pancreatic trypsin genes. *J Biol Chem.* 259(22):14255–14264.

44. Craik CS; Largman C, Fletcher T; Roczniak S, Barr PJ, Fletterick R, Rutter WJ. 1985. Redesigning trypsin: alteration of substrate specificity. *Science*. 228(4697):291-297.
45. Cuervo P, Mesquita-Rodrigues C, D'avila Levy CM, Britto C, Pires FA, Gredilha R, Alves CR, Jesus JB. 2008. Serine protease activities in *Oxysarcodexia thornax* (Walker) (Diptera: Sarcophagidae) first instar larva. *Mem Inst Oswaldo Cruz*. 103(5): 504-506.
46. Cuervo P, De Jesus JB, Saboia-Vahia L, Mendonça-Lima L, Domont GB, Cupolillo E. 2009. Proteomic characterization of the released/secreted proteins of *Leishmania (Viannia) braziliensis* promastigotes. *J Proteomics*. 73(1):79-92.
47. Dana AN, Hong YS, Kern MK, Hillenmeyer ME, Harker BW, Lobo NF, et al. 2005. Gene expression patterns associated with blood-feeding in the malaria mosquito *Anopheles gambiae*. *BMC Genomics*. 6:5.
48. Darboux I, Nielsen-LeRoux C, Jean-François C, Pauron D. 2001. The receptor of *Bacillus sphaericus* binary toxin in *Culex pipiens* (Diptera: Culicidae) midgut: molecular cloning and expression. *Insect Biochem Mol Biol*. 31(3):981-990.
49. David JP, Strode C, Vontas J, Nikou D, Vaughan A, Pignatelli PM, et al. 2005. The *Anopheles gambiae* detoxification chip: a highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. *Proc Natl Acad Sci USA*. 102:4080–4084.
50. de Almeida RW, Tovar FJ, Ferreira II, Leoncini O. 2003. Chymotrypsin genes in the malaria mosquitoes *Anopheles aquasalis* and *Anopheles darlingi*. *Insect Biochem Mol Biol*. 33(3):307-15.
51. De Jesus JB; Cuervo P; Britto C; Sabóia-Vahia L; Silva-Filho FC; Borges-Veloso A; Petrópolis DB; Cupolillo E; Domont GB. 2009. Cysteine peptidase expression in *Trichomonas vaginalis* isolates displaying high and low virulence phenotypes. *J Proteome Res*. 8(3): 1555-1564.
52. Devenport M, Alvarenga PH, Shao L, Fujioka H, Bianconi ML, Oliveira PL, Jacobs-Lorena M. 2006. Identification of the *Aedes aegypti* peritrophic matrix protein AeIMUC1 as a heme-binding protein. *Biochemistry*. 45(31):9540-9549.
53. Dias-Lopes G, Borges-Veloso A, Saboia-Vahia L, Domont BG, Britto C, Cuervo P, et al. 2015. Expression of active trypsin-like serine peptidases in the midgut of sugar-feeding female *Anopheles aquasalis*. *Parasit Vectors*. 8:296

54. Dillon RJ, El-Kordy. 1997. Carbohydrate digestion in sandflies: α -glucosidase activity in the midgut of *Phlebotomus langeroni*. *Comp Biochem Physiol B Biochem Mol Biol.* 116(1):35-40.
55. Dinardo-Miranda LL, Contel EPB. 1996. Variability in natural populations of *Aedes aegypti* (Diptera, Culicidae) from Brazil. *J Med Entomol.* 33(5):726-733.
56. Dinglasan RR, Kalume DE, Kanzok SM, Ghosh AK, Muratova O, Pandey A, et al. 2007. Disruption of *Plasmodium falciparum* development by antibodies against a conserved mosquito midgut antigen. *Proc Natl Acad Sci USA.* 104(33):13461-13466.
57. Dinglasan RR, Devenport M, Florens L, Johnson Jr, McHugh CA, Donnelly-Doman M, Carucci DJ, Yates Jr, Jacobs-Lorena M. 2009. The *Anopheles gambiae* adult midgut peritrophic matrix proteome. *Insect Biochem Mol Biol.* 39(2):125-134.
58. Duncan AB, Agnew P, Noel V, Demettre E, Seveno M, Brizard JP, et al., 2012. Proteome of *Aedes aegypti* in response to infection and coinfection with microsporidian parasites. *Ecol Evol.* 2(4):681–694.
59. Eng JK, McCormack AL, Yates JR. 1994. An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database. *J Am SW Mass Spectrom.* 5:976-989.
60. Farajollahi A, Fonseca DM, Kramer LD, Kilpatrick AM: 2011. Bird biting mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infect Genet Evol.* 11(7):1577-1585.
61. Faran ME, Linthicum KJ. 1981. A handbook of the Amazonian species of *Anopheles* (*Nyssorhynchus*) (Diptera: Culicidae). *Mosquito Sys* 13: 1-81.
62. Fazito do Vale V, Pereira MH, Gontijo NF. 2007. Midgut pH profile and digestion in the larvae of *Lutzomyia longipalpis* (Diptera: Psychodidae). *J Insect Physiol.* 53(11):1151-1159.
63. Feldmann AM, Billingsley PF, Savelkoul E. 1990. Bloodmeal digestion by strains of *Anopheles stephensi* Liston (Diptera: Culicidae) of differing susceptibility to *Plasmodium falciparum*. *Parasitology.* 101(Pt 2): 193-200.
64. Felix CR, Betschart B, Billingsley PF, Freyvogel TA. 1991. Post-feeding induction of trypsin in the midgut of *Aedes aegypti* L. (Diptera: Culicidae) is separable in two cellular phases. *Insect Biochem.* 21(2): 197-203.

65. Ferec C, Ragueneas O, Salomon R, Roche C, Bernard JP, Guillot M, et al. 1999. Mutations in the cationic trypsinogen gene and evidence for genetic heterogeneity in hereditary pancreatitis. *J Med Genet.* 36(3): 228- 232.
66. Fields S. 2001. Proteomics is genomeland. *Science.* 291:1221-1224.
67. Fisk FW. 1950. Studies on proteolytic digestion in adult *Aedes aegypti* mosquitoes. *Ann. Entomol. Soc. Am.* 439(4):555-572.
68. Forattini OP, Kakitani I, Massad E, Marucci D. 1994. Studies on mosquitoes (Diptera: Culicidae) and anthropic environment. 5. Breeding of *Anopheles albitarsis* in flooded rice fields in south-eastern Brazil. *Rev Saude Publica.* 28(5):329-331.
69. Forattini OP. 2002. *Culicidologia Médica.* São Paulo, EDUSP. 860p.
70. Foster AW. 1995. Mosquito sugar feeding and reproductive energetics. *Annu Rev Entomol.* 40:433-474.
71. Gary RE Jr, Foster WA. 2001. Effects of available sugar on the reproductive fitness and vectorial capacity of the malaria vector *Anopheles gambiae*. *J. Med. Entomol.* 38(1):22-28.
72. Galán JE, Pace J, Hayman MJ. 1992. Involvement of the epidermal growth factor receptor in the mammalian cells by *Salmonella typhimurium*. *Nature.* 357(6379): 588-589.
73. Gallien S, Kim SY, Domon B. 2015. Large-Scale Targeted Proteomics Using Internal Standard Triggered-Parallel Reaction Monitoring. *Mol Cell Proteomics.*14(6):1630-1644.
74. Gan CS, Chong PK, Pham TK, Wright PC. 2007. Technical, Experimental, and Biological Variations in Isobaric Tags for Relative and Absolute Quantitation (iTRAQ). *J Proteome Res.* 6(2):821–827.
75. Gass RF, Yeates RA. 1979. In vitro damage of cultured ookinetes of *Plasmodium gallinaceum* by digestive proteinases from susceptible *Aedes aegypti*. *Acta Trop.* 36(3):243–252.
76. Giannoni F, Müller HM, Vizioli, J, Catteruccia F, Kafatos FC, Crisanti A. 2001. Nuclear factors bind to a conserved DNA element that modulates transcription of *Anopheles gambiae* trypsin genes. *J Biol Chem.* 276(1):700–707.

77. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R. 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol.* 17(10):994-9.
78. Goddard LB, Roth AE, Reisen WK, Scott TW. 2002. Vector competence of California mosquitoes for West Nile virus. *Emerg Infect Dis.* 8(12):1385–1391.
79. Gontijo NF, Almeida-Silva S, Costa FF, Mares-Guia ML, Williams P, Melo MN. 1998. *Lutzomyia longipalpis*: pH in the gut, digestive glycosidases, and some speculations upon *Leishmania* development. *Exp. Parasitol.* 90(3):212-219.
80. Gooding RH. 1966. Physiological aspects of digestion of the blood meal by *Aedes aegypti* (Linnaeus) and *Culex fatigans* (Wiedemann). *J. Med. Ent.* 3(1):53-60.
81. Gooding RH. 1966. In vitro properties of proteinases in the midgut of adult *Aedes aegypti* L. and *Culex fatigans* (Wiedemann). *Comp. Biochem. Physiol.* 17(1): 115-127.
82. Gorman MJ, Paskewitz SM. 2001. Serine proteases as mediators of mosquito immune responses. *Insect Biochem Mol Biol.* 31(3): 257–262.
83. Graça-Souza AV, Maya-Monteiro C, Paiva-Silva GO, Braz GRC, Paes MC, Sorgine MHF, et al. 2006. Adaptations against heme toxicity in blood-feeding arthropods. *Insect Biochem Mol Biol.* 36(4):322-335.
84. Graf R, Briegel H. 1985. Isolation of trypsin isozymes from the mosquito, *Aedes aegypti* (L.). *Insect Biochem.* 15(5): 611-618.
85. Graf R, Binz H, Briegel H. 1988. Monoclonal antibodies as probes for *Aedes aegypti* trypsin. *Insect Biochem.* 18(5):463-470.
86. Graf R, Briegel H. 1989. The synthetic pathway of trypsin in the mosquito *Aedes aegypti* L. (Diptera: Culicidae) and in vitro stimulation in isolated midguts. *Insect Biochem.* 19(2):129-137.
87. Graves PR, Haystead TAJ. 2002. Molecular biologist's guide to proteomics. *Microbiol Mol Biol Rev.* 66(1): 39-63.
88. Grimstad PR, DeFoliart GR. 1974. Nectar sources of Wisconsin mosquitoes. *J Med Entomol.* 11(3):331-341.
89. Harbach RE, Besansky NJ. 2014. Mosquitoes. *Curr Biol.* 24(1):R14-5.

90. Hebert AS, Richards AL, Bailey DJ, Ulbrich A, Coughlin EE, Westphall MS, Coon JJ. 2014. The one hour yeast proteome. *Mol Cell Proteomics*. 13(1):339-347.
91. Hedstrom L; Szilagyl L, Rutter WJ. 1992. Converting trypsin to chymotrypsin the role of surface loops. *Science*. 255(5049):1249-1253.
92. Hedstrom L. 2002. Serine protease mechanism and specificity. *Chem Rev*. 102(12): 4501-4523.
93. Heger A, Ponting CP. 2008. Evolutionary rate analyses of orthologs and paralog from 12 *Drosophila* genomes. *Genome Res*. 17(12):1837–1849
94. Hecker H. 1977. Structure and function of midgut epithelial cells in culicidae mosquitoes (Insecta, diptera). *Cell Tissue Res*. 184(3): 321–341.
95. Hillyer JF, Schmidt SL, Christensen BM. 2003. Hemocyte-mediated phagocytosis and melanization in the mosquito *Armigeres subalbatus* following immune challenge by bacteria. *Cell Tissue Res*. 313(1):117-127
96. Holder P, Browne G, Bullians M. 1999. The mosquitoes of New Zealand and their animal disease significance. *Surveillance*. 26(4):12-15.
97. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, et al. 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*. 298(5591):129–149.
98. Horler E, Briegel H. 1995. Proteolytic enzymes of female *Anopheles*: biphasic synthesis, regulation and multiple feeding. *Arch. Insect Biochem. Physiol*. 28(2): 189-205.
99. Hu Q, Noll RJ, Li H, Makarov A, Hardman M, Graham Cooks R. 2005. The orbitrap: a new mass spectrometer. *J Mass Spectrom*. 40(4):430–443
100. Huber R, Kukla D, Bode W, Schwager P, Bartels K, Deisenhofer J, et al. 1974. Structure of the complex formed by bovine trypsin and bovine pancreatic trypsin inhibitor II. Crystallographic refinement at 1.9 Å resolution. *J Mol Biol*. 89(1):73–101
101. Huber M, Cabib E, Miller LH. 1991. Malaria parasite chitinase and penetration of the mosquito peritrophic membrane. *Proc Natl Acad Sci USA*. 88(7): 2807–2810.
102. Hunter DM. 1977. Sugar-feeding in some Queensland black flies. *J. Med. Entomol*. 14(2):229-232.

103. Isoe J, Rascón AA, Kunz S, Miesfeld RL. 2009. Molecular genetic analysis of midgut serine proteases in *Aedes aegypti* mosquitoes. *Insect Biochem Mol Biol.* 39(12), 903-912.
104. Jacobson RL, Schlein Y. 2001. *Phlebotomus papatasi* and *Leishmania major* parasites express α -amylase and α -glucosidase. *Acta Trop.* 78(1):41-49.
105. Jahan N, Docherty PT, Billingsley PF, Hurd H. 1999. Blood digestion in the mosquito, *Anopheles stephensi*: the effects of *Plasmodium yoelii nigeriensis* on midgut enzyme activities. *Parasitology.* 119 (Pt 6):535-541.
106. James AA, Blackmer K, Racioppi JV. 1989. A salivary gland-specific, maltase-like gene of the vector mosquito, *Aedes aegypti*. *Gene.* 75(1):73-83.
107. Jamnback H. 1961. Observations on *Culicoides obsoletus* (Meigen) in the laboratory (Diptera: Ceratopogonidae). *Mosq. News.* 21:48-53.
108. Jiang Q, Hall M, Noriega FG, Wells M. 1997. cDNA cloning and pattern of expression of an adult, female specific chymotrypsin from *Aedes aegypti* midgut. *Insect Biochem Mol Biol.* 27(4):283–289.
109. Junqueira M, Carvalho PC. 2012. Tools and challenges for diversity-driven proteomics in Brazil. *Proteomics.* 12(17):2601–2606.
110. Kalhok SE, Tabak LM, Prosser DE, Brook W, Downe AE, White BN. 1993. Isolation, sequencing and characterization of two cDNA clones coding for trypsin-like enzymes from the midgut of *Aedes aegypti*. *Insect Mol. Biol.* 2(2): 71-79.
111. Kalume DE, Okulate M, Zhong J, Reddy R, Suresh S, Deshpande N, et al. 2005. A proteomic analysis of salivary glands of female *Anopheles gambiae* mosquito. *Proteomics.* 5(14):3765-3777.
112. Kilpatrick AM, Randolph SE: 2012. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *Lancet.* 380(9857):1946-1955.
113. Kitamoto Y, Yuan X, Wu Q, McCourt DW, Sadler JE. 1994 Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of a distinctive assortment of domains. *Proc Natl Acad Sci U S A.* 91(16):7588-7592.
114. Klein B, Sellos D, Van Wormhoudt A. 1998. Genomic organisation and polymorphism of a crustacean trypsin multi-gene family. *Gene.* 216(1):123–9.

115. Kunitz, M, Northrop, JH. 1935. Crystalline chymo-trypsin and chymo-trypsinogen. I. Isolation, crystallization, and general properties of a new proteolytic enzyme and its precursor. *J Gen Physiol.* 18(4):433-458.
116. Lefevre T, Thomas F, Schwartz A, Levashina E, Blandin S, Brizard JP, et al. 2007. Malaria Plasmodium agent induces alteration in the head proteome of their *Anopheles mosquito* host. *Proteomics.* 7(11): 1908-1915.
117. Lehane MJ, Billingsley PF. 1996. Biology of the insect midgut. London, Chapman & Hall, 486p.
118. Lemos FJA, Cornel AJ, Jacobs-Lorena M. 1996. Trypsin and aminopeptidase gene expression is affected by age and food composition in *Anopheles gambiae*. *Insect Biochem Mol Biol.* 26(7): 651–658.
119. Liu Y, Zhang F, Liu Z, Xiao X, Zhang S, Qin C, et al. 2014. Transmission-Blocking Antibodies against Mosquito C-Type Lectins for Dengue Prevention. *PLoS Pathog.* 13;10(2):e1003931.
120. Lopes AR, Juliano MA, Juliano L, Marana SR, Terra WR. 2006. Substrate specificity of insect trypsins and the role of their subsites in catalysis. *Insect Biochem Mol Biol.* 36:130-140
121. Lu S, Pennington JE, Stonehouse AR, Mobula MM, Wells MA. 2006. Reevaluation of the role of early trypsin activity in the transcriptional activation of the late phase trypsin gene in the mosquito *Aedes aegypti*. *Insect Biochem Mol Biol.* 36(4):336-343.
122. Ludwig GV, Christensen BM, Yuill TM, Schultz KT. 1989. Enzyme processing of La Crosse virus glycoprotein G1: a bunyavirus-vector infection model. *Virology.* 171(1):108–113.
123. Ludwig GV, Israel BA, Christensen BM, Yuill TM, Schultz KT. 1991. Role of La Crosse virus glycoproteins in attachment of virus to host cells. *Virology.* 181(2): 564–571.
124. Mackenzie JS, Gubler DJ, Petersen LR. 2004. Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and Dengue viruses. *Nat Med.* 10:98-109.
125. Mancini E, Tammaro F, Baldini F, Via A, Raimondo D, George P, et al. 2011. Molecular evolution of a gene cluster of serine proteases expressed in the *Anopheles gambiae* female reproductive tract. *BMC Evol Biol.* 11:72.

126. Mann M. 2006. Functional and quantitative proteomics using SILAC. *Nat Rev Mol Cell Biol.* 7(12):952-8.
127. Marinotti O, James AA. 1990. An α glucosidase in the salivary gland of the vector mosquito *Aedes aegypti*. *Insect Biochem.* 20(6):619-623.
128. Marinotti O, Brito M, Moreira CK. 1996. Apyrase alfa-glucosidase in the salivary glands of *Aedes albopictus*. *Comp Biochem Physiol B Biochem Mol Biol.* 113(4):675-679.
129. Mastrobuoni G, Qiao H, Iovinella I, Sagona S, Niccolini A, Boscaro F, et al. 2013. A Proteomic Investigation of Soluble Olfactory Proteins in *Anopheles gambiae*. *PLoS ONE* 8(11): e75162.
130. Megy K, Emrich SJ, Lawson D, Campbell D, Dialynas E, Hughes DS, et al. 2012. VectorBase: improvements to a bioinformatics resource for invertebrate vector genomics. *Nucleic Acids Res.* 40:729-734.
131. Mertens PP, Burroughs JN, Walton A, Wellby MP, Fu H, O'Hara RS, et al. 1996. Enhanced infectivity of modified bluetongue virus particles for two insect cell lines and for two *Culicoides* vector species. *Virology.* 217(2):582–593.
132. Mesquita-Rodrigues C, Saboia-Vahia L, Cuervo P, Levy CM, Honório NA, Domont GB, De Jesus JB: 2011. Expression of Trypsin-like serine peptidases in pré-imaginal stages of *Aedes aegypti* (Diptera:Culicidae). *Arch Insect Biochem Physiol.* 76(4): 223-235.
133. Molina-Cruz A, Gupta L, Richardson J, Bennett K, Black Wt, et al. 2005. Effect of mosquito midgut trypsin activity on dengue-2 virus infection and dissemination in *Aedes aegypti*. *Am J Trop Med Hyg.* 72(5): 631–637.
134. Moll RM, Romoser WS, Modrzakowski MC, Moncayo AC, Lerdthusnee K. 2001. Meconial peritrophic membranes and the fate of midgut bacteria during mosquito (Diptera: Culicidae) metamorphosis. *J Med Entomol.* 38(1):29–32.
135. Moreira-Ferro CK, Marinotti O, Bijovsky AT. 1999. Morphological and biochemical analyses of the salivary glands of the malaria vector, *Anopheles darlingi*. *Tissue Cell.* 31(3):264-73.
136. Muhlia-Almazán A, Sánchez-Paz A, García-Carreño FL. 2008. Invertebrate trypsins: a review. *J Comp Physiol B.* 178(6):655–672

137. Müller HM, Crampton JM, della Torre A, Sinden R, Crisanti A. 1993. Members of a trypsin gene family in *Anopheles gambiae* are induced in the gut by blood meal. *EMBO J.* 12(7):2891–2900.
138. Müller HM, Catteruccia F, Vizioli J, della Torre A, Crisanti, A. 1995. Constitutive and blood meal-induced trypsin genes in *Anopheles gambiae*. *Exp. Parasitol.* 81(3): 371-385.
139. Nayar JK, Sauerman DM. 1971. Physiological effects of carbohydrates on survival, metabolism, and flight potential of female *Aedes taeniorhynchus*. *J Insect Physiol.* 17:2221-2233.
140. Neafsey DE, Waterhouse RM, Abai MR, Aganezov SS, Alekseyev MA, Allen JE, et al., 2015. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. *Science.* 347(6217):43-53.
141. Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, et al. 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science.* 316:1718–1723.
142. Neurath H. 1984. Evolution of proteolytic enzymes. *Science.* 224(4647):350-357.
143. Noriega FG, Pennington, JE, Barrilla-Mury C, Wang WY, Wells MA. 1996. *Aedes aegypti* midgut early trypsin is post-transcriptionally regulated by blood feeding. *Insect Mol Biol.* 5(1), 25-29.
144. Noriega FG, Wang XY, Pennington JE, Barillas-Mury CV, Wells MA. 1996. Early trypsin, a female-specific midgut protease in *Aedes aegypti*: isolation, amino terminal sequence determination, and cloning and sequencing of the gene. *Insect Biochem Mol Biol.* 26(2):119-126.
145. Noriega FG, Wells MA. 1999. A molecular view of trypsin synthesis in the midgut of *Aedes aegypti*. *J Insect Physiol.* 45(7): 613–620.
146. Noriega FG, Edgar KA, Bechet R, Wells MA. 2002. Midgut exopeptidase activities in *Aedes aegypti* are induced by blood feeding. *J. Insect Physiol.* 48(2): 205-212.
147. Northrup JH, Kunitz M. 1931. Isolation of protein crystals possessing tryptic activity. *Science.* 73(1888):262–263.
148. Ong SE, Mittler G, Mann M. 2004. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat Methods.* 1(2):119-26.

149. Organização Mundial da Saúde Saúde/OMS (World Health Organization/WHO) 2015. <http://www.who.int/mediacentre/factsheets/fs117/en/>. Acessado em 05/06/2015
150. Okuda K, de Souza Caroci A, Ribolla PE, de Bianchi AG, Bijovski AT: 2002. Functional morphology of adult female *Culex quinquefasciatus* midgut during blood digestion. *Tissue Cell*. 34(3):210–219.
151. Okuda K, Caroci A, Ribolla P, Marinotti O, de Bianchi AG, Bijovsky AT 2005. Morphological and enzymatic analysis of the midgut of *Anopheles darling* during blood digestion. *J Insect Physiol*. 51(7):769-776.
152. Pandey A, Mann M. 2000. Proteomics to study genes and genomes. *Nature*. 405:837-846.
153. Parish LA, Colquhoun DR, Ubaida Mohien C, Lyashkov AE, Graham DR, Dinglasan RR. 2011. Ookinete-interacting proteins on the microvillar surface are partitioned into detergent resistant membranes of *Anopheles gambiae* midguts. *J Proteome Res*. 10(11): 5150–5162.
154. Paskewitz SM, Brown MR, Lea AO, Collins FH. 1988. Ultrastructure of the encapsulation of *Plasmodium cynomolgi* (B strain) on the midgut of a refractory strain of *Anopheles gambiae*. *J Parasitol*. 74(3): 432–439.
155. Paskewitz SM, Shi L. 2005. The hemolymph proteome of *Anopheles gambiae*. *Insect Biochem Mol Biol*. 35(8):
156. Patterson SD, Aebersold RH. 2003. Proteomics: the first decade and beyond. *Nature Geetics*. 33:311-323.
157. Patil DP, Atanur S, Dhotre DP, Anantharam D, Mahajan VS, Walujkar SA et al. 2009. Generation, annotation, and analysis of ESTs from midgut tissue of adult female *Anopheles stephensi* mosquitoes. *BMC Genomics*. 10:386.
158. Perona JJ; Craik CS. 1995. Structural basis of substrate specificity in the serine proteases. *Protein Sci*. 4(3):337-360.
159. Popova-Butler A, Dean DH. 2009. Proteomic analysis of the mosquito *Aedes aegypti* midgut brush border membrane vesicles. *J Insect Physiol*. 55(3):264–272.
160. Prakash A, Tomazela DM, Frewen B, Maclean B, Merrihew G, Peterman S, et al. 2009. Expediting the development of targeted SRM assays: using data from

- shotgun proteomics to automate method development. *J Proteome Res.* 8(6):2733–2739.
161. Rascón AA Jr, Gearin J, Isoe J, Miesfeld RL. 2011. In vitro activation and enzyme kinetic analysis of recombinant midgut serine proteases from the Dengue vector mosquito *Aedes aegypti*. *BMC Biochem.* 12:43.
 162. Rawlings ND, Barrett AJ. 1993. Evolutionary families of peptidases. *Biochem J* 290: 205-218.
 163. Rawlings ND, Barrett AJ, Bateman, A. 2012 MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res.* 40:343-350.
 164. Réaumur M de. 1740. Mémoires pour servir à l’histoire des insectes. Amsterdam, Chez Pierre Mortier, t.4, pt 2.
 165. Reddy BN, Labbé P, Corbel V. 2012. *Culex* genome is not just another genome for comparative genomics. *Parasit Vectors.* 5:63.
 166. Redi F. 1688. Esperienze intorno alla generazione degl’insecti. 5. Impr. Firenze, Stamparia Piero Matini.
 167. Ribeiro JM, Charlab R, Pham VM, Garfield M, Valenzuela JG. 2004. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. *Insect Biochem Mol Biol.* 34(6):543–563.
 168. Ribeiro JM, Arca B, Lombrado F, Calvo E, Pham VM, Chandra PK, Wikel SK. 2007. An annotated catalogo f salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. *BMC genomics.* 8:6.
 169. Rinderknecht H. 1986. Activation of pancreatic zymogens. *Digestive diseases and sciences.* 31(3): 314-321.
 170. Roach JC, Wang K, Gan L, Hood L. 1997. The molecular evolution of the vertebrate trypsinogens. *J Mol Evol.* 45(6):640-652.
 171. Robich RM, Denlinger DL. 2005. Diapause in the mosquito *Culex pipiens* evokes a metabolic switch from blood feeding to sugar gluttony. *Proc Nat Acad Sci USA.* 102(44):15912–15917.
 172. Romoser WS. 1996. The vector alimentary system. BJ Beaty, WC Marquardt. The Biology of Disease Vectors. 1º edição. Colorado: University

173. Rosenfeld A, Vanderberg JP. 1998. Identification of electrophoretically separated proteases from midgut and hemolymph of adult *Anopheles stephensi* mosquitoes. *J Parasitol.* 84(2):361-365.
174. Ross J, Jiang H, Kanost MR, Wang Y. 2003. Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships. *Gene.* 304:117-131.
175. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, et al. 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics.* 3(12):1154-69.
176. Saboia-Vahia L, Borges-Veloso A, Cuervo P, Junqueira M, Mesquita-Rodrigues C, Britto C, Domont GB, De Jesus JB. 2012. Protein expression in the midgut of sugar-fed *Aedes albopictus* females. *Parasit Vectors.* 5:290.
177. Saboia-Vahia L, Borges-Veloso A, Mesquita-Rodrigues C, Cuervo P, Lopes GD, Britto C, Silva AP, De Jesus JB. 2013. Trypsin-like serine peptidase profiles in the egg, larval, and pupal stages of *Aedes albopictus*. *Parasit Vectors.* 6:50.
178. Saboia-Vahia L, Cuervo P, Borges-Veloso A, de Souza NP, Britto C, Dias-Lopes G, et al. 2014. The midgut of *Aedes albopictus* females expresses active trypsin-like serine peptidases. *Parasit Vectors.* 7:253.
179. Sardelis MR, Turell MJ, Dohm DJ, O'Guinn ML. 2001. Vector competence of selected North American *Culex* and *Coquillettidia* mosquitoes for West Nile virus. *Emerg Inf Dis.* 7(6):1018- 1022.
180. Schaefer CH, Miura T. 1972. Sources of energy utilized by natural populations of the mosquito, *Culex tarsalis*, of overwintering. *J Insect Physiol.* 18(4):797-805.
181. Schechter I, Berger A. 1968. On the active site of proteases. 3. Mapping the active site of papain; specific peptide inhibitors of papain. *Biochem Biophys Res Commun.* 32(5):898-902.
182. Shahabuddin M, Lemos FJA, Kaslow DC, Jacobsloreña M. 1996. Antibody-mediated inhibition of *Aedes aegypti* midgut trypsins blocks sporogonic development of *Plasmodium gallinaceum*. *Infect Immun.* 64(3): 739-743.
183. Serrano-Pinto V, Acosta-Pérez M, Luviano-Bazán D, Hurtado-Sil G, Batista CV, Martínez-Barnetteche, J, et al. 2010. Differential expression of proteins in the

- midgut of *Anopheles albimanus* infected with *Plasmodium berghei*. *Insect Biochem Mol Biol.* 40(10):752-758.
184. Shen Z, Edwards MJ, Jacobs-Lorena M. 2000. A gut-specific serine protease from the malaria vector *Anopheles gambiae* is downregulated after blood ingestion. *Biol.* 9(3): 223-229.
185. Shiao, SH, Higgs S, Adelman Z, Christensen BM. 2001. Effect of prophenoloxidase expression knockout on the melanization of microfilariae in the mosquito *Armigeres subalbatus*. *Insect Mol. Biol.* 10(4):315–321.
186. Simpson JE, Hurtado PJ, Medlock J, Molaei G, Andreadis TG, Galvani AP, et al. 2012. Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile virus as a model system. *Proc Biol Sci.* 279(1730):925-933.
187. Sinka ME, Rubio-Palis Y, Manguin S, Patil AP, Temperley WH, Gething PW, et al. 2010. The dominant *Anopheles* vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasit Vectors.* 3:72.
188. Sinkins SP. 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science.* 316(5832): 1718-1723.
189. Sirot LK, Hardstone MC, Helinski ME, Ribeiro JM, Kimura M, Deewatthanawong P, et al. 2011. Towards a semen proteome of the dengue vector mosquito: protein identification and potential functions. *PLoS Negl Trop Dis.* 5(3):e989.
190. Spiro-Kern A 1974. Untersuchungen über die proteasen bei *Culex pipiens*. *Rev Suisse Zool.* 79:1151-1159.
191. Sreenivasamurthy SK, Dey G, Ramu M, Kumar M, 1, Gupta MK, Mohanty AK, et al. 2013. A compendium of molecules involved in vector-pathogen interactions pertaining to malaria. *Malar J.* 12:216.
192. Sriwichai P, Rongsriyam Y, Jariyapan N, Apiwathnasorn C, Sattabongkot J, Paskewitz S. 2008. Expression of three serine protease genes from the South East Asian malaria vector, *Anopheles dirus*, in relation to blood feeding and parasite infection. *Dev Comp Immunol.* 32(9):1011–1014.
193. Somboon P, Prapanthadara LA. 2002. Trypsin and aminopeptidase activities in blood-fed females *Anopheles dirus* (Diptera: Culicidae) of differing susceptibility to *Plasmodium yoelii nigeriensis*. *Southeast Asian J Trop Med Public Health.* 33(4): 691–693.

194. Souza-Neto JA, Machado FP, Lima JB, Valle D, Ribolla PE. 2007. Sugar digestion in mosquitoes: identification and characterization of three midgut alpha-glucosidases of the neo-tropical malaria vector *Anopheles aquasalis* (Diptera: Culicidae). *Comp Biochem Physiol A Mol Integr Physiol*. 147(4):993-1000.
195. Strode C, Wondji CS, David JP, Hawkes NJ, Lumjuan N, Nelson DR, et al. 2008. Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. *Insect Biochem Mol Biol*. 38:113–123.
196. Tchankouo-Nguetcheu S, Khun H, Pincet L, Roux P, Bahut M, Huerre M, Guette C, Choumet V. 2010. Differential protein modulation in midguts of *Aedes aegypti* infected with chikungunya and dengue 2 viruses. *PLoS One*. 5(10):e13149.
197. Teich N, Ockenga J, Hoffmeister A, Manns M, Mossner J, Keim V. 2000. Chronic pancreatitis associated with an activation peptide mutation that facilitates trypsin activation. *Gastroenterology*. 119(2):461-465.
198. Terra WR, Ferreira C. 1994. Insect digestive enzymes: properties, compartmentalization and function. *Com Biochem Physiol*. 109B: 1–62.
199. Terra WR, Ferreira C, Jordão BP, Dillon RJ. 1996. Digestive enzymes. In: Lehane, M. J., Billingsley, P. F. (Eds) *Biology of the insect midgut*. 1st edition. London: Chapman & Hall.
200. Turell MJ. 2012. Members of the *Culex pipiens* complex as vectors of viruses. *J Am Mosq Control Assoc*. 28(4 Suppl):123-6.
201. Twyman RM. 2014. *Principles of gene manipulation and genomics* (8th edn). Wiley-Blackwell, Oxford UK (forthcoming).
202. Ubaida, MC, Colquhoun DR, Mathias DK, Gibbons JG, Armistead JS, Rodriguez MC, et al. 2012. A bioinformatics approach for integrated transcriptomic and proteomic comparative analyses of model and non-sequenced anopheline vectors of human malaria parasites. *Moll Cell Proteomics*. 12(1):120-131.
203. Valledor L, Jorrín J. 2011. Back to the basics: Maximizing the information obtained by quantitative two dimensional gel electrophoresis analyses by an appropriate experimental design and statistical analyses. *J Proteomics*. 74:1-18.

204. Van Handel E. 1984. Metabolism nutrients in the adult mosquito. *Mosquito News*. 44:573-579.
205. Venâncio TM, Cristofolletti PT, Ferreira C, Verjovski-Almeida S, Terra WR. 2009. The *Aedes aegypti* larval transcriptome: a comparative perspective with emphasis on trypsins and the domain structure of peritrophins. *Insect Mol Biol*. 18(1):33–44.
206. Veitia RA, Caburet S. 2009. Extensive sequence turnover of the signal peptides of members of the GDF/BMP family: exploring their evolutionary landscape. *Biol Direct*. 4(22):1-9.
207. Vizioli J, Catteruccia F, della Torre A, Reckmann I, Müller HM. 2001. Blood digestion in the malaria mosquito *Anopheles gambiae*: Molecular cloning and biochemical characterization of two inducible chymotrypsins. *Eur. J. Biochem*. 268(14): 4027-4035.
208. Volz J, Müller HM, Zdanowicz A, Kafatos FC, Osta MA. 2006. A genetic module regulates the melanization response of *Anopheles* to *Plasmodium*. *Cell Microbiol*. 8(9):1392–1405.
209. Walsh KA, KauVman DL, Sampath-Kumar KSV, Neurath H. 1964. On the structure and function of bovine trypsinogen and trypsin. *Proc Natl Acad Sci USA*. 51:301–308.
210. Wang SJ, Magoulas C, Hickey D. 1999. Concerted evolution within a trypsin gene cluster in *Drosophila*. *Mol Biol Evol*. 16(9): 117-1124.
211. Wasinpiyamongkol L, Patramool S, Luplertlop N, Surasombatpattana P, Doucoure S, Mouchet F, et al. 2010. Blood-feeding and immunogenic *Aedes aegypti* saliva proteins. *Proteomics*. 10(10):1906–1916.
212. Waterhouse RM, Kriventseva EV, Meister S, Xi Z, Alvarez KS, Bartholomay LC, et al. 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science*. 316:1738–1743.
213. Wilkerson RC, Sallum MAM. 1999. *Anopheles* (*Anopheles*) *forattinii*: a new species in Series *Arribalzagia* (Diptera: Culicidae). *J Med Entomol*. 36(3): 345-354.
214. Wilkins MR, Appel, RD, Williams KL, Hochstrasser DF (Eds.) 2007. *Proteome Research: Concepts, Technology and Application*. Springer-Verlag Berlin Heidelberg.

215. Wu DD, Wang GD, Irwin DM, Zhang YP: 2009. A profound role for the expansion of trypsin-like serine protease family in the evolution of hematophagy in mosquito. *Mol Biol Evol.* 26(10):2333-2341.
216. Yang YJ, Davies DM. 1971. Trypsin and chymotrypsin during metamorphosis in *Aedes aegypti* and properties of the chymotrypsin. *J Insect Physiol.* 17(1): 117-131.
217. Yates RA, Steiger S. 1981. Ultrastructural damage of in vitro cultured ookinetes of *Plasmodium gallinaceum* (Brumpt) by purified proteinases of susceptible *Aedes aegypti* (L.). *Z Parasitenkd.* 66(1):93–97.
218. Yates JR. 1998. Mass spectrometry and the age of the proteome. *J Mass Spectrom.* 33(1):1-19.
219. Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, Copley RR, et al. 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. *Science.* 298(5591):149–159.
220. Zhang M, Zheng X, Wu Y, Gan M, He A, Li Z, Zhang D, Wu X, Zhan X. 2013. Differential proteomics of *Aedes albopictus* salivary gland, midgut and C6/36 cell induced by dengue virus infection. *Virology.* 444(1-2):109-118.
221. Zhao Z, Russell PJ. 2003. Trypsin activity assay in substrate-specific one- and two dimensional gels: a powerful method to separate and characterize novel proteases in active form in biological samples. *Electrophoresis.* 24(18): 3284-3288.
222. Zheng L, Whang LHS, Kumar V, Kefatos FC. 1995. Two genes encoding midgut-specific maltase-like polypeptides from *Anopheles gambiae*. *Exp Parasitol.* 81(3):272-283.
223. Zhou D, Zhang D, Ding G, Shi L, Hou Q, Ye Y. 2014. Genome sequence of *Anopheles sinensis* provides insight into genetics basis of mosquito competence for malaria parasites. *BMC Genomics.* 18:15:42.
224. Zhu YC, Kramer KJ, Dowdy AK, Baker JE. 2000. Trypsinogen-like cDNAs and quantitative analysis of mRNA levels from the Indian meal moth, *Plodia interpunctella*. *Insect Biochem Mol Biol.* 30(11):1027–1035.
225. Zhu H, Bilgin M, Snyder M. 2003. Proteomics. *Annu Rev Biochem.* 72:783-812.

8. ANEXOS

Anexo I

Proteínas de *Culex quinquefasciatus* envolvidas em processos biológicos importantes em mosquitos.

Anexo I Proteínas de *Culex quinquefasciatus* envolvidas em processos biológicos importantes em mosquitos

Identificador do UNIPROT	Proteínas Identificadas	Massa Molecular	Peptídeos Únicos	Espectros totais	Cobertura %	Sequência do Peptídeo	Mascot Score
Quimiotripsinas							
B0W653_CULQU	Chymotrypsin 1	25 kDa	8	63	46	(K)AAWVLTASCVNDK(T)	71.81
						(K)RVHPSYDTTGVNNVALLK(L)	22.13
						(R)VHPSYDTTGVNNVALLK(L)	41.33
						(K)AVTLNDAAVTTGIPTIFYGWGSLK(Y)	56.87
						(K)SNILQTLNQK(A)	62.62
						(K)ALSAADCADKVPGLPTGFICAQIQHGQAACAR(D)	18.97
						(K)VPGLPTGFICAQIQHGQAACAR(D)	28.16
						(R)DEGGPLINYNDGK(L)	62.33
B0W174_CULQU	Chymotrypsin-2	29 kDa	4	6	22	(R)VVTFEIHPGYQGGVSPNDIAVIR(V)	82.7
						(K)VVLPLVEYDACYR(L)	52.16
						(R)LWDFSADLAK(S)	46.15
						(R)SACSADSGGPLVK(Q)	25.87
B0WH74_CULQU	Chymotrypsin-2	28 kDa	1	2	4	(R)LESDGALPTR(L)	23.52
Aminopeptidases							
B0WS64_CULQU	Leucyl aminopeptidase	55 kDa				(K)TVISDEATLTK(E)	31.71
						(K)GVTYDTGGADIK(A)	39.89

B0XAQ5_CULQU	Leucine aminopeptidase 1	55 kDa	1	1	4	(K)SVLHPAPTDCSLYLDLAK(L)	33.0
B0W6N8_CULQU	Leucine aminopeptidase	50 kDa	7	25	21	(K)YPAVAVAGLGDASK(W)	53.67
						(R)VEVEDLEDAQAAAEGALLANYK(F)	102.45
						(R)GWILGGAQNWAR(V)	69.41
						(K)SICVDNTDAEGR(L)	41.92
						(R)LILADALCYASTFEPK(F)	93.93
						(K)FIVDIATLTGAIK(V)	123.83
						(R)TLVQFITNAK(-)	60.63
B0WH45_CULQU	Xaa-Pro aminopeptidase 1	59 kDa	6	20	19	(K)GLGSINAYIIPSDDAHQSEYLAK(R)	41.88
						(R)AFISGFDFGSAGTAVVTEK(E)	78.12
						(K)VGVDANLISTR(A)	92.39
						(K)GQFLDTIAR(K)	50.28
						(R)IEDIVQVVSTNIGDNFDGR(G)	79.0
						(R)NALNNYHTTVWETLSPLLK(N)	41.89
B0X730_CULQU	Xaa-pro dipeptidase	53 kDa	3	6	11	(R)LPAEYAVWMGPLLTPDDFK(T)	48.42
						(K)TKYEVDVAVHYADEIDVK(I)	50.46
						(K)FVVDSDILFPVIAECR(V)	54.41
Carboxipeptidases							
B0WS12_CULQU	Zinc carboxypeptidase A 1	49 kDa	2	3	9		

						(R)NIAENYDWHFFPVANPDGYVYTR(T)	26.25
						(K)GNYQTPIAFAYELR(D)	37.46
B0XCT6_CULQU	Glutamate carboxypeptidase	53 kDa	15	20	37		
						(K)TLFSHIDANK(A)	54.34
						(K)YIAALAETVAIK(S)	91.02
						(K)SVSAWPESRPEIFR(M)	38.67
						(R)MVNWVADR(L)	26.17
						(R)TLGATIELADV GK(Q)	89.35
						(R)VIDLPNVILGTLGNDPAK(K)	96.69
						(R)VIDLPNVILGTLGNDPAKK(T)	54.43
						(R)GICYFDVQVCAGK(D)	68.47
						(R)DQLGATQLQHK(E)	42.45
						(R)WRQPSLSIHGIEGAFYEPGQK(T)	16.24
						(R)QPSLSIHGIEGAFYEPGQK(T)	18.45
						(R)GSPNSFAVR(M)	31.42
						(R)EGGSIPVTLTLQQTGK(N)	92.75
						(K)NVL L L P M G A S D D G A H S Q N E K (I)	18.84
						(K)LLGAYLYEVSK(I)	45.85
B0WTW4_CULQU	Carboxypeptidase B	46 kDa	1	1	4		
						(R)EWIAPATALYVIDQLVK(H)	40.58
B0WNR6_CULQU	Vitellogenic carboxypeptidase	53 kDa	2	5	5		
						(K)YVPAVSHAIHR(N)	29.67
						(K)EAGNLVEVLVR(N)	77.57
Dipeptidases							
B0WDL4_CULQU	Dipeptidyl-peptidase 3	82 kDa	7	29	19		

						(K)AQHVLPNTQPIVELECAPAFK(A)	41.74
						(R)IFLGEGQDQLR(K)	46.84
						(R)KPIYLLTER(T)	41.7
						(K)VAEALTEANK(H)	56.15
						(R)GEFEGFVAMVNK(E)	50.44
						(K)FGTLVSNANFIK(L)	60.17
						(K)NVSLGNVLANTNK(T)	52.2
						(K)TDSIPFLSDEDQALLK(K)	26.37
						(K)VLEAGEGLVSVEETEPNK(N)	56.9
						(K)SYESTFDGYIR(S)	37.45
B0X281_CULQU	Alpha-aspartyl dipeptidase	27 kDa	6	19	27		
						(R)WGFDCGIHR(Q)	20.62
						(K)SLYDNALVEPIR(D)	57.37
						(K)SLYDNALVEPIRDR(V)	23.12
						(R)IVQFHELNEAPVLGLR(E)	79.72
						(R)ATLIGVNK(A)	34.62
						(R)GQEPQEFESGADFSFLMK(-)	41.42
Metallo peptidases							
B0W1Y3_CULQU	Protease m1 zinc metalloprotease	113 kDa	11	14	14		
						(R)LPTTITPTR(Y)	33.94
						(K)YLATTQFESTSAR(E)	61.11
						(R)SNALDEATFAQEAGEK(I)	56.6
						(K)ILAALDSHLK(M)	42.07
						(R)NIVGDENWATGLR(L)	52.74
						(K)GANVEHLHEGLQAAIEGK(N)	61.37

						(R)SYDTGDVIISQER(F)	33.59
						(R)AQLIDDAFWLAR(S)	59.64
						(R)MQIIDALGCSR(N)	28.91
						(R)IIQSVYSGSR(V)	54.38
						(R)LGINTLNSAIANIASR(T)	57.9
B0W1Y6_CULQU	Protease m1 zinc metalloprotease	100 kDa	3	3	4		
						(K)NSDLLNIPVIK(F)	38.79
						(R)TSYTTDENVVK(Y)	28.41
						(K)YAGLTQFQACDAR(S)	34.95
B0WH17_CULQU	Protease m1 zinc metalloprotease	120 kDa	4	14	5		
						(K)SWFASTQFSPIDAR(R)	28.62
						(R)VFETFIQNELQK(A)	52.74
						(R)SQLVDDAFNLAR(A)	48.52
						(K)LEEFLNQLIR(D)	50.62
B0W1Y2_CULQU	Protease m1 zinc metalloprotease	101 kDa	1	12	2		
						(R)AQLIDDSLNLAR(S)	66.27
						(R)NSFLFIWDR(M)	40.44
Catepsinas							
B0WI10_CULQU	Cathepsin I	38 kDa	1	10	12		
						(K)LVSLSEQNLVDCSTK(Y)	98.52
						(K)GFVDIPQGDEK(A)	50.15
						(R)DNHCGIATAASYPLV(-)	24.59
B0W0V3_CULQU	Cathepsin L	37 kDa	3	5	13		
						(K)SSYSIASHEQQIQK(E)	43.68
						(R)ILGWGVENDTK(F)	38.87

						(R)GSDHLGIESSIAAGLPK(-)	33.23
Lipase							
B0WDK6_CULQU	Lipase	43 kDa	1	2	5		
						(K)TVVHYAQEIHESGNFQR(F)	32.16
Inibidores de serina peptidases							
B0WQ11_CULQU	Serine protease inhibitor, serpin	72 kDa	1	2	3		
						(K)LFGELVSSEPSLNTPVPWR(Q)	49.21
B0X5Z9_CULQU	Alaserpin	42 kDa	6	23	23		
						(K)SFYSEAEVNFGENDAAAK(K)	62.86
						(K)DLISSDCLGQDTR(M)	53.06
						(R)MVLVNAIHFK(G)	33.55
						(K)LQSINIADLTSK(M)	26.24
						(K)ADFSELLEQPDPLYVSK(V)	60.5
						(R)FAVEHPFYFALLDR(K)	51.84
Detoxificação							
B0XJT6_CULQU	Cytochrome P450 6A1	57 kDa	1	8	7		
						(K)ALFAIVEELASR(M)	71.84
						(K)RVDYLEHLISLR(N)	25.46
						(K)FNFVGVFNIR(Q)	49.43
B0WRU2_CULQU	Cytochrome P450	58 kDa	1	1	2		
						(K)ILEQFVEVFDK(Q)	48.5
B0VZI7_CULQU	Cytochrome P450 6B7	57 kDa	4	6	12		
						(R)VGNVDGPSPLPLFGNTLEQVLGTK(H)	25.9
						(K)DILVGDFASFNK(N)	35.2
						(K)NVAENLLQHVTDSR(G)	25.36
						(K)IGIATLLSK(Y)	27.15

B0XCA1_CULQU	Cytochrome P450 93A3	58 kDa	2	4	5	(R)IGLAYLLTNFR(F)	51.55
						(K)SFILAPEGGLWLK(V)	32.98
B0WTS8_CULQU	Cytochrome P450	49 kDa	1	2	3	(R)LYVDKLDEVAR(E)	35.64
B0WTS9_CULQU	Cytochrome P450	60 kDa	2	4	6	(K)DAHPFLFLPFGGPR(T)	34.64
						(K)ILTALVNIPSNELR(F)	59.06
B0W673_CULQU	Cytochrome p450 family protein 44A1	58 kDa	2	4	6	(K)IYDDFKNEPVVGYFSVR(T)	23.92
						(R)YFASNQFSDLVDEK(S)	64.89
B0X581_CULQU	Cytochrome P450	33 kDa	1	1	4	(R)DVADFYEDIVTR(T)	32.87
B0W6Y4_CULQU	Cytochrome P450 9b1	61 kDa	7	41	19	(K)YTLAGFVEIYNK(Y)	72.83
						(K)VFGLFDNITPIFVVR(D)	67.49
						(R)IFANQFFPTLVGK(L)	52.85
						(K)LGIDVIDR(E)	57.45
						(R)YFSGLIMNAIR(E)	53.0
						(R)NDLINLLLQAR(A)	73.67
						(K)EQHADGFATAESDLGK(V)	31.43
						(R)ATLSPAFTGSK(M)	37.42
B0WTQ1_CULQU (+1)	Cytochrome P450 17A1	61 kDa	3	16	13	(R)RPVFGDNANQNSNVLFSK(T)	31.28
						(K)HQEETESSAGFATVEESHVGK(V)	54.74
						(K)GNINPAAYLPFGVGPR(N)	49.79

						(R)ATLSPAFTGSK(M)	37.42
B0WTP9_CULQU	Cytochrome P450 9b2	62 kDa	1	4	2		
						(K)FLTVVAFPLAR(K)	40.86
B0XJY3_CULQU	Cytochrome P450 6A1	57 kDa	5	31	12		
						(K)TLFAIVEALASR(M)	94.68
						(K)LVEPSFGSLLTVILR(S)	52.69
						(K)RADYLEHLISLR(N)	35.51
						(R)ADYLEHLISLR(N)	37.52
						(R)FSPETGGVNPYR(E)	30.61
						(K)FNFVGVFNIR(Q)	49.43
B0WTS7_CULQU	Cytochrome P450 98A1	60 kDa	6	18	16		
						(R)YHNLPIIDHR(R)	41.17
						(R)FREDFGDLLVIPGILGR(K)	39.91
						(K)GLGGLVNEQGESWQQFR(T)	103.29
						(K)NELPADFNQWLN(R)	51.8
						(R)FLFLPFGGPR(A)	66.59
						(R)LVSNLINAPANELR(Y)	65.04
B0WFB9_CULQU	Cytochrome P450 11A1, mitochondrial	59 kDa	4	22	13		
						(K)SDPLSGHLFLLEGAPWR(A)	31.19
						(K)VFEQTPFQLAK(F)	49.48
						(K)IGLITLLR(S)	56.09
						(R)FKPTAATPNPLVDFPK(S)	35.41
						(K)SFVLSPIGGNHLR(V)	20.45
B0WFC1_CULQU	Cytochrome P450	57 kDa	3	13	11		
						(K)NNVVRPDFLQLLMQLK(N)	30.1
						(R)IANQPYQLTSPNVVIEK(G)	63.68
						(R)HTHTFLPFGDGPR(N)	21.43

						(K)FGIAQLLSR(L)	32.74
B0WTS5_CULQU	Cytochrome P450 12b1, mitochondrial	60 kDa	2	14	19		
						(R)SVQAQPVSSTSEGHGEVEYTNALPYSK(I)	36.56
						(R)IALDTFVHYR(K)	30.2
						(R)YNNLTFPDLHR(S)	34.86
						(R)LDKNPSQSDSQSVLEK(L)	21.65
						(K)GTDVAMASMILHSGEEYFER(G)	42.26
						(K)DVHPFLFLPFGFGPR(T)	52.41
B0WTS3_CULQU	Cytochrome P450 12b1, mitochondrial	60 kDa	1	13	17		
						(R)SVQAQPVSSTSEGHGEVEYANALPYSK(I)	36.55
						(R)YNNLTFPDLHR(S)	34.86
						(R)LDKNPSQSDSQSVLEK(L)	21.65
						(K)GTDVAMASMILHSGEEYFER(G)	42.26
						(K)DVHPFLFLPFGFGPR(T)	52.41
B0WIH9_CULQU	Carboxylesterase-6	71 kDa	3	13	7		
						(K)AQLIATTVGCPPTSSTK(M)	46.41
						(R)YLYNPFSPPLGVVIEK(Q)	73.02
						(R)EEYLLEIDER(W)	36.06
B0W1X8_CULQU	Palmitoyl-protein thioesterase 1	34 kDa	2	6	13		
						(R)MNNLITLGGQHQGVFGLPDCPSISSK(T)	37.08
						(K)ESSTFLADINNER(T)	79.47
A5Y5K2_CULQU (+2)	Esterase A11	61 kDa	1	4	3		
						(R)ILFSFGPVIEPYVTK(N)	38.09

B0W5S7_CULQU	Ubiquitin thioesterase OTUB1	32 kDa	2	4	12	(K)VKPEGDNTEAALAELHK(L)	27.92
						(K)LLNEQGHSDYVVVYLR(L)	31.55
R9RIG8_CULPP	Esterase A4-B4, B	61 kDa	2	93	39	(R)AIVMSGSTYSSWSLTR(Q)*	71.03
						(R)AAKPEDIVAHQEK(L)*	54.62
						(R)FKAPVPPQK(W)	27.51
						(K)WTETLDCTQQCEPCYHFDR(R)	27.43
						(K)IVGCEDSLK(I)	42.2
						(K)DIVLVSFNYR(I)	36.45
						(R)IGALGFLCCQSEQDGVPGNAGLK(D)	17.83
						(R)WVLENIAAFGGDPK(R)	79.33
						(K)RVTLAGHSAGAASVQYHLISDASK(D)	38.33
						(R)VTLAGHSAGAASVQYHLISDASK(D)	80.85
						(K)AIGWDGQGGESGALR(F)	73.86
						(R)VFWHGLHR(T)	23.29
						(R)ICLDSEFYNHYR(I)	82.73
						(R)GTAHADELSYLFNFTQQVPGK(E)	18.66
						(K)GGVVFEFNAQTKPTFK(C)	28.69
EST1_CULPI	Esterase B1	61 kDa	1	35	32	(K)LHPELLSHPHLFLGNVPPNLK(I)*	28.32
						(K)WTETLDCTQQCEPCYHFDR(R)	27.43
						(R)IGALGFLCCQSEQDGVPGNAGLK(D)	17.83
						(R)WVLENIAAFGGDPK(R)	68.44
						(K)RVTLAGHSAGAASVQYHLISDASK(D)	38.33
						(R)VTLAGHSAGAASVQYHLISDASK(D)	80.85
						(K)AIGWDGQGGESGALR(F)	72.12

						(R)VFWHGLHR(T)	20.85
						(R)ICLDSEFYNHRY(I)	25.93
						(R)GTAHADELSYLSNFTQQVPGK(E)	19.41
						(K)GGVVFEPNAQTKPTFK(C)	28.69
B0XGK3_CULQU	Glutathione-s-transferase theta, gst	24 kDa	1	1	7		
						(K)ATNQEGAEELAALYR(A)	47.22
B0W6C9_CULQU	Glutathione S-transferase	23 kDa	1	8	10		
						(R)LYFDMGTLYQR(F)	47.39
						(K)VAAGLDLSR(Y)	41.36
B0XGJ5_CULQU	Glutathione-s-transferase theta, gst	25 kDa	2	7	7		
						(K)RLEQVIPDYHEFNTVR(A)	35.88
						(R)LEQVIPDYHEFNTVR(A)	20.08
B0X3C8_CULQU	Glutathione-s-transferase theta, gst	27 kDa	1	1	4		
						(K)LAESVAILR(Y)	28.16
B0WFX0_CULQU	Glutathione s- transferase	27 kDa	3	29	31		
						(K)SDLVANHPNLQR(L)	41.87
						(K)LVTLNNEVIPFYLEK(L)	51.56
						(R)LVDNVTSIDSIAAWIAK(R)	73.63
						(K)KLVTLNNEVIPFYLEK(L)	24.05
						(K)IAVVSYPDDDVKEK(K)	31.21
						(K)IAVVSYPDDDVKE(E)	41.62
						(R)FLLSYGNLPPFDDIR(I)	28.61

B0VZ90_CULQU	Glutathione S-transferase	30 kDa	2	9	16	(R)IHLILDAK(N)	34.47
						(R)KLYPADPFQK(A)	16.34
						(R)FNGAVISPYR(I)	26.78
						(K)FALGDKYELDK(E)	32.28
B0W6C9_CULQU	Glutathione S-transferase	23 kDa	1	8	10	(R)LYFDMGTLYQR(F)	47.39
						(K)VAAGLDLSR(Y)	41.36
B0XAJ0_CULQU	Glutathione S-transferase 1	25 kDa	4	11	27	(M)PITLYTPISPPAR(A)	64.69
						(R)ELGLNVEFKPVDVMAGGTR(T)	41.09
						(R)VLNHDLISAFYPK(T)	21.05
						(K)LDEGLTNLELFLVR(N)	48.81
B0W6C3_CULQU	Glutathione S-transferase 1	25 kDa	1	2	5	(K)VIDNSVQGAAR(E)	67.73
B0XG85_CULQU	Microsomal glutathione S-transferase 1	17 kDa	2	3	15	(M)TSVFDSIDSTVFR(S)	47.2
						(K)VVTNDPDPVER(V)	33.85
C4B4V7_CULQU	Glutathione transferase	25 kDa	1	33	31	(K)EAPGAALNEAGLEDFK(A)	90.77
						(R)AIQIYLVEK(Y)	32.38

						(R)AVQMTAAAVGVELNLK(L)	46.95
						(K)LTNLMAGEHMKPEFLK(I)	27.29
						(R)LFFDQGTLYQR(F)	70.98
C4B4V8_CULQU	Glutathione transferase	23 kDa	2	65	62		
						(K)EAPGAAINEAGCEEFK(K)	26.01
						(K)EAPGAAINEAGCEEFK(Y)	19.9
						(K)LNPQHCIPTLVDGSFPVWESR(A)	43.43
						(R)AIMIYLVEK(Y)	49.23
						(R)FADYFYPQIFAK(Q)	57.81
						(K)MLDGLDFLNTFLGGSK(Y)	90.32
						(K)YPNVAGWYAR(L)	43.53
						(R)AVQMTAAAVGVELNLK(L)	46.95
						(K)LTNLMAGEHMKPEFLK(L)	27.29
						(R)LFFDQGTLYQR(F)	64.43
B0WFX0_CULQU	Glutathione s-transferase	27 kDa	3	29	31		
						(K)SDLVANHPNLQR(L)	41.87
						(K)LVTLNNEVIPFYLEK(L)	51.56
						(R)LVDNVTSIDSIAAWIAK(R)	73.63
						(K)KLVTLNNEVIPFYLEK(L)	24.05
						(K)IAVVSYPEDDDVKEK(K)	31.21
						(K)IAVVSYPEDDDVK(E)	41.62
						(R)FLLSYGNLPPFDDIR(I)	28.61
B0W6B0_CULQU	Glutathione S-transferase 1-6	82 kDa	0	64	16		
						(K)LNPQHCIPTLVDGSFPVWESR(A)	43.43
						(R)AIMIYLVEK(Y)	49.23

						(R)FADYFYPQIFAK(Q)	57.65
						(K)MLDGLDFLNTFLGGSK(Y)	90.32
						(K)YPNVAGWYAR(L)	43.53
						(R)AIQIYLVEK(Y)	32.38
						(R)AVQMTAAAVGVELNLK(L)	46.95
						(K)LTNLMAGEHMKPEFLK(L)	27.29
						(R)LFFDQGTLYQR(F)	64.43
<hr/>							
BOWSC1_CULQU	Glutathione synthetase	56 kDa	3	6	12		
						(R)ETYPFVQVIR(R)	40.29
						(R)AGYEPGHYYGPNEWSAR(L)	30.23
						(R)EGGGNNVYGEDIPGALEK(M)	56.43
						(K)VDAFTGSLFEIYETVLK(E)	49.08
<hr/>							
Carboidrases							
BOXJR7_CULQU	Alpha-glucosidase	70 kDa	11	104	28		
						(R)ELVSDWWEK(A)	47.03
						(K)DSNGDGIGDLK(G)	40.0
						(K)AFWLSPIYK(S)	48.76
						(K)LILDFVPNHSSDEHEWFVK(S)	44.92
						(R)VAGYEDYYVWNDGIVGSDGQR(S)	97.01
						(R)SPPNNWNEAFR(G)	41.09
						(R)GSAWQWSATR(Q)	57.05
						(R)NPAVVEAMK(N)	35.82
						(R)IDAVPWLFEDEQLR(D)	60.47
						(R)EVLDEYKK(A)	41.6
						(K)TVIDSWLDAVPVGHAPNWWLGNHDK(R)	49.3
						(K)SFDTAVVR(V)	34.53

						(K)AGFYQIYPR(S)	57.9
						(K)KQPDLNYR(N)	40.14
B0X220_CULQU	Maltase 1	69 kDa	1	3	6	(K)DLGVDGVWFSPFK(S)	39.49
						(K)SPMADFGYDISDFR(D)	58.0
						(K)KQPDLNYR(N)	40.14
B0XAA1_CULQU	Alpha-glucosidase	67 kDa	12	72	38	(K)DSDGNGIGDLGVTKE(L)	54.11
						(R)LILDFVPNHTSNQHEYFK(K)	32.34
						(K)SVQKEDPYTNYVWHPGVTNEDGTK(S)	32.65
						(K)EDPYTNYVWHPGVTNEDGTK(S)	66.02
						(K)SPPSNWISVFR(G)	60.33
						(R)GSAWEWNEDRQEYLLHQFLK(E)	35.38
						(R)IDAVPYLFESVEVENR(Y)	70.32
						(R)YVDEPLSR(A)	38.29
						(R)AVLDEFTAK(D)	46.39
						(R)IMMTEGYTSIPNVMK(F)	32.69
						(R)YANLWLDNLPAGR(R)	85.38
						(R)SNWVLGNHDNNR(I)	33.62
						(R)AGPDDYTLYSR(D)	66.21
						(K)NAGFSSADK(T)	29.59
						(R)QILTHGDLDLR(T)	42.02
						(R)TTDSNLVLYK(R)	42.89
						(K)WINVADTELPGESAIVLQK(L)	71.85
B0XBN3_CULQU	Alpha-glucosidase	72 kDa	28	176	52	(K)VVQTITLGR(D)*	37.96
						(K)SIDGEEFEITK(D)*	45.83
						(K)DVDESGLFTVAR(K)*	108.01

						(K)ISGANWYGGPQQK(Y)*	52.15
						(K)YQYWPIQK(L)*	32.48
						(K)EADNCAVGDR(Y)*	52.64
						(R)YWLNSLGSFVFDVDR(T)*	33.12
						(R)TAPLFVDQNYGEPGYLCLEAK(K)*	46.72
						(K)KTLPYDVYDATYSFEYK(V)*	31.78
						(K)TLPYDVYDATYSFEYK(V)*	33.76
						(K)VGVANHAK(A)*	53.81
						(K)AVGHILGKPTGYPAEEMVK(Y)*	40.66
						(K)YPIWSTWAR(Y)*	40.74
						(R)DIDHDLVLK(F)*	47.56
						(R)VTIWIHPFINK(G)*	40.47
						(R)LANVLAVSGIDSFK(F)*	113.3
						(K)FDAGEASWSPPDALLSGDR(K)*	86.3
						(K)FDAGEASWSPPDALLSGDRK(D)*	53.56
						(R)KDNPNIVDYLR(T)*	50.57
						(K)DNPNIVDYLR(T)*	81.44
						(R)TVASTGGNLVEVR(T)*	108.94
						(R)TAHNTQDLPVFVR(M)*	61.37
						(K)YTPQIMAR(F)*	29.22
						(K)LAVSDGLPVNPPLWWISPDDTNAQK(V)*	27.23
						(K)VFDQFLLGDDIIAAPVIR(E)*	61.8
						(R)SRDIYLPAGTWIDGNSNAEHTGPK(W)*	34.99
						(R)DIYLPAGTWIDGNSNAEHTGPK(W)*	65.81
						(R)NYSVPLSMLPYFVR(K)*	46.8
BOWS02_CULQU	Alpha-glucosidase	68 kDa	7	44	17	(R)GSAWEWSNLR(E)	35.31

						(R)EQYYLHQFTVEQPDLNYR(N)	35.53
						(K)DVITFWLDK(G)	45.47
						(R)IDAVPFLFEVEK(N)	61.4
						(R)ELVDAYQETHGGETR(V)	53.66
						(R)LGSHMPFNFGLITDITK(A)	30.53
						(R)VVNDNVLAIVR(E)	51.6
						(K)AGFYQIYPR(S)	57.9
BOWS45_CULQU	Glucosidase 2 subunit beta	63 kDa	4	8	9		
						(K)SDAEALESVALK(V)	91.87
						(R)DGAVTVVEAK(Y)	31.44
						(K)VEYDPDTAELIR(K)	70.87
						(R)SVLVHLECGLDTR(I)	34.44
BOWQR9_CULQU	Neutral alpha-glucosidase AB	106 kDa	1	1	1		
						(K)LLVAPVLK(A)	25.2
BOWFP7_CULQU	Oligo-1,6-glucosidase	66 kDa	2	132	29		
						(K)LVLDFVPNHSSDEHANFLK(S)*	51.45
						(R)EAGYEDYYLWHPGK(L)*	57.4
						(R)SFMSDSDGVDGDLR(G)	63.08
						(R)QELGVDAIWLSPIFK(S)	126.28
						(K)SPMADFGYDISDFK(D)	91.0
						(R)VEPSNWISVFR(G)	71.7
						(K)VVQDMKDVLTFWLQK(G)	36.12
						(K)DVLTFWLQK(G)	52.26
						(R)AGAQIPFNFEVLSNIFK(D)	62.61
						(K)DSTAQDFYDNAMR(F)	54.46
						(K)ALPSDQFANWVLGNHDNK(R)	20.76

						(K)ALPSDQFANWVLGNHDNKR(L)	29.72
						(K)VVDNMLIYK(R)	37.06
						(K)KQPDLNYR(S)	40.14
B0W934_CULQU	78 kDa glucose-regulated protein	72 kDa	9	211	43		
						(R)EFTDATVQGDIK(L)*	78.87
						(K)DAGTIAGLNVMR(I)*	86.48
						(R)ALSSSHQVR(I)*	42.43
						(K)VLEDADMNKK(D)*	41.34
						(K)DVDEIVLVGGSTR(I)*	73.06
						(K)WLDENQDAESEYKQ(Q)*	35.26
						(K)KELEDVVQPIIAK(L)*	49.42
						(K)ELEDVVQPIIAK(L)*	72.51
						(K)LYASSGGAPPPAGGEDLDKDEL(-)*	60.63
						(R)ITPSYVAFTADGER(L)	81.82
						(K)NQLTTNPENTVFDK(R)	85.04
						(K)VFAPEEISAMVLGK(M)	93.4
						(K)VTHAVVTPAYFNDAQR(Q)	77.3
						(R)AKFEELNMDLFR(S)	78.95
						(K)FEELNMDLFR(S)	65.68
						(K)SQIFSTASDNQHTVTIQVYEGERPMTK(D)	23.68
						(K)FDLTGIPPAPR(G)	40.27
						(K)IVITNDQNR(L)	53.85
						(R)NELESYAYSLK(N)	74.04
						(K)MEEAIDEK(I)	39.09
						(K)MKETAEAYLGK(K)	58.38
						(K)ETAEAYLGK(K)	51.41
						(R)IINEPTAAAIAYGLDKK(D)	37.96

						(R)VEIANDQGNI(I)	63.31
						(R)IINEPTAAAIAYGLDK(K)	136.31
BOXGHO_CULQU	Alpha-amylase	69 kDa	1	17	6	(R)AGFYQVYPR(S)	44.58
						(R)LITELTADSSAYDFK(N)	112.39
						(K)SPMADFGYDISDFR(D)	58.0
BOXJR6_CULQU	Alpha-amylase	67 kDa	7	169	48	(K)DIHHEFGTIADLESAAECK(A)*	77.15
						(K)LVLDVFPNHSSDEHVNFK(S)*	39.98
						(K)SVNREEGYEDYYLWHPGK(L)*	50.15
						(R)GSAWEWNDVR(K)*	56.18
						(R)IDAVPYLFEIEMTDGQYPDEPK(T)*	34.61
						(K)TGWTNDPTNPDYVQHIYTQNR(E)*	58.01
						(K)DTVDPAGCNTNPK(E)*	57.87
						(R)NSFELPADAGIVLEGSVK(N)*	60.07
						(R)SFMDSGDGVGDLR(G)	63.08
						(R)QELGVDAIWLSPFK(S)	126.28
						(K)SPMADFGYDISDFK(D)	91.0
						(R)VEPSNWISVFR(G)	71.7
						(K)VVDQMKDVLTFWLQK(G)	36.12
						(K)DVLTFWLQK(G)	52.26
						(R)AGAQIPFNFEVLSNIFK(D)	62.61
						(K)DSTAQDFYDNAMR(F)	54.46
						(K)ALPSDQFANWVLGNHDNK(R)	20.76
						(K)ALPSDQFANWVLGNHDNKR(L)	29.72
						(K)VVDNMLIYK(R)	37.06
						(K)KQPDLYR(S)	40.14

Resposta Imune

B0WIF2_CULQU	Salivary C-type lectin	22 kDa	1	2	7	(R)LASVNSAEDDAALK(L)	57.71
B0WIE3_CULQU (+1)	Salivary C-type lectin	18 kDa	1	3	7	(R)GPWWIAGTDLGK(S)	42.51
B0WID7_CULQU	Salivary C-type lectin	21 kDa	1	1	8	(R)LASVNSAEDDTALR(L)	49.25
B0XG12_CULQU	Truncated ER mannose-binding lectin	53 kDa	4	6	11	(R)IGADGLAFWYTAEK(G)	87.66
						(R)AENVVLPK(N)	32.01
						(R)HEVDALLQNQNVLAQTVR(E)	90.66
						(R)ADSILNNQAR(A)	42.72
B0XEF4_CULQU	Galectin	16 kDa	3	40	33	(K)MTHDQFNINLQAGPNVNPR(D)	90.94
						(R)DDAPLHISIRPK(D)	53.92
						(R)FIHIGEGAQVDAMITE(-)	30.91
B0W0J6_CULQU	Galectin	36 kDa	5	8	19	(R)VNINFLSGK(S)	46.15
						(R)EDNLNELTAPNPVSPGEIFR(L)	40.18
						(R)LYILVGDDR(F)	26.87
						(R)SGGFPPVLDQQFK(L)	29.04
						(K)LAIALTESEFK(F)	78.47
B0X1F3_CULQU	Galectin	17 kDa	3	25	25	(R)TLGDGDELELK(G)	61.41
						(R)FGSGIAYHFGVR(L)	44.25
						(R)LSGDEIWILVDGEK(Q)	58.14
B0WLS0_CULQU	Arginine kinase	45 kDa	9	182	45		

						(K)TDKHPASDFGDVSTFGNVDPTGEFVVSTR(V)*	64.64
						(K)HPASDFGDVSTFGNVDPTGEFVVSTR(V)*	58.72
						(R)SMEGYPFNPCLTEAQYK(E)*	68.84
						(K)VSATLSGLEGELK(G)*	100.37
						(K)AVQQQLIDHFLFK(E)*	49.91
						(R)FLQAAACR(Y)*	72.52
						(K)TFLVWCNEEDHLR(I)*	49.08
						(R)IISMQMGGDLGQVYR(R)*	74.33
						(R)RLTSAVNEIEK(R)*	36.89
						(R)LTSAVNEIEK(R)*	53.06
						(K)RVPFSHNR(L)*	48.93
						(R)VPFSHNR(L)*	41.29
						(R)LGFLTCPTNLGTTIR(A)*	100.92
						(R)GEHSEAEGGIYDISNK(R)*	86.19
						(R)MGLTEYEAVK(E)*	58.98
						(K)EMYDGISELIK(I)*	28.2
BOWR64_CULQU	Fructose-bisphosphate aldolase	39 kDa	12	213	67		
						(R)QLFTADER(L)*	50.68
						(R)LSQNISGVILFHETLYQK(A)*	54.99
						(K)ADDGTPLVELLK(K)*	75.94
						(K)GVVDLYGSEGETTQGLDDLGAR(C)*	97.51
						(K)NTPSYQSILENANVLAR(Y)*	103.34
						(R)YASICQSQR(I)*	49.99
						(K)ALSDHHVYLEGTLKPNMVTAGMSCAK(K)*	50.49

						(K)KPTAQEIGLATVLALR(R)*	78.96
						(K)AAQDELVK(R)*	57.41
						(K)AAQDELVKR(A)*	35.97
						(K)ANGEASLGK(Y)*	31.28
						(K)YGGGVVGAAGTGSFLIANHAY(-)*	41.02
						(K)GILAADESTATCGK(R)	98.93
						(R)FADIGVENNEDNR(R)	78.56
						(R)FADIGVENNEDNRR(Q)	44.6
						(K)KGILAGIK(V)	56.91
						(R)IVPIVEPEILPDGDHDLER(C)	78.31
						(K)VTETVLAAYK(A)	95.94
						(R)ALQASVLR(A)	52.58
BOWTE9_CULQU	F-actin capping protein subunit beta	31 kDa	5	11	24		
						(R)RLPPQIEK(N)	25.74
						(R)KLEIEANHAFDQYR(E)	35.31
						(K)QSSGTINLGGSLTR(Q)	38.81
						(R)QIEQDAPVSESPHIANIGR(I)	31.79
						(K)DLAAALLR(R)	39.19
BOWPC7_CULQU	Leucine-rich transmembrane protein	56 kDa	3	6	9		
						(R)LTADVLRPDVFR(G)	29.44
						(R)NPLGEITEDTAAALGSIK(L)	72.28
						(K)HILNDEFQAPHSEAR(C)	23.66
BOXOG6_CULQU	Leucine-rich transmembrane protein	52 kDa	2	3	5		

						(K)FVEVPLFSGDK(L)	28.38
						(K)LEEYALFGLPR(L)	47.75
B0WWMF9_CULQU	Signal recognition particle subunit SRP68	70 kDa	1	3	6		
						(K)VFTIEILR(V)	26.65
						(R)LFLLSIQDLANSIEK(A)	47.27
						(K)IELLENVLLDCK(D)	29.59

Anexo II

Borges-Veloso et al. 2015. In-depth characterization of trypsin-like serine peptidases in the midgut of the sugar fed *Culex quinquefasciatus*. Parasites & Vectors.