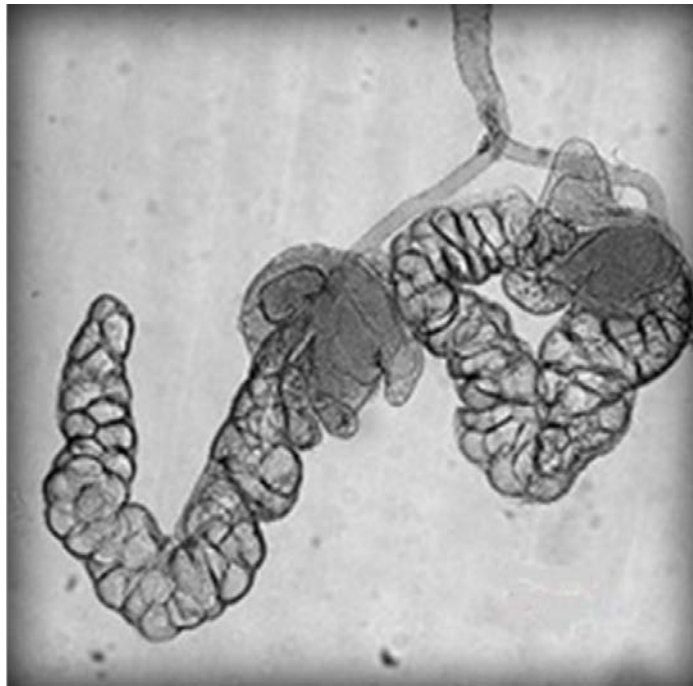


**Ministério da Saúde**  
**Fundação Oswaldo Cruz**  
**Centro de Pesquisa René Rachou**  
**Programa de Pós-graduação em Ciências da Saúde**

**Estudo da glândula salivar do principal vetor da oncocercose no Brasil, *Thysopelma guianense* (Diptera: Simuliidae): Aspectos morfológicos, bioquímicos e moleculares**



**por**

**Andrezza Campos Chagas**

**Belo Horizonte**

**Fevereiro/2011**

**Ministério da Saúde**  
**Fundação Oswaldo Cruz**  
**Centro de Pesquisa René Rachou**  
**Programa de Pós-graduação em Ciências da Saúde**

**Estudo da glândula salivar do principal vetor da oncocercose no Brasil, *Thyrsopelma guianense* (Diptera: Simuliidae): Aspectos morfológicos, bioquímicos e moleculares**

**por**

**Andrezza Campos Chagas**

**Tese apresentada com vistas à obtenção do título de Doutor em Ciências na área de concentração em Doenças Infecciosas e Parasitárias.**

**Orientador: Dr. Paulo Filemon Palloucci Pimenta**

**Belo Horizonte**

**Fevereiro/2011**

Catálogo-na-fonte  
Rede de Bibliotecas da FIOCRUZ  
Biblioteca do CPqRR  
Segemar Oliveira Magalhães CRB/6 1975

C426e

2011

**Chagas, Andrezza Campos.**

Estudo da glândula salivar do principal vetor da oncocercose no Brasil, *Thyrsopelema guianense* (Diptera: Simuliidae): Aspectos morfológicos, bioquímicos e moleculares / Andrezza Campos Chagas. – Belo Horizonte, 2011.

xxii, 162 f.: il.; 210 x 297mm.

Bibliografia: f. 157 - 184

Tese (doutorado) – Tese para obtenção do título de Doutor em Ciências pelo Programa de Pós-Graduação em Ciências da Saúde do Centro de Pesquisas René Rachou. Área de concentração: Doenças Infecciosas e Parasitárias.

**1. Oncocercose/transmissão 2. Simuliidae/parasitologia 3. Glândulas Salivares/secreção 4. Proteínas e Peptídeos Salivares/análise I. Título. II. Pimenta, Paulo Filemon Paolucci (Orientação).**

CDD – 22. ed. – 616.965 2

**Ministério da Saúde  
Fundação Oswaldo Cruz  
Centro de Pesquisa René Rachou  
Programa de Pós-graduação em Ciências da Saúde**

**Estudo da glândula salivar do principal vetor da oncocercose no Brasil, *Thyrsopelma guianense* (Diptera: Simuliidae): Aspectos morfológicos, bioquímicos e moleculares**

**por**

**Andrezza Campos Chagas**

**Foi avaliada pela banca examinadora composta pelos seguintes membros:**

**Prof. Dr. Paulo Filemon Paolucci Pimenta (Presidente)  
Prof. Dr. Breno de Melo Silva  
Prof. Dra. Vanessa Cabreira de Freitas  
Prof. Dr. Marcos Horácio Pereira  
Suplente: Prof. Dr. Ricardo Nascimento Araújo**

**Tese defendida e aprovada em: 28/02/2010**



### **Dedicatória**

Para as duas pessoas mais importante na minha vida

“MEUS PAIS (EDSON & SANDRA)”

À minha querida AVÓ “FLORIPES”

Aos meus irmãos “STEFFEN & ABRAÃO” e as minhas irmãs “SHEILA & EDSANDRA”

Às minhas sobrinhas e sobrinhos “STEFFANY, RAQUEL, EMANUELLE, ISAAC, SOPHIA,

VALENTINA, LUIDI e ANA GABRIELA”

Ao meu cunhado “NETO” e as minhas cunhadas “MÁRCIA e GRACIANA”

.... e ao meu caozinho PAVEL....

## **Agradecimentos**

- ✓ Fundação de Amparo a Pesquisa do Estado do Amazonas (FAPEAM) pela bolsa de doutorado, edital N.19/2006-CAPES/FAPEAM- Programa de Apoio à Formação de Recursos Humanos Pós-Graduados do Estado do Amazonas-RHPOSGRAD.
- ✓ FUNDEP - Processos nº15743, nº16717
- ✓ FAPEMIG, Processo APQ-03790-10
- ✓ CNPQ, Processo 302249/2008-7, Edital Produtividade de Pesquisa – PQ-2008 e Processo 576018/2008-2, Edital MCT/CNPq/CT-Saúde/MS/SCTIE/DECIT nº34/2008

## Agradecimentos

- ✓ Ao “cara” lá de cima.. que nunca me abandonou... e que me deu coragem para superar “TODAS” as “BARREIRAS”... e que me deu “GARRA” para “NUNCA” desistir.. mesmo quando o “CLARO” parecia “ESCURO”, quando o “FÁCIL” parecia “DIFÍCIL”.. e quando o “POSSÍVEL” a cada dia parecia “IMPOSSÍVEL”...e que **“FEZ O SONHO SE TORNAR REALIDADE”...**
- ✓ A MINHA FAMÍLIA **“CHAGAS”**... que sempre me deu força quando eu ESTAVA PERTO e quando eu ESTIVE LONGE (por motivo do trabalho, claro!). E que me APOIOU mesmo quando transformei o nosso “CANTINHO” num verdadeiro INSETÁRIO..O meu MUITO OBRIGADA!
- ✓ A minha querida “Avó FLORIPES”.. por todas às ORAÇÕES... e aos meus tios, tias e primos pelo carinho...
- ✓ Ao meu “ORIENTADOR PAULO” pelo CONVITE, pela CONFIANÇA, pelos ENSINAMENTOS, pela AMIZADE e principalmente pelas MENSAGENS DE INCENTIVO.
- ✓ Ao Dr JOSÉ RIBEIRO” pela OPORTUNIDADE de trabalho juntos, pelo INCENTIVO, pela CONFIANÇA e pelas “VIDAS” no VIDEO GAME do SIALOTRANSCRIPTOMA... e pelo CONVITE do PÓS.
- ✓ Ao Dr ERIC CALVO” pela RECEPÇÃO no NIH, pela DISPONIBILIDADE e ACESSIBILIDADE, pelos “ENSINAMENTOS” em biologia molecular, pelos “MOMENTOS de BANCADA”, pelo APOIO, RESPEITO e principalmente pela AMIZADE.
- ✓ Ao Centro de Pesquisa René Rachou (CpQRR) – FIOCRUZ MINAS e ao Instituto Leônidas e Maria Deane (ILMD) – FIOCRUZ AMAZÔNIA, pelas intalações oferecidas e principalmente pela disponibilização de carros para o trabalho de campo.
- ✓ À Fundação de Amparo à Pesquisa do Estado do Amazonas, pela BOLSA DE ESTUDOS e pelos benefícios oferecidos pelo Programa de Apoio à Formação de Recursos Humanos Pós-Graduados do Estado do Amazonas – RH –POSGRAD.
- ✓ À Biblioteca do CPqRR em prover acesso gratuito local e remoto à informação técnico-científica em saúde custeada com recursos públicos federais, integrante do rol de referências desta tese, também pela catalogação e normalização da mesma.
- ✓ Thanks to the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.
- ✓ À Mineração Taboca S/A pelo acesso e apoio nas coletas de simulídeos. Agradeço à todos os amigos que fiz, e que muito me ajudaram nas coletas: Assis, Dultevir, seu Zé, Capeta, Colombiano, Lima, Arnaldo, Tatu, Joseildo, Torquato, Zé Filho, Gago, Beth e aos nomes que esqueci, peço desculpas, mas manifesto o meu MUITO OBRIGADA.
- ✓ Aos companheiros de coleta: Arturzinho, seu Zé, Joseildo, Capeta, Jansinho, Joãozinho e Jordan. Sem essas pessoas.. realmente eu teria morrido nas cachoeiras, heheh..

- ✓ A equipe da Fiocruz Amazônia, Dr. Roberto Sena, Soraya Moresi, Marizete, pela liberação de carro para as coletas e pela ajuda burocrática no envio de amostras biológicas, e aos motoristas Ricardo e Frank pelas horas de estrada rodada para a coleta de simulídeos.
- ✓ Aos amigos “simulidólogos” Dr. Jansen, Dr. Felipe, Dr. Vitor, MSc. Cláudia, Ulisses, Orlando e Júnior pelas experiências compartilhadas e pela amizade.
- ✓ Ao meu grande e inseparável aprendiz de simulídeo, Jordan Nunes, pelo apoio, incentivo, disponibilidade, companheirismo, simplicidade, carinho, respeito e amor. O meu muito obrigada à quem sempre quis o meu SUCESSO!
- ✓ As minhas amigas e amigos do “CORACÃO MANAUARA”, Samantha, Nete, Marcinha Borboleta, Artur, Jansen, Véia, Felipe e Cláudia, pelo apoio, amizade e momentos de descontração em Manaus...
- ✓ As minhas amigas e amigo do “CORACÃO MINEIRO”, Helena, Carol, Dani, Luciana e Rafael, pelo apoio, amizade, E COM CERTEZA PELOS MUITOSSSS MOMENTOSSSS DE ZOAÇÃOOOO E BEBBEEDEIIIRRRRAAA!!!! AFF!
- ✓ À minha GRANDE IRMÃ DO CORACÃO.. HELENA.. por tudo que VIVEMOS e por TUDO que AINDA IREMOS VIVER..PQ.. “AMIGAS PARA SEMPRE É O QUE NÓS IREMOS SER”..
- ✓ À minha GRANDE AMIGA “CAROL”.. pela AMIZADE, CARINHO, COMPANHEIRISMO, POR TODAS AS FARRAS, PELO BRAÇO AMIGO QUANDO MAIS PRECISEL.. AMIGOS NÃO SE COMPRA, SE CONQUISTA.. E ESPERO Q NOSSA AMIZADE SEJA ETERNA..
- ✓ À minha GRANDE AMIGA LUCIANA.. por tudo que VIVEMOS, PELO APOIO, HORAS DE DESABAFO, COMPREENSÃO, AMIZADE, CARINHO.. ETC.. ESPERO Q VC NUNCA ESQUECE DESSA AMIGA Q TE GOSTA MUITO.
- ✓ AO MEU AMIGO RAFAEL.. PELA AMIZADE, CARINHO, RESPEITO, COMPREENSÃO.. E PELAS HORAS DE BAR..
- ✓ À minha AMIGA DANI.. que só se eu escrevesse palavras LOUCAS eu conseguiria agradecer.. Muito OBRIGADA POR TUDO...
- ✓ Ao meu amigo Gustavo Martins pelas dicas em microscopia!
- ✓ Ao meu amigo Rodrigo Soares pela amizade, carinho, respeito e exemplo!
- ✓ À minha amiga Nágila, por todos os momentos vividos.. e exemplo de dedicação ao trabalho.
- ✓ Aos meus amigos que passaram ou permanecem no Laboratório de Entomologia Médica do CpRRR, Helena, Carol, Dani, Luciana, Rafael Pimenta, Nágila, Rodriguinho, Brenoviski, Ana Paulo, Carina Margonari, Érika, Roberta, Claudia Velasquez, Vanessa, Carol Cunha, Alessandra, Junara, Ana Anhê, Fernanda, Maíra, Marcele, Bruno, Felipe, Igor, Gustavo Vassorito, Rafael Assis, Isabela, Paula. E a grande Belinha pela amizade e pelo seu excelente cházinho de toda manhã !!!
- ✓ Aos amigos da turma de doutorado 2007 do CpQRR: Helena, Ana Anhê, Ana Cláudia, Ana Carolina, José Eloy, Marco Polo, Fernando, Patrícia.

- ✓ À todos os professores da pós-graduação que derão aula para a turma de 2007.
- ✓ As secretárias da pós-graduação, Cristiane Pinheiro e Andréia Dias, pelo EXCELENTE serviço de atendimento aos alunos.
- ✓ À todos os amigos do Laboratory of Malaria and Vector Research of the National Institute of Health (NIH), José Ribeiro, Eric, John, Ivo, Teresa, Nick, Dongying Ma, Xuenxing e Van pela amizade, apoio, carinho, ensinamentos e descontração.
- ✓ À minha amiga TERESA.. pela amizade, apoio, e ajuda em bioinformática.
- ✓ I'd like to say thank you for my BIG FRIEND, John Andersen.. the only American that I could still understand..QueeÊÊ???? Thank you for your support, teachings, friendship, e pelo braço amigo quando chorei porque não entendia INGLÊS.. Thank You very much!!!
- ✓ Ao meu GRANDE AMIGO Eric Calvo por TUDO que FEZ POR MIM durante o meu doutorado! EU NÃO TENHO NEM PALAVRAS PARA AGRADECER..... SACOVISKI!?
- ✓ Ao amigo “GRANDÃO” Ivo Francischetti ... pelo apoio, amizade, ensinamentos, ajuda em bioquímica.
- ✓ O meu muito obrigada aos **SCIENTIFIC STAFFS** do LMVR/NAID/NIH...**JOHN, ERIC e IVO** pela amizade e por **ALEGRAREM TODOS OS MEUS DIAS DE TRABALHO...**
- ✓ Aos amigos do LMVR: Alexandre, Regis, Clarissa, Fabiano, Jannet, Gisele, Janeki pela amizade e momentos de descontração.
- ✓ **MUITO OBRIGADA PARA MINHA FAMÍLIA BRASILEIRA DE TODOS OS MOMENTOS NA AMÉRICA: ALEXANDRE, GUSTAVO, FERNANDA, ALESSANDRA, LILIANE, ANA PAULA, sem vocês teria sido difícil!.**
- ✓ Thank you... for all my friends of the 2300 Veirs Mill: Ms. Edna, Pradeep and Unus.
- ✓ E a todos que me aguentaram nesses QUATROS ANOS DE DOUTORADO

## **SUMÁRIO**

|   |      |
|---|------|
| <b>LISTA DE FIGURAS</b>                           | xv   |
| <b>LISTA DE TABELAS</b>                           | xix  |
| <b>LISTA DE ABREVIATURAS E SÍMBOLOS</b>           | xx   |
| <b>RESUMO</b>                                     | xxi  |
| <b>ABSTRACT</b>                                   | xxii |
| <b>1 INTRODUÇÃO</b>                               | 23   |
| 1.1 Família Simuliidae: Características gerais    | 23   |
| 1.2 A alimentação sanguínea dos simulídeos        | 24   |
| 1.3 Impactos da picada de simulídeos              | 25   |
| 1.4 Oncocercose – A doença                        | 26   |
| 1.5 A espécie – <i>Thyrsopelma guianense</i>      | 28   |
| 1.6 Glândulas salivares de artrópodes hematófagos | 29   |
| 1.7 Cascata de coagulação sanguínea               | 36   |
| 1.8 Enzimas: fator Xa, trombina e lisozima        | 38   |
| <b>2 JUSTIFICATIVA</b>                            | 39   |
| <b>3 OBJETIVOS</b>                                | 40   |
| 3.1 Objetivo geral                                | 40   |
| 3.2 Objetivos específicos                         | 40   |
| <b>4 MATERIAL E MÉTODOS</b>                       | 41   |
| 4.1 Coleta de simulídeos                          | 41   |

|   |    |
|---|----|
| 4.2 Dissecção de glândula salivar   | 42 |
| 4.3 Microscopia de luz  | 42 |
| 4.4 Microscopia eletrônica de varredura (MEV)   | 43 |
| 4.5 Microscopia eletrônica de transmissão (MET)   | 43 |
| 4.6 Microscopia laser confocal (MLC)  | 43 |
| 4.7 Preparo de homogenado glândula salivar (HGS)  | 44 |
| 4.8 Dosagem de proteína do HGS  | 44 |
| 4.9 Ensaio de tempo de recalcificação   | 44 |
| 4.10 Ensaio de tempo de tromboplastina parcial ativada e tempo de trombina                    | 45 |
| 4.11 Ensaio cromogênico de fXa e trombina   | 45 |
| 4.12 Ensaio quantitativo da atividade da enzima lisozima                                      | 46 |
| 4.13 Transcriptoma e proteoma   | 46 |
| 4.13.1 Construção da biblioteca de cDNA   | 46 |
| 4.13.2 Sequenciamento do cDNA   | 47 |
| 4.13.3 Procedimentos e ferramentas de bioinformática  | 48 |
| 4.13.4 Caracterização proteômica usando gel de eletroforese 1D e espectrometria de massa (MS) | 49 |
| <b>5 RESULTADOS</b>   | 51 |
| 5.1 Análise morfológica da glândula salivar de <i>Thyrsopelma guianense</i>                   | 51 |
| 5.2 Análise funcional dos homogenados salivares de <i>Thyrsopelma guianense</i>               | 53 |
| 5.3 Análise do sialotranscriptoma   | 54 |

|  |    |
|--|----|
| 5.3.1 Enzimas  | 55 |
| 5.3.1.1 Glicosidasas   | 55 |
| 5.3.1.2 Serino Proteases   | 56 |
| 5.3.1.3 Hialuronidasas   | 56 |
| 5.3.1.4 Apirase  | 56 |
| 5.3.1.5 Adenosina deaminase e Destabilase                                | 57 |
| 5.3.2 Domínios inibidores de proteases                                   | 57 |
| 5.3.2.1 Serpinas   | 57 |
| 5.3.2.2 Inibidor de protease com domínio Kunitz                          | 57 |
| 5.3.3 Famílias de proteínas ubíquas                                      | 57 |
| 5.3.3.1 Produtos relacionados a imunidade                                | 57 |
| 5.3.3.2 Família Antígeno-5   | 58 |
| 5.3.3.3 Família Yellow   | 58 |
| 5.3.3.4 Família com domínio ML   | 59 |
| 5.3.3.5 Lipocalinas  | 59 |
| 5.3.4 Famílias de proteínas específicas de insetos                       | 59 |
| 5.3.4.1 Família Aegyptina  | 59 |
| 5.3.4.2 Família de proteína secretada em Diptera e conservada em insetos | 59 |
| 5.3.4.3 Família de proteína laminina                                     | 60 |
| 5.3.4.4 Superfamília D7/OBP  | 60 |
| 5.3.4.4.1 Família D7 Longa   | 60 |



|   |    |
|---|----|
| 5.3.4.4.2 Família D7 16kDa  | 60 |
| 5.3.4.4.3 Família D7 ultra-curtas (10-12kDa)                                    | 61 |
| 5.3.5 Famílias específicas de <i>Simulium</i>                                   | 61 |
| 5.3.5.1 Família SVEP/Marydilan  | 61 |
| 5.3.5.2 Família de proteínas ácidas de <i>Simulium</i> ricas em H               | 62 |
| 5.3.5.3 Mucinas   | 62 |
| 5.3.5.3.1 Mucina de <i>Simulium</i>   | 62 |
| 5.3.5.3.2 Mucinas ácidas similar a família básica de 7-13kDa de <i>Simulium</i> | 62 |
| 5.3.5.4 Família de proteína do tipo-colágeno de <i>Simulium</i>                 | 63 |
| 5.3.5.5 Família Sv 7,8kDa   | 63 |
| 5.3.5.6 <i>Simulium</i> Básica 7-13kDa  | 63 |
| 5.3.5.7 Família <i>Simulium</i> 4,8kDa  | 63 |
| 5.3.5.8 Família <i>Simulium</i> Básica 7,4kDa                                   | 64 |
| 5.3.5.9 <i>Simulium</i> Básica 13kDa  | 64 |
| 5.3.5.10 Família 5-Cys <i>Simulium</i>  | 64 |
| 5.3.5.11 Família 8-10 Cys W   | 64 |
| 5.3.5.12 Família ácida de 28kDa   | 65 |
| 5.3.5.13 Família básica de 28kDa  | 65 |
| 5.3.5.14 Família 19kDa  | 65 |
| 5.3.5.15 Família de proteína básica de 8kDa                                     | 65 |
| 5.3.5.16 Família Sn básica de 4,4kDa  | 66 |

|  |     |
|--|-----|
| 5.3.5.17 Novo peptídeo similar para kunitoxina | 66  |
| 5.4 Trabalho de campo                          | 66  |
| <b>6 DISCUSSÃO</b>                             | 119 |
| <b>7 CONCLUSÃO</b>                             | 134 |
| <b>8 PERSPECTIVAS</b>                          | 135 |
| <b>9 ANEXO I - Tabela Suplementar I e II</b>   | 136 |
| <b>10 REFERÊNCIAS BIBLIOGRÁFICAS</b>           | 157 |

## LISTA DE FIGURAS

|                       |  |    |
|-----------------------|--|----|
| <b>Figura da capa</b> | Glândula salivar de fêmea de <i>Thyrsopelma guianense</i> (Diptera: Simuliidae)  | I  |
| <b>Figura 1</b>       | Fêmea de <i>T. guianense</i> (Diptera: Simuliidae) antes e após o repasto sanguíneo em humanos   | 25 |
| <b>Figura 2</b>       | Características gerais da oncocercose com nódulo aparente, perda da elasticidade da pele e cegueira  | 27 |
| <b>Figura 3</b>       | Esquema dos complexos procoagulantes   | 37 |
| <b>Figura 4</b>       | Coleta e manutenção de simulídeos  | 41 |
| <b>Figura 5</b>       | Dimorfismo sexual em glândula salivar de <i>T. guianense</i> (fêmeas e machos) observadas em microscopia eletrônica de varredura (MEV)                                 | 68 |
| <b>Figura 6</b>       | Glândula salivar de fêmea de <i>T. guianense</i> com dois dias de idade com intensa área pigmentada (pig) em diferentes regiões da glândula                            | 70 |
| <b>Figura 7</b>       | Topografias da glândula salivar de fêmeas de <i>T. guianense</i> em diferentes idades de envelhecimento glandular obtidas em microscópio eletrônico de varredura (MEV) | 72 |
| <b>Figura 8</b>       | Detalhes de fibras nervosas na glândula salivar de fêmeas de <i>T. guianense</i>   | 74 |
| <b>Figura 9</b>       | Aspecto geral da saliva de fêmea de <i>T. guianense</i> em diferentes partes da glândula salivar   | 76 |
| <b>Figura 10</b>      | Marcação da secreção salivar de fêmeas de <i>T. guianense</i> com o uso de lectinas fluorescentes em glândulas com dois dias de idade                                  | 78 |
| <b>Figura 11</b>      | Micrografias da glândula salivar de fêmeas de <i>T. guianense</i> em microscopia laser confocal.   | 80 |

|                  |   |    |
|------------------|---|----|
| <b>Figura 12</b> | Histologia da glândula salivar de fêmeas de <i>T. guianense</i>   | 82 |
| <b>Figura 13</b> | Ultraestrutura da glândula principal de fêmeas de <i>T. guianense</i> . As imagens foram obtidas em microscópio eletrônico de transmissão (MET)                                 | 84 |
| <b>Figura 14</b> | Ultraestrutura da glândula acessória de fêmeas de <i>T. guianense</i> . As imagens foram obtidas em microscópio eletrônico de transmissão (MET)                                 | 86 |
| <b>Figura 15</b> | Atividade anticoagulante de homogenados de glândulas salivares (HGS) de fêmeas de dois dias de idade de <i>T. guianense</i>   | 88 |
| <b>Figura 16</b> | Inibição da atividade catalítica de 0.5 nM de fXa por homogenados de glândula salivar de <i>T. guianense</i>  | 90 |
| <b>Figura 17</b> | Ensaio de atividade catalítica de 0.5 nM de fXa e 0.1 nM de trombina por homogenados de glândula salivar de <i>T. guianense</i>   | 90 |
| <b>Figura 18</b> | Atividade de lisozima em homogenados de glândulas salivares (HGS) de fêmeas de <i>T. guianense</i>  | 91 |
| <b>Figura 19</b> | Classificação das classes funcionais das etiquetas de sequências expressas (EST) ou grupos assemblados (GRUPO) derivados da biblioteca de cDNA de fêmeas de <i>T. guianense</i> | 92 |
| <b>Figura 20</b> | Gel de eletroforese 1D de homogenados de glândulas salivares de <i>T. guianense</i>   | 94 |
| <b>Figura 21</b> | Família de proteína amilase/maltase de <i>T. guianense</i>  | 96 |
| <b>Figura 22</b> | Família serino protease de <i>T. guianense</i>  | 97 |
| <b>Figura 23</b> | Família hialuronidase de <i>T. guianense</i>  | 98 |
| <b>Figura 24</b> | Família apirase de <i>T. guianense</i>  | 99 |

|                  |   |     |
|------------------|---|-----|
| <b>Figura 25</b> | Família Kunitz de <i>T. guianense</i>   | 100 |
| <b>Figura 26</b> | Família lisozima de <i>T. guianense</i>   | 101 |
| <b>Figura 27</b> | Família cecropina de <i>T. guianense</i>  | 102 |
| <b>Figura 28</b> | Família antígeno-5 de <i>T. guianense</i>   | 103 |
| <b>Figura 29</b> | Família aegyptina de <i>T. guianense</i>  | 104 |
| <b>Figura 30</b> | Família de proteína secretada em Diptera e conservadas em insetos de <i>T. guianense</i>  | 105 |
| <b>Figura 31</b> | Família D7 longa de <i>T. guianense</i>   | 106 |
| <b>Figura 32</b> | Família D7 16kDa de <i>T. guianense</i>   | 107 |
| <b>Figura 33</b> | Família D7 super curtas de <i>T. guianense</i>  | 108 |
| <b>Figura 34</b> | A proteína eritrema de <i>Simulium vittatum</i> (SVEP) superfamília de <i>Simulium</i>  | 109 |
| <b>Figura 35</b> | Proteína de <i>T. guianense</i> ricas em His  | 110 |
| <b>Figura 36</b> | Família mucina de <i>Simulium</i> de <i>T. guianense</i>  | 111 |
| <b>Figura 37</b> | Análise filogenética de proteínas mucinas ácidas similares a família básica 7-13 <i>Simulium</i> derivadas do sialotranscriptoma de <i>T. guianense</i> | 112 |
| <b>Figura 38</b> | Análise filogenética da família similar a colágeno de <i>Simulium</i> derivadas do sialotranscriptoma de <i>T. guianense</i>                            | 113 |
| <b>Figura 39</b> | Família de proteína Sv 7,8kDa de <i>T. guianense</i>  | 114 |
| <b>Figura 40</b> | Alinhamento e dendograma da proteína básica 7-13kDa de <i>T. guianense</i>  | 115 |

|                  |  |     |
|------------------|--|-----|
| <b>Figura 41</b> | Alinhamento e dendograma da proteína 4,8kDa de <i>T. guianense</i> | 115 |
| <b>Figura 42</b> | Alinhamento e dendograma da proteína 5-Cys de <i>T. guianense</i>  | 116 |
| <b>Figura 43</b> | Família Sn 8-10 Cys W de <i>T. guianense</i>                       | 116 |
| <b>Figura 44</b> | Família proteína ácida 28kDa de <i>T. guianense</i>                | 117 |
| <b>Figura 45</b> | Família de proteína básica 28kDa de <i>T. guianense</i>            | 117 |
| <b>Figura 46</b> | Família de proteína 19kDa de <i>T. guianense</i>                   | 118 |

## LISTA DE TABELAS

|                 |   |    |
|-----------------|---|----|
| <b>Tabela 1</b> | Lectinas fluorescentes-FITC utilizadas para determinação de resíduos de carboidratos de superfície da glândula salivar de simúlideos  | 44 |
| <b>Tabela 2</b> | Quantificação de proteína de homogenados de glândulas salivares (HGS) de fêmeas de <i>T. guianense</i> de dois dias de idade <i>T. guianense</i>  | 88 |
| <b>Tabela 3</b> | Ensaio de tempo de tromboplastina parcial ativada (TTPa) na ausência (controle) e na presença de diferentes concentrações de homogenados de glândulas salivares (HGS) de fêmeas de dois de idade de <i>T. guianense</i> | 89 |
| <b>Tabela 4</b> | Ensaio de tempo de tromboplastina (TP) na ausência (controle) e na presença de diferentes concentrações de homogenados de glândulas salivares (HGS) de fêmeas de dois dias de idade de <i>T. guianense</i>              | 89 |
| <b>Tabela 5</b> | Classificação funcional de transcritos salivares originados da glândula salivar de <i>T. guianense</i>  | 92 |
| <b>Tabela 6</b> | Classificação funcional dos transcritos secretados originados das glândulas salivares de <i>T. guianense</i>  | 93 |
| <b>Tabela 7</b> | Classificação funcional de transcritos celulares originados das glândulas salivares de <i>T. guianense</i>  | 93 |
| <b>Tabela 8</b> | Proteínas secretadas deduzidas da análise do sialotranscriptoma e identificação de expressão por análise proteômica   | 95 |

## LISTA DE ABREVIATURAS E SÍMBOLOS

|            |                                |
|------------|--------------------------------|
| <b>gl</b>  | Glândula                       |
| <b>gp</b>  | Glândula principal             |
| <b>ga</b>  | Glândula acessória             |
| <b>d</b>   | Ducto                          |
| <b>dc</b>  | Ducto comum                    |
| <b>c</b>   | Constricção                    |
| <b>ve</b>  | Vasícula                       |
| <b>rp</b>  | Região proximal                |
| <b>rd</b>  | Região distal                  |
| <b>n</b>   | Núcleo                         |
| <b>lu</b>  | Lúmen                          |
| <b>HGS</b> | Homogenado de glândula salivar |
| <b>PBS</b> | Tampão fosfato salino          |
| <b>BSA</b> | Albumina bovina                |
| <b>A</b>   | Alanina                        |
| <b>C</b>   | Cisteína                       |
| <b>F</b>   | Fenilalanina                   |
| <b>G</b>   | Glicina                        |
| <b>H</b>   | Histidina                      |
| <b>I</b>   | Isoleucina                     |
| <b>L</b>   | Leucina                        |
| <b>M</b>   | Metionina                      |
| <b>P</b>   | Prolina                        |
| <b>S</b>   | Serina                         |
| <b>T</b>   | Treonina                       |
| <b>V</b>   | Valina                         |
| <b>W</b>   | Triptofano                     |
| <b>Y</b>   | Tirosina                       |
| <b>R</b>   | Arginina                       |
| <b>Q</b>   | Glutamina                      |
| <b>N</b>   | Asparagina                     |
| <b>K</b>   | Lisina                         |
| <b>E</b>   | Ácido glutâmico                |
| <b>D</b>   | Ácido aspártico                |



## RESUMO

O simulídeo *Thrysopelma guianense* é o principal vetor da oncocercose no Brasil, e pouco se conhece sobre a composição e função da saliva em simulídeos. Artrópodes hematófagos contém uma “poção mágica” para interagir com a hemostase, inflamação e imunidade de seus hospedeiros. A glândula salivar de fêmeas de *T. guianense* é constituída de dois lobos idênticos contendo uma glândula principal (envolvida com a alimentação com sangue) e uma acessória (envolvida com a alimentação com açúcar), a última similar à glândula de macho. A análise histoquímica demonstrou uma saliva é rica em polissacarídeos, proteínas e lipídios principalmente na glândula principal. Essa região contém cavidades secretoras elétron-densas, enquanto, o lúmen da glândula acessória é elétron-lucente. Os homogenados salivares inibiram a coagulação sanguínea, atuando em fXa mas não em trombina. Esses homogenados também apresentaram atividade enzimática de lisozima em pH 7,0. O sialotranscriptoma isolou e identificou 1.772 transcritos, classificados como secretores (S) (74,7%), celulares (C) (16,2%), produtos de função desconhecida (U) (9%) e elementos de transposição (0,1%). Esse sialotranscriptoma identificou famílias de proteínas ubíquas como antígeno-5, Yellow, domínios ML, lipocalina, lisozima, cecropina, serpina, domínios Kunitz, serino proteases, hialuronidase, apirase, glicosidases, adenosina deaminase e destabilase, além de famílias de proteínas específicas para insetos como Aegyptina e a superfamília D7/OBP. Cerca de 63,4% dos produtos secretores corresponderam a famílias de proteínas encontradas somente em *Simulium*, como SVEP/Marydilan, proteínas ácidas ricas em His, Sv 7,8 kDa, mucinas, tipo-colágeno, básicas 7-13 kDa, Sv 7,8 kDa, 4,8 kDa, básicas 7,4 kDa, básica 13 kDa e família 5-Cys. Várias famílias de proteínas foram deorfanzadas do sialotranscriptoma de *S. nigrimanum* como Sn 8-10 Cys W, ácida 28 kDa, básica 28 kD, 19 kDa, básica 8 kDa e Sn básica 4,4 kDa. Transcritos com similaridade para kunitoxina foram encontrados em *T. guianense*. Observamos um aumento na expressão de transcritos para a classe S e expecionalmente em SVEP, e nos subgrupos síntese de proteína e metabolismo energético da classe C, em relação a outros sialotranscriptomas de *Simulium*. O proteoma confirmou 28 das 32 famílias de proteínas encontradas no transcriptoma e disponibilizamos 101 sequências de proteínas de interesse para o NCBI. Essa é a primeira descrição de um sialoma (do grego Greek Sialo=saliva) de um vetor de oncocercose. Assim, nossos resultados ajudam no entendimento do papel da saliva de *Simulium* na transmissão de *Onchocerca volvulus* e para a evolução de proteínas salivares em simulídeos, além de consistir numa plataforma de novos compostos antihemostáticos e candidatos a vacina contra oncocercose.

## ABSTRACT

The black fly *Thyrsopelema guianense* is the main vector of River Blindness disease in Brazil, and little is known about the composition and function of the saliva in black flies. Hematophagous arthropods contain a “magic potion” that counteracts their host’s hemostasis, inflammation and immunity. The salivary glands is formed by two identical lobes contained main (blood-feeding) and accessory (sugar-feeding) glands, which has similarity to male glands. The saliva is rich in polysaccharides, lipids and proteins mainly present in the secretory cavities of the main gland. These secretory cavities are highly electron-dense, however, the lumen of the accessory gland is electron-lucent. Salivary gland homogenates inhibit the blood coagulation by targeting fXa and not thrombin. The salivary glands secrete lysozyme as a immunity related product, with optimal activity at pH 7.0. We isolated and identified 1,772 transcripts, classified as secreted (S) (74.7%), housekeeping (H) (16.2%), unknown products (U) (9%) and transposable elements (0.1%). Ubiquitous salivary protein families, such as the Antigen-5, Yellow, ML domain, lipocalins, lysozyme, cecropin, serpin, Kunitz domain, serine proteases, hyaluronidase, apyrase, glycosidase, adenosine deaminase and destabilase were reported in the *T. guianense* sialotranscriptome. Also found were insect-specific families such as the Aegyptin and D7/OBP superfamily. About 63.4% of all secreted products revealed protein families only found in *Simulium*, such as SVEP/Marydilan, H-rich acid proteins of *Simulium*, Sv 7.8 kDa, Mucin, collagen-like, basic 7-13 kDa, Sv 7.8 kDa, 4.8 kDa, basic 7.4 kDa, basic 13 kDa, 5-Cys family. Several protein families from *S. nigrimanum* sialotranscriptome were deorphanized such as the Sn 8-10Cys W family, acid 28 kDa, basic 28 kD, 19 kDa, Basic 8 kDa and Sn basic 4.4 kDa. Additionally, we found a novel peptide similar to kunitoxin, found in insects sialotranscriptomes for the first time. Transcripts of the S class, mainly of the SVEP family, and of the H class involved in protein synthesis and energy metabolism were significantly increased when compared with other *Simulium* sialotranscriptomes. We confirmed by proteome analysis 28 of the 32 protein families found in the transcriptome, and offer 101 sequences to NCBI. This is the first description of a sialome (from the Greek Sialo=saliva) of a vector associated with River blindness disease. Thus, our results contribute to understand the role of the *Simulium* saliva in the transmission of *Onchocerca volvulus* and to the evolution of the salivary proteins in black flies. It also consists of a platform for mining novel anti-hemostatic compounds and vaccine candidates against filariasis.

# 1 INTRODUÇÃO

## 1.1 Família Simuliidae: Características gerais

Os simulídeos (Figura 1A) pertencem à ordem Diptera, subordem Nematocera, Infraordem Culicomorpha, Superfamília Chironomoidea, Família Simuliidae. São conhecidos como “borrachudo” ou “pium” no Brasil e como “black flies” em países de idioma inglês. São insetos cosmopolitas, exceto para o continente antártico, alguns arquipélagos e locais desprovidos de água corrente. Atualmente, a família Simuliidae representa uma das famílias de Diptera mais estudadas com 2.072 espécies descritas, sendo 12 fósseis (Adler & Crosskey, 2008).

Os simulídeos são insetos com metamorfose completa: ovo, larva, pupa e adulto. O ciclo de vida compreende a dois estágios: um aquático (ovo-larva-pupa) e um aéreo (adultos: fêmea e macho). A fêmea ovipõe na superfície de águas correntes, em lotes que variam de 100 a 600 ovos com a eclosão da larva depois de vários dias. Ela pode passar por seis a nove mudas dependendo da espécie e das características do criadouro, sendo até sete mudas o mais frequente. A larva constrói um casulo feito de seda e passa para o estágio pupal, que dura de quatro a sete dias, não ocorrendo alimentação (Crosskey, 1990).

Os machos e fêmeas adultos se alimentam de néctar tendo a fêmea uma alimentação sanguínea adicional para a maturação dos ovos. Desse modo, as fêmeas são conhecidas como pragas picadoras e importantes transmissoras de parasitos sanguíneos e de pele para humanos e para animais de sangue quente (Crosskey, 1990).

Os simulídeos são também importantes indicadores de contaminação em ambientes de água doce devido aos estágios imaturos serem altamente susceptíveis à poluição orgânica e inorgânica. Além disso, os estágios imaturos de simulídeos são considerados organismos chaves em ecossistemas aquáticos, sendo fontes importantes de alimentação para invertebrados e vertebrados (Crosskey, 1990).

No mundo cerca de 10% das espécies de simulídeos trazem impactos econômicos para a agricultura, silvicultura, pecuária e turismo, acarretando perdas na produção de leite e carne de bovinos e redução no turismo (Crosskey, 1990). Na América do Norte, algumas espécies proliferam como resultado do desenvolvimento rural, como por exemplo, a construção de usinas hidrelétricas (Shelley *et al.*, 2010).

Os simulídeos são vetores de várias espécies de parasitos, desde vírus e protozoários até filárias. Esses parasitos podem ser transmitidos para humanos e animais domésticos (Crosskey, 1990; Adler, 2005). Entre os protozoários podemos citar espécies do gênero *Leucocytozoon* que

assolam populações de aves (Szymanski & Lovette, 2005) e as filárias *O. volvulus*, *Mansonella ozzardi* (Shelley *et al.*, 1997; Py-Daniel, 2000) que atingem os humanos. Os simulídeos também são suspeitos de serem vetores da Encefalite Equina Venezuelana (Homan *et al.*, 1985).

No Brasil, os simulídeos também foram associados a uma doença endêmica conhecida como Fogo Selvagem (FS) ou Pênfigo Foliáceo (PF) endêmico. É uma doença autoimune, bolhosa, crônica da pele e que se caracteriza histopatologicamente pela formação de bolhas intraepidérmicas com acantólise e, imunologicamente, por autoanticorpos circulantes na epiderme, responsáveis pelas lesões cutâneas e que causa no paciente forte sensação de calor e queimadura (Zaitz *et al.*, 2000). Essa doença é endêmica na América Latina (Argentina, Bolívia, Colômbia, El Salvador, Paraguai, Peru e Venezuela) e em algumas partes do Brasil (Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, São Paulo) tendo sido atribuída à saliva de insetos, inclusive de simulídeos (espécie *Simulium nigrimanum*) além de outros fatores ambientais (Diaz *et al.*, 1989).

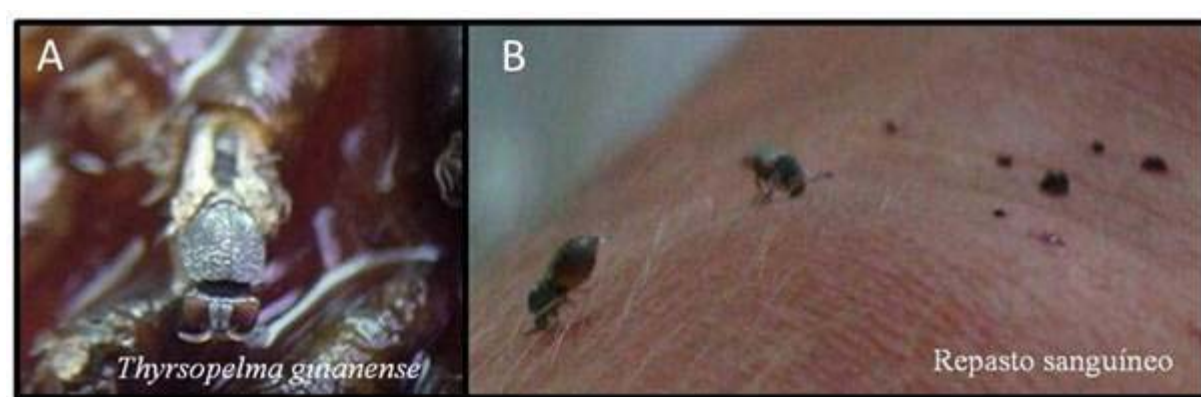
Na região norte do Brasil os simulídeos foram relatados por causar reações autoimunes em migrantes e não em nativos que habitavam a rodovia Transamazônica (Pinheiro *et al.*, 1974). Essa enfermidade ficou conhecida como Síndrome Hemorrágica de Altamira com casos de púrpura trombocitopênica de forma secundária. A ação direta da toxina provocava hemorragias cutâneas, localizadas ou disseminadas (forma de múltiplas petéquias e equimoses) ao redor do local das picadas. Muitos pacientes apresentaram casos de sangramento em mucosas, anemia e mortes. A ocorrência dessa síndrome foi atribuída à saliva de *Simulium amazonicum* (Pinheiro *et al.*, 1974).

Na região amazônica, os simulídeos são importantes na transmissão de *O. volvulus* e *M. ozzardi* para humanos (Py-Daniel, 1997; 2000; Medeiros *et al.*, 2007, 2009), além de causarem grande incômodo para a população. Em algumas regiões da Amazônia têm sido registrado mais de 6.780 picadas/homem/dia (Shelley *et al.*, 1997).

## **1.2 A alimentação sanguínea dos simulídeos**

As fêmeas de simulídeos realizam uma alimentação sanguínea, adicional à de açúcares, para a maturação dos óvulos. Entre os culicomorfas, os simulídeos são conhecidos por dilacerar o tecido do hospedeiro na busca pelo sangue, sendo denominados “pool feeders” (Figura 1B) diferentemente dos mosquitos que possuem um aparelho bucal canular podendo se alimentar diretamente de um vaso sanguíneo e por isso são denominados “vessel feeders”.

O modo de alimentação “pool feeders” para simulídeos é devido ao seu curto aparelho bucal. A dilaceração do tecido do hospedeiro é realizada por uma estrutura denominada sintrófia, constituída de mandíbulas, lacínias maxilares, hipofaringe e labro. As demais peças bucais como lábio e lobos labelares não penetram na pele e são retraídas permitindo a picada (Sutcliffe & McIver, 1987). O modo de alimentação ocorre em quatro momentos: 1) penetração das mandíbulas; 2) consolidação do sintrófia na pele; 3) funcionamento de duas bombas funcionais (uma no cibário e outra na faringe), ambas separadas do canal alimentar; e 4) liberação das peças bucais e do sintrófia (Sutcliffe & McIver, 1987). O sucesso na obtenção da alimentação sanguínea em hematófagos, incluindo simulídeos, tem sido atribuída também à produção de uma “poção mágica” na saliva desses insetos (Ribeiro, 1995).



**Figura 1** – Fêmea de *T. guianense* (Diptera: Simuliidae) antes e após o repasto sanguíneo em humanos. Fotos: Chagas, AC.

### 1.3 Impactos da picada de simulídeos

Em simulídeos somente a fêmea é hematófaga e são comumente conhecidos pela agressividade de suas picadas. Os impactos causados pela picada podem ser agrupados em três categorias: (1) incômodo, causado pelas picadas em enxame; (2) trauma, causado pela injeção de toxinas salivares; e (3) doença, causada pela transmissão de parasitos. Cada impacto pode ser ocasionado por espécies particulares afetando humanos, animais domésticos e selvagens (Adler, 2005). Diferentes espécies apresentam o hábito persistente de voar para os olhos, os ouvidos, o nariz, a boca, entre os cabelos e na pele causando pânico e incômodo proveniente de suas picadas (Mulla & Lacey, 1976), além de ocasionar vários sintomas que incluem febre, inchaço, cefaléia e dores nas articulações (Service, 1997).

Em pessoas e animais não ambientados, a picada de simulídeos pode ocasionar fortes reações devido ao efeito tóxico da injeção das secreções salivares, conhecida como “simuliotoxicose” (intoxicação por simulídeo) (Adler, 2005). O grande número de picadas

também pode resultar em mortes nos mais diversos animais (bovinos, equinos, suínos, ovinos e aves) em função da perda excessiva de sangue, do comportamento descontrolado do animal (resultando em atropelamento, pisadas e esmagamento entre eles), além de sufocamento devido ao entupimento das vias respiratórias (Adler, 2005). Algumas espécies de simulídeos acarretam prejuízo na economia de diferentes países como pragas de búfalos nos Estados Unidos e Canadá (*Cnephia pecuarum* e *Prosimulium mixtum*, respectivamente).

No sudeste e sul do Brasil, a espécie *Chirostilbia pertinax* é a principal espécie-alvo em programas de controle de simulídeos. A picada dessa espécie causa prurido intenso e irritação, podendo acarretar reações imunes graves. Vários programas de controle têm sido implantados no intuito de diminuir a população desses insetos e minimizar o sofrimento das populações rurais que trabalham em atividades tradicionais, como a coleta de uva e agropecuária (Gil-Azevedo *et al.*, 2004). Somente na região norte do Brasil os simulídeos apresentam importância epidemiológica na transmissão de filárias (Py-Daniel & Medeiros, 2000).

#### **1.4 Oncocercose – A doença**

A oncocercose ou “cegueira dos rios” é uma doença causada pela filária *Onchocerca volvulus*, endêmica em 28 países da África Tropical (onde ocorre a maioria dos casos) e em focos dispersos nas Américas Central e do Sul (particularmente na Guatemala, Equador, México, Colômbia, Brasil e Venezuela), na Península Arábica (Arábia Saudita), e no Yêmen (Lane & Crosskey, 1993).

Estima-se que 18 milhões de pessoas estejam infectadas por oncocercose no mundo, com aproximadamente 40 mil casos novos/ano. A doença é a quarta causa de cegueira em países em desenvolvimento. Nas Américas, cerca de 1.8% da população precisa de tratamento entre o Brasil e a Venezuela (WHO, 2003).

No hospedeiro definitivo (humanos), os vermes adultos (macho e fêmea) vivem em nódulos aparentes e palpáveis (Figura 2). As fêmeas de *O. volvulus* vivem cerca de 14 anos e produzem microfilárias que migram ativamente no tecido subcutâneo, podendo migrar para os olhos. Quando presentes na pele humana, elas podem ser ingeridas por fêmeas de simulídeos durante o repasto sanguíneo. No vetor, essas filárias passam por duas mudas ( $L_1$  e  $L_2$ ) e transformam-se em larva infectante  $L_3$ , que é liberada num próximo repasto sanguíneo. No hospedeiro definitivo, a filária passa por duas mudas até se tornar verme adulto (WHO, 2003).

Em áreas endêmicas de oncocercose, a oncodermatite é a principal causa de morbidade, com severas desfigurações dérmicas que levam à perda da elasticidade da pele durante a fase

crônica da doença (Figura 2). Nessa fase, notam-se complicações oftalmológicas que variam desde uma redução da visão periférica até uma cegueira noturna ou uma completa perda visual quando a microfilária migra para o nervo óptico (Figura 2) (WHO, 2003). Na oncocercose, a principal resposta inflamatória na córnea é causada pela presença de bactérias do gênero *Wolbachia* que são simbiossomas nas filárias. Atualmente, tratamentos com antibióticos estão ajudando a diminuir a perda visual das pessoas oncocercóticas (Saint Andre *et al.*, 2002).



**Figura 2** – Características gerais da oncocercose com nódulo aparente, perda da elasticidade da pele e cegueira.

No Brasil o primeiro caso de oncocercose foi relatado em uma criança de três anos de idade, filha de missionários estrangeiros procedentes do estado de Roraima (Bearzoti *et al.*, 1967). Posteriormente a forma endêmica da oncocercose foi encontrada na área Yanomami e em

famílias de missionários que habitavam essa região (Moraes & Dias, 1972; Moraes & Chaves, 1974).

Atualmente, a hipótese mais aceita para a presença de oncocercose nas Américas é que ela tenha sido introduzida através do tráfico de escravos (Sá & Maia-Herzog, 2003). Entretanto, a existência dessa doença entre os Yanomamis ainda é muito discutida, principalmente devido ao isolamento desse grupo. Acredita-se que a doença tenha seguido um caminho indireto através de focos existentes na Venezuela, visto que os indígenas nunca tiveram contato com os escravos e nem com seus descendentes (Moraes, 1991).

Das 12 espécies de simulídeos presentes na área, somente quatro foram envolvidas com a transmissão de *O. volvulus*: *T. guianense* (Wise, 1911), *Psaroniocompsa incrustata* (Lutz, 1910), *Cerqueirellum oyapockense* (Floch & Abonnenc, 1946) e *Notolepria exigua* (Roubaud, 1906), sendo a primeira de maior importância epidemiológica (Py-Daniel, 2000; Shelley *et al.*, 1997).

Após vários anos de acompanhamento do foco amazônico foi relatado a possível presença de um novo foco na região central do Brasil (Maia-Herzog *et al.*, 1999), entretanto, sua fixação e expansão não foi confirmada. Atualmente, não existem programas de controle vetorial para o foco oncocercótico brasileiro e o tratamento de pacientes infectados dá-se através da administração de ivermectina a cada 6 meses.

### 1.5 A espécie - *Thyrsopelma guianense*

*Thyrsopelma guianense* [= *Simulium guianense*] foi descrita por Wise (1911) baseado em onze fêmeas que picaram humanos no rio Essequibo na Guiana. Atualmente existem duas escolas de sistemática para a família Simuliidae: a primeira categoriza todas as espécies dentro de um único gênero - *Simulium* e aceita somente subgêneros, enquanto a segunda escola aceita gêneros adicionais para *Simulium*. Seguidores desta segunda linha taxonômica elevaram alguns subgêneros a nível genérico, como no caso de *Thyrsopelma* (Py-Daniel & Moreira-Sampaio, 1994).

A espécie é amplamente distribuída no Brasil nos estados do Amapá, Amazonas, Goiás, Maranhão, Mato Grosso, Minas Gerais, Pará, Pernambuco, Rio de Janeiro, Roraima, São Paulo e Tocantins, e em países vizinhos como Guiana Francesa, Guiana, Suriname e Venezuela (Adler & Crosskey, 2008). *T. guianense* coloniza terrenos de características geológicas muito antigas (Pré-cambriano: Escudo das Guianas, Escudo do Brasil Central) não ocorrendo em áreas do Terciário ou Quaternário da Amazônia. A fase imatura desse inseto possui forte relação com macrófitas aquáticas da família Podostemaceae (Py-Daniel, 1994).



Essa espécie é altamente antropofílica (Figura 1) e conhecida como vetor primário da oncocercose em áreas montanhosas do foco amazônico (Brasil e Venezuela) (Basanez *et al.*, 1988; Py-Daniel, 1997). Em áreas altas do lado venezuelana do foco essa espécie é considerada antropofílica, enquanto nas áreas baixas é considerada zoofílica (Shelley *et al.*, 1997; Hamada & Fouque, 2001).

*T. guianense* também tem sido suspeita ser um complexo de espécies crípticas devido aos comportamentos de picadas diferenciados. Na maioria das localidades as fêmeas apresentam atividade de picada bimodal sendo mais acentuada nos períodos crepusculares (manhã/tarde) (Lacey & Charlwood, 1980; Lacey, 1981) e com nítida preferência pela região inferior do corpo tanto no Brasil quanto na Venezuela (Takaoka *et al.*, 1985; Basañez, 1988; Py-Daniel, 1997). No rio Tapajós, estado do Pará/Brasil, o grande número de fêmeas picando acarreta a interrupção das atividades de rotina na área (Lacey & Charlwood, 1980; Lacey, 1981).

Uma das características bastante utilizadas em simulídeos para indicar a competência vetorial é a presença de dentes cibariais. Esses dentes promovem uma maior possibilidade de rompimento das microfilárias durante a hematofagia agregando um fator de dificuldade na continuidade do ciclo de vida do parasito (Py-Daniel *et al.*, 2005).

*T. guianense* foi primeiramente sugerida com ausência de dentes cibariais (Shelley *et al.*, 1997), entretanto, trabalhos posteriores mostraram que a espécie possui um grande número de dentes com o ápice rombo e normalmente mais largos que longos (também considerado um fator de impedimento) (Py-Daniel *et al.*, 2005). Devido a importância epidemiológica dessa espécie mais estudos sobre a biologia desse vetor deveriam ser realizados no intuito de entender a competência vetorial de *T. guianense* na transmissão de *O. volvulus*.

## **1.6 Glândulas salivares de artrópodes hematófagos**

Por muitos anos acreditou-se que a saliva de artrópodes hematófagos fosse somente útil para a lubrificação das partes bucais. Atualmente o estudo das glândulas salivares desses artrópodes tem contribuído para um melhor entendimento do papel da saliva frente a uma alimentação sanguínea e/ou rica em açúcares (Ribeiro, 1995). As glândulas salivares de artrópodes hematófagos representam um dos órgãos mais estudados devido vários fatores, entre eles: a) local onde se encontram as formas infectantes dos principais parasitas e vírus responsáveis por importantes doenças parasitárias; b: esse órgão é responsável pela secreção da saliva que contém várias substâncias farmacologicamente ativas que desarmam as respostas hemostáticas do hospedeiro e garante o êxito da alimentação sanguínea e da transmissão de

patógenos (Ribeiro, 1995); c) a glândula salivar é um órgão potencial para busca de novas drogas de interesse farmacológico.

Em geral a glândula salivar de famílias distintas de artrópodes hematófagos varia quanto aos padrões morfológicos, desde formas lobulares, saculares e lobulares-saculares (Barth, 1954; Orr *et al.*, 1961; Gosbee *et al.*, 1969; Wright, 1969; James & Rossignol, 1991; Serrao *et al.*, 2008), podendo secretar diferentes produtos salivares. Como por exemplo, diferentes gêneros de mosquito podem apresentar diferenças morfológicas aparentes na glândula salivar (Orr *et al.*, 1961; Wright, 1969; Barrow *et al.*, 1975; Moreira-Ferro *et al.*, 1999; Sais *et al.*, 2003) que reflete na secreção de distintos compostos salivares que atuam na alimentação de açúcar e/ou sangue (Orr *et al.*, 1961; Wright, 1969; Marinotti & James, 1990). Em simulídeos e ceratopogonídeos, a presença da forma sacular em uma glândula lobular tem sido proposta como um reservatório salivar com diferença de produtos secretados (Wächtler *et al.*, 1971; Perez de Leon *et al.*, 1994).

Em *Cimex hemipterus*, a glândula salivar é formada por duas glândulas (uma principal e outra acessória) sendo similar a forma encontrada em outros heterópteros, entretanto, ela é diferente da forma pêra vista nas glândulas de *Cimex lectularius* (Baptist, 1941; Forattini, 1990). Essas diferenças podem ser reflexos de suas funções. A glândula salivar de *C. lectularius* apresentam anticoagulantes, vasodilatadores e enzimas (Valenzuela *et al.*, 1995; 1998; Ribeiro & Francischetti, 2003). Em Triatominae os lobos da glândula principal estão envolvidos na síntese de anticoagulantes e substâncias hemolíticas (Barth, 1954; Lacombe, 1999).

A glândula salivar também possui um importante papel para a manutenção do ciclo de parasitos, e seu estudo morfológico tem ajudado a reconhecer sítios de invasão desses parasitas nos modelos mosquito/*Plasmodium* e moscas tsé-tsé/*Trypanosoma brucei* (Perrone *et al.*, 1986).

Vários aspectos fisiológicos sobre a glândula salivar desses hematófagos permanecem obscuros, como por exemplo, o estímulo necessário para salivação. Em alguns grupos acredita-se que a presença de um grande número de fibras nervosas seja responsável por esse estímulo. Em *Aedes aegypti*, a serotonina tem sido proposta como um neurotransmissor para o estímulo da salivação (Novak *et al.*, 1995).

Acredita-se que pelo menos 20 famílias de artrópodes hematófagos tiveram uma evolução independente quanto ao hábito da hematofagia e isso aconteceu aproximadamente 30 vezes dentro desse grupo (Ribeiro, 1995; Grimaldi & Engel, 2005). Esse fato poderia explicar a grande diversidade de compostos anti-hemostáticos, em adição à rápida evolução das várias famílias de proteínas salivares possivelmente como resultado da pressão imune de seus hospedeiros. Desse modo, postula-se que todos os gêneros de artrópodes hematófagos que divergiram antes da

irradiação dos mamíferos possam apresentar consideráveis variações na suas secreções salivares (Ribeiro & Francischetti, 2003; Ribeiro & Arca, 2009).

Os hematófagos tem sido batizados como verdadeiros “farmacologistas” toda vez que eles precisam desarmar as respostas hemostáticas do hospedeiro, que são ativadas para prevenir a perda sanguínea após uma injúria vascular. Hoje sabemos que eles desenvolveram uma saliva contendo uma “poção mágica”, composta de poderosas substâncias bioativas capazes de superar todas as barreiras impostas pelo hospedeiro durante a hematofagia (Ribeiro, 1995). Esses hematófagos elaboraram pelo menos um anticoagulante, um antiplaquetário e um vasodilator, e em muitos casos, mais de uma molécula para cada categoria. A presença desse tripé anti-hemostático é suficiente para o sucesso da laborosa tarefa desses hematófagos na remoção de sangue de um sistema fisiológico tão equilibrado como é o dos vertebrados (Ribeiro, 1987; Law *et al.*, 1992; Ribeiro, 1995; Ribeiro & Francischetti, 2003).

Durante a injúria vascular a resposta das plaquetas é extremamente rápida, sendo relevante no processo da alimentação sanguínea. Desse modo, esses hematófagos elaboraram diferentes substâncias farmacologicamente ativas para inibir a agregação plaquetária. Vários desses inibidores plaquetários já foram molecularmente caracterizados e seu modo de ação revelado (Francischetti, 2010).

Na maioria desses hematófagos tem sido observada a presença de apirase (ATP-difosfohidrolases). Essa enzima possui a função de prevenir a agregação plaquetária por meio da conversão de ATP e ADP para AMP, eliminando assim o sinal (ADP) para a agregação (Ribeiro *et al.*, 1991). Em mosquitos, a síntese dessa enzima ocorre no lobo médio e na região distal dos lobos laterais da glândula (Ribeiro *et al.*, 1984) e tem sido proposto que a quantidade dessa enzima na saliva é inversamente proporcional ao tempo que uma fêmea necessita para picar e localizar o vaso sanguíneo. Assim, a presença de grandes quantidades dessa enzima nas glândulas salivares proporcionaria uma vantagem adaptativa aos mosquitos (Ribeiro, 2000).

A primeira apirase salivar foi caracterizada em *Aedes aegypti* como um homólogo da família 5'-nucleotidase de vertebrado (Champagne *et al.*, 1995a,b). As 5'-nucleotidasas são enzimas ubíquas e comumente encontradas ligadas à face extracelular de membranas fisiológicas por uma âncora de glicoinositol fosfato (GPI) (Misumi *et al.*, 1990, Ogata *et al.*, 1990). Em dípteros, as apirases salivares de mosquitos (Champagne *et al.*, 1995a,b), de triatomíneos (Faudry *et al.*, 2004) e de simulídeos (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010a) não apresentam o domínio carboxiterminal onde o glicolípídeo é ancorado, permitindo sua secreção. Em outros artrópodes hematófagos como os cimiídeos e flebotomíneos, a apirase é do

tipo *Cimex* (Valenzuela *et al.*, 1998, 2001) e em pulgas apresentam o tipo proteína CD-39 (Andersen *et al.*, 2007).

Os anticoagulantes tem sido relatados em glândulas salivares de hematófagos desde o início do século XX, entretanto, somente nas últimas décadas teve início a caracterização dessas moléculas assim como seu uso em ensaios clínicos (Ribeiro, 1987; 1995; Law *et al.*, 1992; Ribeiro & Francischetti, 2003). Os anticoagulantes estão entre os antihemostáticos mais estudados e sua origem polifilética sugere um processo de evolução convergente para hematofagia (Stark & James, 1996).

Entre os anticoagulantes mais estudados encontra-se a “Hirudina” identificada a mais de 50 anos na glândula salivar da sanguessuga *Hirudo medicinalis* (Markwardt, 1994). O sucesso da hirudina recombinante mostrou a utilidade desses produtos naturais no desenvolvimento de drogas (Koh & Kini, 2009), permitindo que outros inibidores de trombina fossem caracterizados como bivalirudina (Warkentin *et al.*, 2008), argatroban (Yeh & Jang, 2006), dabigatran (Eriksson *et al.*, 2008), assim como o inibidor de fXA rivaroxaban (Abrams & Emerson, 2009). Nota-se que os anticoagulantes mais comuns são os inibidores de fXa e/ou antitrombina (Champagne, 2004).

Eles podem conter diferentes classes e até mesmo mais de uma classe dentro da mesma família de artrópodes, como: inibidores de protease do tipo Kazal e da família lipocalina em barbeiros, do tipo anophelina em mosquitos anofelinos, do tipo thrombostasina em moscas de chifre, do tipo TTI em mosca tsetse. Assim como inibidores do tipo hirudina, granulina e antistasina em sanguessugas e do tipo Kunitz, madarina e chimadarina, NTI-I, microphilinas, BmAP, BmGTI e variegina em carrapatos. Além de outros que não foram caracterizados em *Tabanus*, *Simulium*, *Boophilus calcaratus*, *Anopheles stephensi* e em *Panstrongylus megistus* (Koh & Kini, 2009).

Da mesma forma, inibidores de fXa foram caracterizados e identificados como do tipo Ascaris em ancilostomídeos, draculina em morcegos e serpinas atípicas no mosquito *Aedes*. Assim como antistasina e da família hemeritrina em sanguessugas, do tipo Kunitz, da família Salp e inibidores não competitivos de *Hyalomma* em carrapatos. Assim como para inibidores de trombina, vários inibidores de fXa não foram caracterizados em *Simulium vittatum*, *Culicoides*, *Amblyomma americanum*, *Haementeria ghilianii*, *Triatoma infestans*, *Rhipicephalus appendiculatus* e em *Dermacentor andersoni* (Koh & Kini, 2009).

Alguns hematófagos também são hábeis a elaborar proteínas inibidoras do sistema de contato, como: do tipo antistasina em *Hirudo nipponia*, do tipo Kazal em *T. infestans*, e do tipo

Kunitz no mosquito *Anopheles stephensi* e nos carrapatos *Haemaphysalis longicornis*, *Boophilus microplus* e *Rhipicephalus sanguineus*. Por outro lado, *Rhodnius prolixus* elaboram inibidores do complexo tenase intrínseco do tipo lipocalina. Enquanto inibidores do complexo tenase extrínseco do tipo Kunitz foram caracterizados para os carrapatos *Ixodes scapularius* e *Dermacentor andersoni* e do tipo Ascaris em ancilostomídeos (Koh & Kini, 2009).

Estudos funcionais e de caracterização dos anticoagulantes de hematófagos têm revelado diferentes classes de inibidores (hirudina, Kunitz, Kazal) tendo o mesmo alvo de inibição, neste caso inibidores de trombina. Assim como a mesma classe, por exemplo, Kunitz, atuando em diferentes alvos de inibição, como por exemplo, trombina e fXa (Koh & Kini, 2009).

Essa diversidade molecular entre os anticoagulantes de hematófagos é resultado da evolução durante milhões de anos, do desenvolvimento de estratégias de sucesso para a alimentação sanguínea para a sobrevivência desses animais frente aos seus hospedeiros específicos, e pode ser discutida com base na divergência funcional e a convergência funcional de proteínas estruturalmente independentes. Isso explica porque membros de uma mesma família de artrópodes hematófagos podem apresentar diferentes estratégias na inibição na cascata de coagulação (Koh & Kini, 2009).

Essa diversidade de anticoagulantes pode ser vista dentro de famílias e também gêneros de dípteros. Enquanto mosquitos culicíneos elaboraram uma serpina para inibir o fXa, os anofelinos elaboraram um pequeno peptídeo de 6,5 kDa para inibir trombina (Stark & James, 1996, 1998; Francischetti *et al.*, 1999; Valenzuela *et al.*, 1999). Peptídeos também foram elaborados para inibir trombina em *Glossina morsitans* (Cappello *et al.*, 1996, 1998), no carrapato *Boophilus microplus* (Horn *et al.*, 2000; Ciprandi *et al.*, 2006) e na glândula salivar de *Triatoma pallidipennis* (Noeske-Jungblut *et al.*, 1995). Por outro lado, o nematócero *Culicoides variipennis* elaborou uma proteína ainda desconhecida para atuar na inibição de fXa (Perez de Leon *et al.*, 1998).

Em simulídeos, homogenados salivares de *Simulium vittatum* foram demonstrados inibir fXa e trombina. Posteriormente, homogenados de quatro espécies de simulídeos revelaram que as espécies zoofílicas (*S. vittatum* e *Simulium argus*) apresentam inibidores de trombina e fXa, enquanto espécies antropofílicas (*Simulium ochraceum* e *Simulium metallicum*) elaboraram somente inibidores de fXa. Entretanto, altas atividades de apirase foram reveladas para espécies antropofílicas, compensando possivelmente a ausência de inibidores de trombina (Abebe *et al.*, 1994, 1995; Cupp *et al.*, 1993).

Durante a injúria vascular pelo artrópode hematófago, as plaquetas liberam serotonina e tromboxano A<sub>2</sub>, levando à contração do vaso. As proteínas salivares com atividade vasodilatadora são importantes para esses hematófagos por dilatar os vasos e deixar o sangue fluído no local da picada, diminuindo assim o tempo de repasto sanguíneo. O estudo de vasodilatadores de hematófagos têm se revelado como um bom exemplo de diversidade salivar e evolução convergente (Ribeiro, 1987; Law *et al.*, 1992; Ribeiro, 1995; Ribeiro & Francischetti, 2003).

O uso de óxido nítrico como vasodilatador é observado no triatomíneo *R. prolixus* e no cimicídeo *C. lectularius* (Ribeiro *et al.*, 1993, Valenzuela *et al.*, 1995). Devido a instabilidade desse gás, essas espécies desenvolveram diferentes hemeproteínas para sua estabilização e transporte devido esse gás ser extremamente reativo para o hospedeiro. Entre eles estão um membro da família lipocalina para a nitroforina de *Rhodnius* (Champagne *et al.*, 1995) e um membro da família inositol fosfatase para a nitroforina de *Cimex* (Valenzuela & Ribeiro, 1998). Diferentemente, os carrapatos tem revelado prostaglandinas como a PGE<sub>2</sub> e PGF<sub>2</sub> (Dickinson *et al.*, 1976; Higgs *et al.*, 1976) para ajudar na vasodilatação.

O flebotomíneo *Lutzomyia longipalpis* é famoso por conter um peptídeo vasodilatador de 6.5 kDa denominado de “Maxadilan”, o mais potente vasodilatador conhecido (Ribeiro *et al.*, 1986; 1989; Lerner *et al.*, 1991; Lerner & Shoemaker, 1992). Este vasodilatador não foi encontrado na saliva de flebotomos do Velho Mundo, como *Phlebotomus papatasi*, o qual contém adenosina para atuar nessa atividade. Outro importante exemplo de vasodilatador foi relatado em *S. vittatum* que possui uma proteína de 15 kDa denominada SVEP (*S*imulium *V*asodilator *E*ritema *P*roteína) com atua com efeito vasodilatador prolongado em pele de coelho, em canais de cálcio dependente de ATP (Cupp *et al.*, 1994, 1998).

Em diferentes gêneros de mosquitos observa-se diferentes moléculas para atuar na vasodilatação. Enquanto os culicíneos elaboraram um decapeptídeo taquiquinina denominado sialoquinina (Champagne & Ribeiro, 1994) os anofelinos elaboraram uma peroxidase de 65 kDa, que destrói os vasoconstritores norepinefrina e serotonina (Ribeiro & Nussenzveig, 1993; Valenzuela *et al.*, 1999).

Essa diversidade de moléculas anti-hemostáticas elaboradas por diferentes gêneros e famílias de artrópodes hematófagos ajudam a entender um pouco mais sobre a evolução dos compostos salivares. Por exemplo, *Aedes* e *Anopheles* divergiram mais de 150 milhões de anos atrás ou 100 milhões de anos antes da irradiação dos mamíferos. Do mesmo modo, *Lutzomyia* e *Phlebotomus* divergiram antes da última separação das placas tectônicas que formaram a África e

a América do Sul, que também coincide com o tempo da irradiação dos mamíferos. Dessa forma, a história evolutiva desses hematófagos ajuda a explicar o porquê membros de uma mesma família, mas não do mesmo gênero, como *Lutzomyia* e *Phlebotomus* e/ou *Aedes* e *Anopheles* possuem vasodiladores completamente diferentes. Assim como *Aedes* possui inibidores de fXa e *Anopheles* apresentam inibidores de trombina. Ou ainda, porquê diferentes famílias de apirase possuem membros da família 5'nucleotidase e da família *Cimex*. Dessa forma, postula-se que todos os gêneros que evoluíram antes da irradiação dos mamíferos devam apresentar consideráveis variações na sua composição salivar (Ribeiro & Francischetti, 2003).

Alguns hematófagos têm elaborado substâncias que agem contra a dor no hospedeiro. Substâncias anestésicas verdadeiras inibem a condução do estímulo nervoso, enquanto, as substâncias que inibem a ação de agonistas de dor possuem efeitos analgésicos (Ribeiro & Francischetti, 2003). O barbeiro *T. infestans* é conhecido por inibir atividade de canais de cálcio em nervos (Dan *et al.*, 1999) enquanto outros componentes salivares com efeito potencial antidor, incluindo apirase, histamina, proteínas que ligam em serotonina foram experimentadas em carrapatos e barbeiros (Ribeiro, 1982; Paesen *et al.*, 1999), assim como cininase (à qual destrói bradicinina) em *Ixodes scapularis* (Ribeiro & Mather, 1998).

Os carrapatos são conhecidos como especialistas na elaboração de um complexo coquetel imunossupressor e anti-hemostático importante para um “farmacologista” que necessita ficar vários dias em seus hospedeiros para finalizar sua alimentação sanguínea (Ribeiro & Francischetti, 2003). O estudo da saliva desses hematófagos tem revelado componentes anti-inflamatórios e imunomodulatórios que podem impedir ou retardar o efeito deletério da resposta do hospedeiro ou contrariar farmacologicamente mediadores imunofarmacológicos de seus hospedeiros, o qual incluem anticomplemento (Valenzuela *et al.*, 2000), anti-IL-2 (Gillespie *et al.*, 2001) e inibidores de protease (Leboulle *et al.*, 2002).

A composição da saliva associada aos modos de alimentação dos hematófagos pode indicar a fase evolutiva desses animais para a hematofagia. Enquanto alguns hematófagos foram eficientes para elaborar potentes coquetéis, outros parecem ainda encontrar-se em fase de pré-adaptação para a hematofagia, como a mosca de chifre *Haematobia irritans*. Nessa espécie, somente uma substância anticoagulante e não antiplaquetários e vasodiladores foram encontrados (Cupp *et al.*, 1998).

Nos últimos 10 anos, iniciou-se uma nova fase de “descobertas” desses “farmacos” salivares. Com o advento das técnicas de biologia molecular e na tentativa de aumentar o conhecimento da complexidade das proteínas e dos transcritos expressos nesses insetos foi

possível estudar o sialoma (do grego, sialo = saliva) que mudou a forma de estudo e aumentou o conhecimento dos compostos salivares dos hematófagos, ao somar a construção de bibliotecas de cDNA com a identificação de proteínas via sequenciamento N-terminal, após sua separação por eletroforese (Francischetti *et al.*, 2002).

Antes do sialoma, um grande número de glândulas salivares era necessário para a confecção de extratos salivares, os quais eram avaliados sobre a sua função e caracterização. Para muitas espécies, a aquisição de uma grande quantidade de glândulas salivares era um fator limitante. Com o sialoma, várias ESTs são obtidas no transcriptoma, permitindo a expressão dessas proteínas e seu uso em ensaios funcionais e de caracterização. Esses avanços possibilitaram a obtenção de uma grande parte das sequências salivares depositadas no GenBank proporcionando um melhor entendimento das moléculas salivares e de sua importância para a alimentação sanguínea em um curto espaço de tempo (Ribeiro & Francischetti, 2003; Ribeiro & Arca, 2009).

Os sialotranscriptomas disponíveis revelam cerca de 40% das seqüências com função desconhecida. Observa-se que algumas famílias de proteínas são abundantemente expressas em determinadas famílias de hematófagos, como as lipocalinas em barbeiros (Assumpção *et al.*, 2008) e as proteínas D7 em mosquitos (Calvo *et al.*, 2006a) e simulídeos (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010). O estudo do sialotranscriptoma dos “farmacologistas” disponibilizam plataformas com genes de interesse disponíveis em bancos de dados públicos, que inclusive já permitiu a identificação de uma proteína candidata para vacina, como no modelo *Phlebotomus papatasi/Leishmania major*. Desse modo, podemos dizer que entramos numa nova fase do conhecimento dos “fármacos” desses “farmacologistas”.

## **1.7 Cascata da coagulação sanguínea**

Em 1964, MacFarlane (1964) e Davie & Ratnoff (1964) propuseram a hipótese de cascata para explicar a fisiologia da coagulação. O sistema de coagulação é uma série complexa de interações nas quais o sangue perde sua característica de fluidez. A formação do coágulo de fibrina envolve complexas interações entre proteases plasmáticas e seus cofatores, que culminam na gênese da enzima trombina, que, por proteólise, converte o fibrinogênio solúvel em fibrina insolúvel (Franco, 2001).

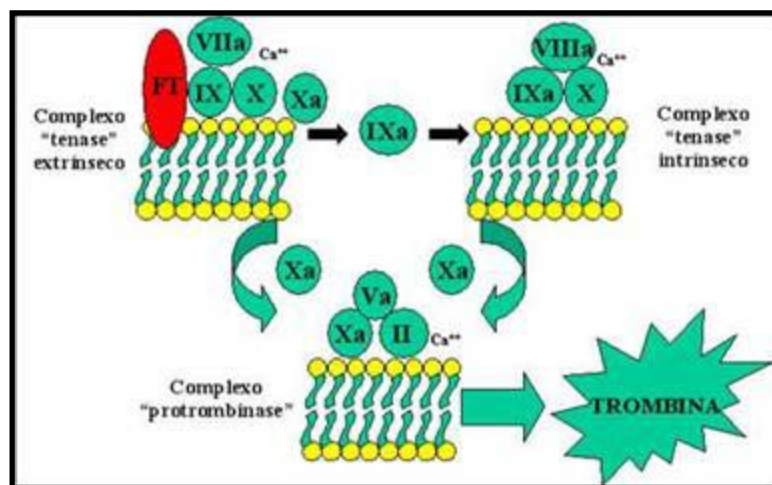
Atualmente aceita-se que os mecanismos hemostáticos estejam associados com três complexos enzimáticos procoagulantes, os quais envolvem serinoproteases dependentes de vitamina K (fatores II, VII, IX e X) associadas à cofatores (V e VIII) todos localizados em uma



superfície de membrana contendo fosfolipídeos (Jenny & Mann, 1998; Colman *et al.*, 2001). Estudos indicam que a via intrínseca e extrínseca não exibem funcionamento independente mas a utilização dos termos são úteis para a interpretação de ensaios de avaliação da hemostasia (TP e TTPA) (Franco, 2001).

As diversas enzimas da coagulação convertem seus substratos procofatores em cofatores, os quais convertem as proteases sobre as superfícies celulares, contendo fosfolipídeos. O início da coagulação se faz mediante expressão de fator tecidual (FT) que funciona como receptor para o fator VII. Em indivíduos normais, níveis mínimos de fVIIa estão presentes na circulação correspondendo a 1% da concentração total. Este se liga ao FT formando o complexo FT/fVIIa (complexo tenase extrínseco) com função enzimática ativa e com capacidade de ativar o fVII, processo de “auto-ativação”. Esse complexo tem como substratos o fator IX e X, cuja clivagem resulta na formação do fIXa e fXa, respectivamente com subsequente formação de trombina e fibrina (Franco, 2001).

Quantidades mínimas de trombina são geradas a partir do complexo “protrombinase” extrínseco. Uma vez que há gênese inicial de trombina, esta enzima é capaz de ativar o fV em fVa e o fVIII em fVIIIa. As duas reações envolvendo ativação de procofatores são fundamentais para a geração do complexo “tenase” intrínseco (fVa/fXa) que converte a protrombina em trombina. Esse complexo ativa o fator X com uma eficiência de 50 vezes maior que o complexo FT/fVIIa. O produto dessas reações, a trombina (IIa), exibe atividades procoagulantes convertendo o fibrinogênio em fibrina. Adicionalmente, os íons cálcio são necessários em diversos passos das reações da coagulação (Figura 3).



**Figura 3** – Esquema dos complexos procoagulantes (Franco, 2001).

### 1.8 Enzimas: fator Xa, trombina e lisozima

O fator X é uma proteína vitamina K-dependente, zimógeno que é sintetizado no fígado e circula no plasma. A ativação do zimógeno, fator X, é catalizada pelos fatores do complexo tenase extrínseco (fator VIIa, fator tissular, superfície celular e íons cálcio) ou intrínseco (fator IXa, fator VIIIa, superfície celular e íons cálcio). É uma serino protease de 46 kDa. O fator Xa é um componente enzimático no complexo protrombinase (fator Xa, fator Va, membrana celular carregada negativamente e íons cálcio), que cataliza a conversão rápida de protrombina para trombina. Apesar do fator Xa converter sozinho protrombina para trombina, o resultado do complexo protrombinase é aumentado 300 mil vezes na taxa de conversão de protrombina. A atividade de coagulação do fator Xa é regulada pela inativação do cofator, fator Va ou por inibição direta do fator Xa pelos inibidores, como a anti-trombina III.

A trombina é uma serino protease com propriedades polifuncionais no sistema hemostático do vertebrado (Fenton, 1986). Ela existe em três formas:  $\alpha$ -,  $\beta$ - e  $\gamma$ - (Fenton *et al.*, 1977; Workman *et al.*, 1977). As formas  $\beta$ - e  $\gamma$ - são de pouca importância com capacidades de coagulação consideravelmente diminuídas (Fenton *et al.*, 1977). A  $\alpha$ -trombina (EC 3.4.21.5) possui um papel central na coagulação sanguínea e com função primária de quebrar fibrinogênio gerando fibrina e um coágulo insolúvel. É altamente específica com peso molecular de 36 kDa gerada pela ativação proteolítica do zimógeno protrombina.

Além disso, ela tem função regulatória na ativação de fatores da coagulação sanguínea (V, VIII, e XIII e proteína C na cascata de coagulação). A principal importância da trombina para o processo de coagulação geral é estabelecido pelo fato que qualquer perturbação dentro do sistema de coagulação sanguínea que resulte numa ampliação significativa ou deficiente bem como na geração de trombina acelerada ou atrasada irão resultar em eventos trombóticos ou hemorrágicos clinicamente relevantes. A atividade de trombina é regulada pela inibição dos cofatores, fator Va e VIIIa ou pela inibição direta de trombina pelo seu inibidor principal, a antitrombina III.

A lisozima (EC 3.2.1.17) é uma enzima que hidrolisa a ligação glicosídica  $\beta$ -(1,4) entre os resíduos glicídicos, ácido N-acetilmurâmico e N-acetilglicosamina, componentes do peptidoglicano em paredes celulares bacterianas. Em bactérias gram-negativas, além da ação enzimática, a lisozima parece interagir com a membrana externa desestabilizando-a. Essa enzima atua como primeira linha de resistência contra bactéria (Daffre *et al.*, 1994). Em insetos sua função tem sido relatada para defesa imune e digestão (Daffre *et al.*, 1994; Moreira-Ferro *et al.*, 1998; 1999).

## 2 JUSTIFICATIVA

A glândula salivar de artrópodes é um órgão especializado em secretar saliva, que é fundamental para o sucesso da alimentação sanguínea e ajuda na transmissão de patógenos. Em simulídeos, pouco se conhece sobre os aspectos morfológicos e moleculares desse órgão. Nesse trabalho é proposto pela primeira vez o estudo da glândula salivar de uma espécie de simulídeo (*T. guianense*) associada com a transmissão de *Onchocerca volvulus*. Estudos de caracterização morfológica, bioquímica e molecular da glândula salivar desse simulídeo fornecerá importantes informações sobre o funcionamento do órgão e sobre a composição da saliva desse inseto. O estudo do sialotranscriptoma através da construção de bibliotecas de cDNA proporcionará a identificação dos genes codificantes das proteínas, possibilitando aos genes de interesse uma posterior clonagem e investigação em ensaios funcionais. A análise proteômica permitirá a validação do transcriptoma e representará uma importante ferramenta no campo biotecnológico na descoberta de novas proteínas. Além disso, todas as informações obtidas sobre a glândula salivar de *T. guianense* permitirá um melhor entendimento do papel da saliva desse inseto e sua possível função na transmissão da filárias, assim como permitirá o estudo da evolução de proteínas salivares dentro da família Simuliidae e fornecerá uma plataforma de proteínas anti-hemostáticas e candidatas para uma vacina contra oncocercose.

### **3 OBJETIVOS**

#### **3.1 Objetivo geral**

Estudar aspectos morfológicos, bioquímicos e moleculares da glândula salivar de fêmeas de *T. guianense*.

#### **3.2 Objetivos específicos**

Caracterizar a histologia, topografia e microanatomia, ultraestrutura e a presença de resíduos de carboidratos nas glândulas salivares de fêmeas de *T. guianense*.

Identificar atividades biológicas como anticoagulantes (fator Xa e trombina) e atividade da enzima lisozima no homogenado de glândula salivar de *T. guianense*.

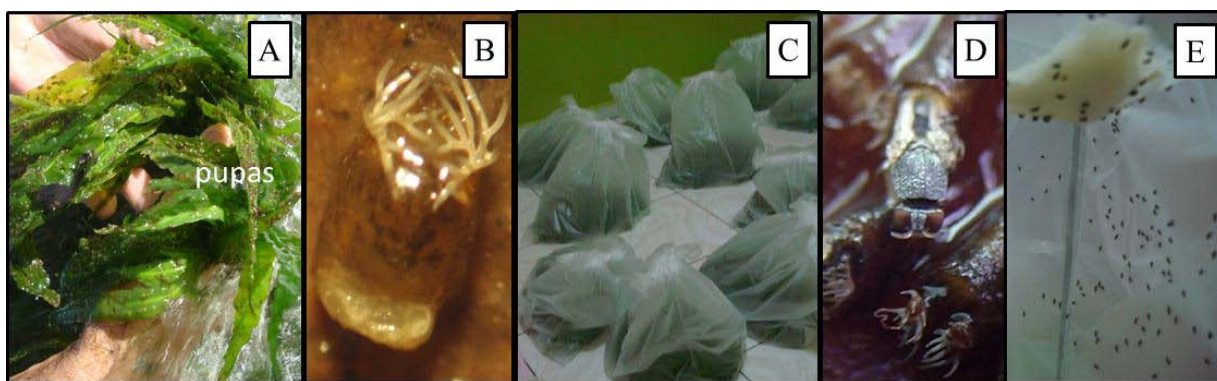
Identificar as proteínas expressas na glândula salivar de fêmeas de *T. guianense* com a construção de uma biblioteca de cDNA e com o estudo do proteoma via espectrometria de massa.

## 4 MATERIAL E MÉTODOS

### 4.1 Coleta de simulídeos

As coletas de simulídeos foram realizadas de janeiro a março nos anos de 2007 a 2010 na cachoeira do Travessão, localizada no rio Jauaperi, município de Rorainópolis, estado de Roraima, Brasil. A coleta consistiu da remoção da parte vegetativa de macrófitas da família Podostemaceae com presença de pupas de simulídeos (Figura 4A e B). As macrófitas foram acondicionadas em sacos plásticos de 50 litros com pequenos furos, e esses sacos foram transportados dentro de caixas de isopor com gelo artificial reutilizável (gelox) para manter uma baixa temperatura até o laboratório localizado na cidade de Manaus (distante oito horas de carro do local de coleta).

No laboratório, o material coletado foi parcialmente enxugado e dividido em novas sacolas de 50 litros com abertura lacrada e pequenos furos feitos com agulha para favorecer a entrada de ar (Figura 4C). As sacolas foram mantidas em ambiente escuro e com temperatura de  $\pm 30^{\circ}\text{C}$ . Após a emergência, as fêmeas de *T. guianense* foram identificadas (Figura 4D) segundo as chaves de identificação de Shelley *et al.* (1997) e somente as fêmeas foram transferidas para gaiolas de filó contendo um chumaço de algodão embebido em 10% de mel comercial (Karo®) e outro chumaço embebido com água (Figura 4E). As gaiolas foram mantidas no escuro para diminuir a atividade dos simulídeos, a  $18^{\circ}\text{C}$  até o momento da dissecação das glândulas salivares das fêmeas.



**Figura 4** – Coleta e manutenção de simulídeos. A: Coleta de pupas em macrófitas da família Podostemaceae; B: Pupa de *T. guianense*; C: Sacolas de emergência; D: Emergência de simulídeo e E: Alimentação das fêmeas com mel comercial e água. Fotos: Chagas, AC.

## **4.2 Dissecção da glândula salivar**

A dissecção da glândula salivar foi realizada através de imobilização do inseto vivo em uma gota de NaCl 150 mM, pH 7,4 em lâmina escavada. Um pequeno corte foi feito próximo à intersecção da cabeça com o tórax, na região ventral da fêmea, para aumentar a área de passagem da glândula salivar durante a dissecção. Com o auxílio de estiletes, a cabeça do inseto foi levemente puxada em direção oposta ao restante do corpo promovendo a remoção da glândula intacta.

A glândula salivar de fêmeas recém emergidas até o quinto dia foi acondicionada em tubos contendo glutaraldeído 2,5% ou paraformaldeído 4% para estudos morfológicos. Amostras para ensaios de atividade biológica foram armazenadas em grupos (n = 1, 5, 10 e 20 pares de glândula salivar) em tubos contendo 4µL/par de glândula de NaCl 150 mM pH 7,4. As amostras foram mantidas em freezer -70°C até o uso.

Para a construção da biblioteca de cDNA, 25 glândulas de fêmeas de cada idade (recém-emergidas, 1º dia e 2º dia) foram armazenadas em 50 µL de RNA later RNAlater (Ambio, Inc., Austin, TX), mantidas a 4°C por três dias e posteriormente a -70° C até o momento da extração do RNA. Para a análise proteômica foram dissecadas 50 glândulas salivares de fêmeas com dois dias de idade, armazenadas em tubos contendo 50 µL de NaCl 150 mM, pH 7,2 e mantidas a -70°C até o uso.

## **4.3 Microscopia de luz**

As glândulas salivares de fêmeas foram fixadas por duas horas em solução de 2,5% de glutaraldeído em 100 mM de tampão cacodilato de sódio, pH 7,2. Em seguida, as glândulas foram lavadas (3X) em tampão PBS. A desidratação foi realizada em séries crescentes de etanol (30% a 100% - a última três vezes) por 10 minutos cada. A inclusão foi feita em três passos: i) em uma mistura de 1:1 de etanol absoluto e historesina Leica® por cinco minutos; ii) somente em historesina Leica® por um período de 24 horas em temperatura ambiente e iii) em 15:1 de historesina Leica® com endurecedor. Cortes seriados de 1 µm de espessura foram produzidos no material com o uso do micrótomo rotativo, modelo Leica®. Para a marcação de lipídios, as glândulas foram previamente incubadas com ósmio por duas horas, antes do passo da desidratação. Todos os cortes histológicos foram aderidos em lâminas e submetidos à coloração com azul de toluidina, PAS, ósmio e mercúrio bromofenol. O material foi visualizado e documentado em sistema fotográfico acoplado ao microscópio laser confocal (Zeiss®-LSM 510).

#### **4.4 Microscopia eletrônica de varredura (MEV)**

As glândulas salivares foram fixadas conforme descrito no item 4.3. As amostras foram lavadas três vezes em tampão PBS e pós-fixação em solução 1:1 de 1% de tetróxido de ósmio e 0,8% de ferricianeto de potássio (em 100 mM de tampão cacodilato, pH 7,2) por duas horas a temperatura ambiente e ao abrigo de luz (Pimenta & De Souza, 1985). Em seguida, as glândulas salivares foram desidratadas em séries crescentes de acetona (30% a 100% - a última três vezes) por 10 minutos cada. As glândulas foram submetidas à passo de secagem com o uso do aparelho de ponto crítico em CO<sub>2</sub>. O passo final consistiu da montagem das glândulas em suportes cobertos com fita de carbono, que foram submetidos à metalização com ouro (Pimenta & De Souza, 1986). As amostras foram observadas e documentadas em microscópio eletrônico de varredura (Jeol® JSM-5600).

#### **4.5 Microscopia eletrônica de transmissão (MET)**

As glândulas salivares foram fixadas conforme descrito no item 4.3 e pós-fixadas conforme descrito no item 4.4. Em seguida, as amostras foram lavadas três vezes em tampão PBS e desidratadas em concentrações crescentes de acetona (30% a 100% - a última três vezes) por 10 minutos. Em seguida as glândulas foram infiltradas em 1:1 de acetona/Epon, e posteriormente em Epon puro até a polimerização dos blocos. Os cortes ultrafinos foram obtidos em ultra-micrótomo (Leica®) utilizando navalha de diamante, sendo coletados em grades de 200 “mesh”. Os cortes foram contrastados por 20 minutos em acetato de uranila aquoso e por três minutos em citrato de chumbo, seguida de sucessivas lavagens com água destilada (Pimenta & De Souza, 1986). O material foi observado e documentado no microscópio eletrônico de transmissão, modelo JEM 1011, JEOL (Tokyo, Japão) e os registros fotográficos foram obtidos por meio da câmera digital Gatan BioScan Modelo 792 e software de captação de imagens Gatan Digital Micrograph 3.6.5 (Gatan Inc., Warrendale, USA).

#### **4.6 Microscopia laser confocal (MLC)**

As amostras foram fixadas em solução de 4% de paraformaldeído por duas horas e lavadas três vezes em tampão PBS, cinco minutos cada, e incubadas em meio RPMI por 30 minutos em temperatura ambiente. As amostras foram incubadas em solução de 1% de PBS/BSA por 30 minutos em câmara úmida. As etapas seguintes foram processadas ao abrigo da luz. As glândulas foram incubadas por uma hora à temperatura ambiente na presença de diferentes tipos de lectinas fluorescentes (Sigma Co., USA) (Tabela 1). Posteriormente, as amostras foram

lavadas em solução de 1% de PBS/BSA, três vezes, cinco minutos cada, e incubadas com marcador nuclear (DAPI - 4,6 diamidino-2-phenylindole por 30 minutos. As glândulas foram novamente lavadas em PBS 1X, três vezes por cinco minutos cada e posteriormente montadas em Mowiol entre lâmina e lamínula. As glândulas foram analisadas quanto à presença de diferentes resíduos de carboidratos e documentadas em microscópio laser confocal (Zeiss®-LSM 510).

**Tabela 1** - Lectinas fluorescentes-FITC utilizadas para verificação de resíduos de carboidratos de superfície na glândula salivar de simúlideos

| Lectinas                                     | SIGLA | Resíduos                               |
|--|-------|--|
| <i>Canavalia ensiformis</i> (Concanavalin A) | Con A | $\alpha$ -man, $\alpha$ -glc           |
| <i>Bandeiraea simplicifolia</i>              | BS1   | $\alpha$ -gal, $\alpha$ -galNAc        |
| <i>Arachis hypogaea</i> (peanut)             | PNA   | B-gal(1→3)galNAc                       |
| <i>Helix pomatia</i>                         | HPA   | GalNAc                                 |
| <i>Triticum vulgare</i> (Wheat germ)         | WGA   | (GlcNAc) <sub>2</sub> , NeuNAc         |
| <i>Limulus polyphemus</i> (horseshoe crab)   | LPL   | NeuNAc, GalNAc, GlcNAc                 |
| <i>Ulex europaeus</i>                        | UEA   | $\alpha$ -L-Fuc, (GlcNAc) <sub>2</sub> |

Catálogo Sigma (2001)

#### 4.7 Preparo de homogenado de glândula salivar (HGS)

As amostras de glândulas salivares em “pool” foram sonicadas durante três minutos e posteriormente centrifugadas a 10.000 g, 4°C, 10 minutos. O sobrenadante foi transferido para um novo tubo eppendorf e denominado como homogenado de glândula salivar (HGS).

#### 4.8 Dosagem de proteína do HGS

A quantificação de proteína presente no HGS foi mensurada para cinco grupos de amostras, cada uma contendo 20 glândulas salivares de dois dias de idade. Para cada leitura de proteína foi utilizado 2  $\mu$ L do HGS. Uma unidade de absorbância a 280<sub>nm</sub> equivale a 1mg/ml de proteína, no instrumento Nanodrop 3300 da Thermo Scientific.

#### 4.9 Ensaio do tempo de recalcificação

O ensaio de tempo de recalcificação foi realizado no leitor de microplaca, SpectraMax (Molecular Devices, USA) utilizando o módulo cinético disponível no software do equipamento. O plasma humano normal utilizado no ensaio foi obtido na forma de liofilizado da American Diagnostica, sendo reconstituído em 1 mL de água miliQ conforme o protocolo específico do fabricante. A concentração de proteína presente na amostra de HGS utilizada no ensaio foi previamente mensurada.



O ensaio foi conduzido em microplacas de 96 poços de fundo reto contendo: 25 µL de NaCl 150 mM, pH 7,4 (controle) ou 25 µL de amostra teste. A amostra teste consistiu de diluições seriadas do HGS. Nesse ensaio a concentração inicial foi 73 até 0,07 µg/ml. Em cada amostra foi adicionado 25 µL de plasma humano normal e posteriormente a placa foi incubada a 37°C por 15 minutos. A leitura das amostras foi iniciada pela adição de 25 µL de CaCl<sub>2</sub> 20 mM pré-aquecido e a formação do coágulo de fibrina foi acompanhada pela leitura da absorbância em 600<sub>nm</sub>, em intervalos de 10 segundos por um período de uma hora a 37°C. O valor do Onset Time estabelecido foi de 0,02. O ensaio foi conduzido em triplicata e os resultados foram interpretados como aumento relativo do tempo de recalcificação do plasma.

#### **4.10 Ensaio de tempo de tromboplastina parcial ativada e tempo de trombina**

Os ensaios de tempo de tromboplastina parcial ativada (TTPa) e tempo de protrombina (TP) foram conduzidos em coagulômetro ST art 4 analyser da Diagnostica Stago. O plasma humano normal utilizado nesse experimento foi o mesmo obtido para o ensaio de tempo de coagulação. Os ensaios de rotina TP e TTPa foram avaliados na ausência (controle) e presença de HGS. O ensaio de TP mede a atividade de fatores da via extrínseca (II, V, VII, X), enquanto o TTPa mede a atividade dos fatores da via intrínseca (II, V, VIII, IX, X, XI, XII).

O ensaio de TP foi conduzido em cubetas específicas do coagulômetro na presença de uma bola magnética da ST art 4 da Diagnostica Stago. Na cubeta controle foi adicionado 50 µL de plasma humano normal e nas cubetas testes foi adicionado além do plasma, diferentes concentrações de HGS (3,2; 9,6; 28,8; 48 e 96 µg/ml). Todas as amostras foram incubadas por um minuto a 37°C no coagulômetro e a reação foi iniciada com a adição de 50 µL de TP reagente. O ensaio de TTPa foi realizado similarmente ao ensaio de TP, constituído de: 50 µL de plasma humano, 50 µL de TTPa reagente, incubação por três minutos a 37°C no coagulômetro e adição de 50 µL de CaCl<sub>2</sub> 0,025 mM (controle) para o início da reação. As concentrações das amostras testes foram de 2,1; 6,3; 18,9; 31,5 e 63µg/ml.

#### **4.11 Ensaio cromogênico de fator Xa e trombina**

As enzimas humanas fXa e trombina foram obtidas da Haematologic Technologies. Os substratos cromogênicos S-2222 (Methoxycarbonil-D-ciclohexylglicil-glicil-argina-para-nitroanilide acetato) e S-2238 (H-D-hexahidrotirosol-alanil-arginina-para-nitroanilide diacetato) foram obtidos da Diapharma Group.

Inicialmente foi preparado 10 mL da enzima fXa 0,5 nM em tampão Tris 25 mM pH 7,4, NaCl 150 mM, 0,5% BSA, CaCl<sub>2</sub> 5mM e 10 mL de trombina 0,1 nM em tampão Tris 25 mM pH 7,4, NaCl 150 mM, 0,5% BSA. Os ensaios cromogênicos de fXa e trombina foram realizados no leitor de microplaca SpectraMax (Molecular Devices, USA), utilizando o módulo cinético disponível no software do equipamento. O experimento foi montado em placas de 96 poços de fundo reto e realizados em triplicatas.

O ensaio de fXa consistiu de 25 µL de TBS (Tris HCl 150 mM pH 7,4, NaCl 150 mM) para o controle ou 25 µL de HGS com concentração de 83 até 0,08 µg/mL. Para ambos os casos, foi adicionado 75 µL da enzima fXa 0,5 nM. A placa foi incubada por 15 minutos à 37°C e a reação foi iniciada pela adição de 5 µL do substrato S-2222 (5 mM), específico para fXa. A degradação do substrato cromogênico foi medida em espectrofotometro a 405<sub>nm</sub>, 37°C, com leitura a cada 10 segundos durante 1h. O ensaio com a enzima trombina foi feito da mesma forma, exceto para a concentração da enzima trombina (0,1 nM) com a concentração do HGS inicial de 57 até 0,05 µg/mL e pelo substrato específico da enzima (S-2238).

#### **4.12 Ensaio quantitativo da atividade da enzima lisozima**

A atividade de lisozima em glândula salivar de fêmeas de 24 e 48h após a emergência foi medida em diferentes pHs, conforme o protocolo de Moreira-Ferro *et al.*, 1998, 1999 em gel de agarose (1%) contendo 1 mg/ml da bactéria *Micrococcus lysodeikticus* em tampão fosfato de sódio 22 mM pH 5.0 com NaCl 117 mM ou em tampão fosfato 67 mM pH 6,0 a 8,0 com NaCl 50 mM. Os homogenados salivares foram aplicadas em poços de 2 mm de diâmetro e as placas foram incubadas por 16 horas a 37°C. Em caso de atividade enzimática da lisozima positiva, uma zona de degradação na camada de bactéria será aparente. O experimento foi conduzido em duplicata, utilizando lisozima comercial (50 mg/mL) (Sigma-Aldrich) como controle positivo e água destilada como controle negativo.

#### **4.13 Transcriptoma e proteoma**

##### **4.13.1 Construção da biblioteca de cDNA**

O isolamento do RNA das 75 glândulas salivares foi realizado com o kit mRNA Micro-FastTrack (Invitrogen, San Diego, CA). Para a construção da biblioteca de cDNA foi utilizado o kit SMART (switching mechanism at 5' end of RNA transcript) (Clontech, Palo Alto, CA). A síntese da primeira fita foi obtida usando a transcriptase reversa PowerScript a 42°C por 1

hora na presença de SMART IV e primers CDS III (3'). A síntese da segunda fita foi obtida usando o protocolo baseado numa PCR de longa distância (LD), usando o mix Advantage Taq polymerase (Clontech) na presença do primer PCR 5' e do primer CDS III (3').

A síntese do cDNA resultou na criação de locais para enzima de restrição *SfiI* A e B nas extremidades do produto do PCR que foram usadas para clonagem dentro do vetor de fago (lambda TriplEx2, Clontech). As condições da PCR foram: 95°C por 1 minuto; 26 ciclos de 95°C por 15 segundos, 68°C por 6 minutos. Uma parte do cDNA obtido pela PCR foi analisado em E-Gel® 1,2% com SYBR Safe (Invitrogen) possibilitando avaliar a qualidade e o tamanho do cDNA sintetizado. O cDNA dupla-fita foi imediatamente tratado com proteinase K (0,8 µg/mL) a 45°C por 20 minutos, e posteriormente a enzima foi removida por ultrafiltração com filtros de 10kDa (YM-100 Microcon filter, Amicon Inc., Beverly, CA).

O cDNA dupla-fita foi digerido com *SfiI* a 50°C por 2 horas seguido do fracionamento em coluna ChromaSpin- 400 (Clontech). Os transcritos foram classificados baseados no perfil eletroforético em pequenos (P), médios (M) e grandes (G) após visualização em E-Gel® 1,2% com SYBR Safe (Invitrogen). As frações selecionadas foram agrupadas e concentradas usando um filtro 10kDa Microcon YM-100.

O cDNA concentrado de cada tamanho (P, M e G) foi ligado separadamente no vetor λ TriplEx2 (Clontech) e o resultado da ligação foi empacotado em GigaPack® III Plus (Stratagene, La Jolla, CA) de acordo com as instruções do manufaturado. A biblioteca empacotada foi plaqueada por infectar a fase log de células XL1- Blue de *Escherichia coli* (Clontech). A porcentagem de clones recombinantes foi determinada pela seleção azul-branco em placas de LB/MgSO<sub>4</sub> contendo X-gal/IPTG. Os recombinantes foram avaliados pelo PCR usando os primers do vetor PT2F1 (AAG TAC TCT AGC AAT TGT GAG C) e PT2R1 (CTC TTC GCT ATT ACG CCA GCT G) que flaqueiam o cDNA inserido, com subsequente visualização dos produtos em E-Gel® 1,2% com SYBR Safe (Invitrogen).

#### **4.13.2 Sequenciamento do cDNA**

Vinte placas de 96 poços foram preparadas para o sequenciamento contendo 94 clones e dois DNAs controles. Inicialmente a biblioteca de cDNA foi plaqueada em placas de petri de 150 mm com meio LB/MgSO<sub>4</sub> contendo X-gal/IPTG. Os clones recombinantes (branco) foram selecionados aleatoriamente e transferidos para microplacas de 96 poços (Nunc, Rochester, NY) contendo 75 µL de água ultrapura (KD Medical) em cada poço. As placas foram seladas e

transferidas para uma plataforma agitadora durante trinta minutos a temperatura ambiente. A suspensão do fago foi imediatamente usada no PCR ou armazenada por 4°C até o uso.

Para a amplificação do cDNA, uma alíquota de 5 µL da amostra do fago foi utilizada como molde para a reação do PCR. Os primers foram sequenciados do vetor λ TriplEx2 e denominados PT2F1 (AAG TAC TCT AGC AAT TGT GAG C) e PT2R1 (CTC TTC GCT ATT ACG CCA GCT G), posicionado no 5' e no 3' final do inserto do DNA, respectivamente. A reação foi realizada em microplacas de PCR com 96 poços (Applied Biosystems, Inc., Foster City, CA) usando FastStart Taq polymerase (Roche diagnostics). A reação foi feita em máquina de PCR modelo GeneAmp PCR system 9700 (Perkin Elmer Corp., Foster City, CA) e as condições da PCR foram: 75°C por 3 minutos, 94°C por 4 minutos, 30 ciclos de 94°C por 1 minuto, 49°C por 1 minuto; 72°C por 4 minutos. Os produtos amplificados foram analisados em E-Gel® 1,2% com SYBR Safe (Invitrogen).

Os clones foram amplificados por PCR e os produtos que demonstraram uma única banda foram selecionados para sequenciamento. Aproximadamente 200–250 ng de cada produto de PCR foi transferido para uma microplaca de PCR com 96 pocos (Applied Biosystems) e congeladas a –20°C. As vinte placas contendo 1 µL do primer (PT2F3, TCT CGG GAA GCG CGC CAT TGT) e 5 µL do molde foram enviadas em gelo seco para o Laboratório de Genômica de Rocky Mountain, USA. As reações de sequenciamento foram realizadas conforme recomendado pelo kit Applied Biosystems BigDye® Terminator v3.1 Cycle Sequencing ao adicionar 1 µL de ABI BigDye® Terminator Ready Reaction Mix v3.1 (P/N 4336921), 1,5 µL do 5X tampão de sequenciamento ABI (P/N 4336699) e 3,5 µL de água para completar um volume final de 10 µL. O ciclo do sequenciamento foi de 96°C por 10 segundos, 50°C por 5 segundos, 60°C por 4 minutos por 27 ciclos em termociclador Bio-Rad Tetrad 2 (Bio-Rad Laboratories, Hercules, CA) ou ABI 9700 (Applied Biosystems, Inc., Foster City, CA). A extensão dos produtos marcados fluorescentemente foram purificados pelo protocolo de purificação da Applied Biosystems BigDye® XTerminator™ e subsequentemente processados em um analisador de DNA ABI 3730xL (Applied Biosystems, Inc., Foster City, CA).

#### **4.13.3 Procedimentos e ferramentas de bioinformática**

As seqüências dos primers e dos vetores foram removidas das etiquetas de seqüências expressas (EST). As ferramentas do BLAST (Altschul *et al.*, 1997), CAP3 (Huang & Madan, 1999) e o software ClustalW (Thompson *et al.*, 1997) foram usados para comparar, assemblar e

alinhar as sequências, respectivamente. As análises filogenéticas e as análises parcimônicas das filogenias foram feitas com o programa Mega (Kumar *et al.*, 2004). A anotação funcional dos transcritos foi baseada na ferramenta do BlastX (Altschul *et al.*, 1997) para comparar as sequências de nucleotídeos contra o banco de dados de proteínas não-redundantes do Centro Nacional para Informação em Biotecnologia (NCBI, National Library of Medicine, NIH) e para o banco de dados de Ontologia Gênica (GO) (Ashburner *et al.*, 1997). A ferramenta blast específico de posição reversa (RPSBLAST) (Altschul *et al.*, 1997) foi usada para pesquisar domínios conservados no Pfam (Bateman *et al.*, 2000), SMART (Schultz *et al.*, 2000), Kog (Tatusov *et al.*, 2003) e no banco de dados de domínios conservados (CDD) (Marchler-Bauer *et al.*, 2002). Os dados também foram comparados com outros subgrupos de sequências de nucleotídeos de RNA ribossomal e mitocondriais baixadas do NCBI. Por último, as sequências foram submetidas para o servidor Signal P (Nielsen *et al.*, 1997) para ajudar a identificar produtos traduzidos que poderiam ser secretados. Os locais de O-glicosilação nas proteínas foram preditos com o programa NetOGlyc (Julenius *et al.*, 2005). A anotação funcional dos transcritos foi baseada em todas as comparações acima que resultaram na classificação dos transcritos em secretórios (S), celulares (H) ou de função não conhecida (U), com subdivisões baseadas na função e/ou famílias de proteínas. Sequências derivadas de elementos de transposição (TE) também foram encontradas.

#### **4.13.4 Caracterização proteômica usando gel de eletroforese 1D e espectrometria de massa (MS).**

A fração da proteína solúvel dos homogenados salivares de *T. guianense* correspondente a 50 µg de proteína foi carregada em tampão de redução Laemmli. A amostra foi aquecida por 10 minutos e aplicada em gel pré-fabricado Bis-Tris NuPAGE 4-12%. As proteínas foram separadas e visualizadas após coloração com SimplyBlue (Invitrogen). O gel foi cortado em 32 pedaços individuais e posteriormente descorados e digeridos durante toda a noite com tripsina a 37°C. Os peptídeos foram extraídos e desalinizados usando ZipTips (Millipore, Bedford, MA) e resuspendidos em 0,1% TFA antes da análise S. A cromatografia de massa de fase líquida nanofluxo (RPLS-MS/MS) foi realizada usando o sistema Agilent 1100 nanoflow LC (Agilent technologies, Palo Alto, CA) acoplado online com uma espectrometria de massa linear de captura de íons (LIT) (LTQ, ThermoElectron, San José, CA), conforme protocolo estabelecido pela Departamento de Tecnologias em Pesquisa do Twinbrook/NIH.

Colunas NanoRPLC foram acondicionadas com 5  $\mu\text{m}$ , tamanho do poro 300-Å fago C-18 (Jupiter, Phenomenex, CA) em um capilar de sílica fundida 75- $\mu\text{m}$  i.d. x 10-cm (Polymicro Technologies, Phoenix, AZ). Depois da injeção da amostra, a coluna foi lavada por 30 minutos com 98% de fase móvel A (0,1% ácido fórmico em água) a 0,5  $\mu\text{L}/\text{min}$  e os peptídeos foram eluídos usando um gradiente linear de 2% de fase móvel B (0,1% ácido fórmico em acetonitrila) para 42% de fase móvel B em 40 min a 0,25  $\mu\text{L}/\text{min}$ , em seguida para 98% B por uns 10 minutos adicionais. O espectrômetro de massa-LIT foi operado em um modo dependente dos dados MS/MS, no qual cada varredura completa do MS foi seguida por sete varreduras MS/MS onde os sete mais abundantes íons foram dinamicamente selecionados para dissociação induzida por colisão (CID) usando uma energia de colisão normalizada de 35%. A exclusão dinâmica foi aplicada para minimizar a seleção repetida de peptídeos previamente selecionados pelo CID.

O espectro de massa foi pesquisado usando “SEQUEST” contra o banco de proteínas de *T. guianense*. Após a identificação dos peptídeos trípticos por MS, eles foram convertidos para o formato de blocos do Prosite (Hulo *et al.*, 2006) por um programa que escreve em Visual Basic. Esse banco de dados foi usado para pesquisar similaridades no banco de dados de proteínas salivares do formato em Fasta, usando um programa Seedtop que é parte do pacote Blast.

## 5 RESULTADOS

### 5.1 Análise morfológica da glândula salivar de *Thyrsopelma guianense*

A glândula salivar de *T. guianense* apresentou dimorfismo sexual (Figura 5A-B), tendo em machos uma forma sacular e em fêmeas uma forma lobular-sacular. Em fêmeas a glândula salivar apresentou-se como um órgão constituído de duas glândulas (principal e acessória) de cada lado (direito e esquerdo) (Figura 5A). Da glândula acessória parte um ducto que se conecta com o ducto da glândula principal formando o ducto individual, que ao se encontrar com o ducto individual do lado oposto formam o ducto comum por onde a saliva é liberada.

A glândula principal da fêmea apresenta forma lobular e a glândula acessória uma forma sacular (Figura 5A). A glândula acessória de fêmeas de *T. guianense* revelou um aspecto sacular não uniforme (lembrando a união de vários pequenos lobos). Trabalhos anteriores tem denominado essa região sacular como reservatório salivar. A glândula acessória da fêmea se assemelha a forma sacular vista em glândulas de machos após dois dias de emergência (Figura 5A-B, 7B ).

A glândula principal da fêmea apresenta-se em forma de U sendo subdividida em região proximal e distal (Figura 5A) com tamanhos variando de 520 a 977  $\mu\text{m}$  de comprimento e 47 a 85  $\mu\text{m}$  de largura (para as regiões proximal e distal) até o 4º dia após a emergência. A glândula salivar de machos com dois dias de idade apresenta 117 x 46,8  $\mu\text{m}$  de diâmetro.

Na microscopia de luz foi observada uma fina membrana envolvendo toda a glândula salivar de fêmeas (Figura 6). Essa membrana é constituída por partículas pigmentadas que fornecem uma determinada “coloração” para a glândula. O acúmulo dos pigmentos foram observados principalmente em glândulas de dois dias de idade podendo ocorrer em diferentes regiões da glândula (Figura 6A-B). Os pigmentos também foram marcados pela lectina WGA na glândula principal da fêmea (Figura 11D). A intensidade desses pigmentos dificultou a observação das estruturas internas da glândula.

Ambas as glândulas salivares de fêmeas de *T. guianense* apresentaram um epitélio frouxo logo após a emergência, entretanto, com o envelhecimento glandular notou-se um maior intumescimento do epitélio das cavidades secretoras da glândula principal (Figura 7D-F) e no epitélio da glândula acessória (Figura 7A-C).

A glândula salivar de fêmeas de *T. guianense* apresentaram poucas estruturas externas, com destaque para uma constrição proeminente na região distal da glândula principal (Figura 7E-

F). Essa constrição foi proeminente em glândulas jovens e menos evidente ao longo dos dias devido o intumescimento glandular. Aberturas e/ou poros foram observados no epitélio da região proximal da glândula principal (Figura 8E). Essas aberturas possuem aproximadamente 2,22 x 2,51  $\mu\text{m}$  de diâmetro com presença ou não de fibras nervosas (Figura 8E). O epitélio glandular possui várias fibras nervosas além de traquéias diminutas, ambas localizadas na glândula acessória e na região proximal da glândula principal (Figura 8A-F).

A glândula principal apresenta cavidades secretoras extracelulares para estocar a saliva enquanto a glândula acessória apresenta um lúmen para exercer essa função (Figura 13-14). O ducto cuticular interno foi observado somente na região proximal da glândula principal (Figura 11E). Após o fraturamento e tratamento para microscopia eletrônica de varredura da glândula principal, da glândula acessória e do ducto comum nós observamos diferentes perfis de consistência da secreção sendo mais frouxa dentro das glândulas (Figura 9B-C) e mais densa dentro do ducto comum (Figura 9D). No interior da glândula principal nota-se a presença de pequenas aberturas com tamanho variando de 3 a 7  $\mu\text{m}$  de diâmetro aparentando ser os ductos que partem de cada cavidade secretora extracelular (Figura 9B). A secreção salivar foi revelada pelas marcação com várias lectinas entre elas LPA, ConA e PNA (Figura 10A-D).

A glândula principal de fêmea de *T. guianense* apresentou poucas células de tamanho grande com um único núcleo grande (Figura 11A), além de um citoesqueleto em forma de ácino como visto pela marcação com faloidina-FITC (Figura 11F). A glândula acessória apresentou muitas células com um núcleo pequeno cada uma (Figura 11A). A glândula salivar de macho apresentou perfil celular similar ao observado para glândula acessória da fêmea (11B). A delimitação das cavidades secretoras da glândula principal foram marcadas com a lectina WGA (Figura 11C). Os núcleos foram evidenciados por DAPI (Figuras 11A-B).

A histologia da glândula salivar de fêmeas revelou cavidades secretoras ricas em substâncias de natureza glicídica (marcação de carboidratos neutros com PAS – Figura 12A), de caráter basófilo (marcação com azul de toluidina) (Figura 12B), sendo mais basófila nas cavidades secretoras do que no lúmen da glândula acessória, além de lipídicas (marcação com ósmio – Figura 12D) e proteicas (marcação com mercúrio bromofenol – Figura 12C). Várias vesículas secretoras foram visualizadas nas duas glândulas.

A microscopia eletrônica de transmissão revelou uma grande quantidade de pequenas células na região de contato das glândulas (Figura 14D). A glândula principal apresenta dois tipos celulares (células claras e escuras – Figura 13A) apoiada numa fina lâmina basal (Figura 13A) contendo cavidades secretoras altamente elétron-densas (Figuras 13A-D) e circundadas por



pequenas microvilosidades (Figuras 13A-D), além de diferentes tipos de invaginações de membrana (Figura 13C). As células contém uma grande quantidade de mitocôndrias e vesículas, e núcleo com nucléolo aparente (Figura 13A, D). O núcleo apresentou formato alongado e quase completamente comprimido contra a membrana celular (Figura 13D). Delimitações celulares foram observados entre as cavidades secretoras (Figura 13B).

A glândula acessória de fêmea apresentou um grande lúmen responsável por secretar e armazenar a saliva, além de secreção diferente da observada para a glândula principal, sendo altamente elétron-lucente (Figura 14A, C). Pelo menos dois tipos celulares (células claras e escuras) foram observados na glândula acessória (Figura 14B). Várias microvilosidades foram observadas na região de contato da região proximal da glândula principal com o lúmen da glândula acessória (Figura 14D).

## **5.2 Análise funcional dos homogenados salivares de *Thyrsopelma guianense***

A quantificação de proteína presente nos homogenados salivares de fêmeas de *T. guianense* de dois dias de idade foi realizada para cinco grupos, cada um contendo 20 pares de glândulas salivares. Essa quantificação revelou uma grande quantidade de proteína por par de glândula ( $2,6 \mu\text{g}, \pm 0.2$ ) (Tabela 2). Os homogenados provenientes de cada grupo foram avaliados quanto à presença de anticoagulantes, sendo confirmada essa atividade por ensaios de tempo de recalcificação (Figura 15), tempo de tromboplastina parcial ativada (TTPa – Tabela 3) e tempo de protrombina (TP – Tabela 4).

O tempo de recalcificação do plasma humano na presença de diferentes concentrações de HGS demonstrou inibição dose-dependente (Figura 15). O aumento relativo do tempo de recalcificação em relação ao controle (dobro de tempo necessário em relação ao controle) foi observado quando  $1,14 \mu\text{g/ml}$  de concentração de HGS foi adicionado ao ensaio (Figura 15). A inibição máxima da coagulação foi registrada para a concentração de  $36,5 \mu\text{g/ml}$  (equivalente a  $3,1 \text{ HGS}$ ), representado pelo tempo máximo de observação (uma hora). Observou-se que  $0,4 \text{ HGS}$  (comparativamente o equivalente a um lobo salivar) foi capaz de aumentar 6 vezes o tempo de formação do coágulo (Figura 15).

Os ensaios de rotina TTPa e TP foram realizados para avaliados na ausência (controle) e presença de HGS (Tabela 3 e 4). Observou-se que  $48 \mu\text{g/mL}$  e  $31,5 \mu\text{g/mL}$  de HGS prolongaram o tempo de formação do coágulo de fibrina em relação ao controle com um incremento de seis e de vinte e sete segundos para os ensaios de TP e TTPa respectivamente (Tabela 3 e 4).

Os HGS foram avaliados quanto a presença de inibição da fXa e trombina, devido à participação de ambos os fatores na via comum da coagulação (Figuras 16 - 17). Os HGS de fêmeas de *T. guianense* revelaram inibição dose-dependente para o fXa (Figuras 16 - 17) e ausência de atividade de inibição para trombina (Figura 17).

Os HGS de fêmeas de dois dias de idade foram utilizados para verificar a presença e/ou ausência de atividade de lisozima. A atividade dessa enzima ocorreu principalmente em pH 7,0 com decréscimo para os demais pHs testados (6,0; 5,0; 8,0), respectivamente e sem atividade para o pH 9,0 (Figura 18). O HGS preparado com uma glândula salivar de fêmeas de diferentes idades foram avaliadas em pH 7,0 resultando em maior atividade para HGS de 48 h, com aproximadamente 1,2 mm de diâmetro para a zona de degradação vista no gel (Figura 18). Os HGS foram testados após serem submetidos a diferentes temperaturas (-70°C por 30 min, 56°C por 30 min, -70°C por 30 min) revelando os mesmos resultados para essa enzima (dados não demonstrados).

### 5.3 Análise do sialotranscriptoma

O sialotranscriptoma de *T. guianense* sequenciou 1.974 clones sendo 1.772 sequências de boa qualidade resultando em 752 grupos de sequências associadas das quais 491 dessas contendo apenas uma EST (etiquetas de sequências expressas). Os resultados do assemblamento das sequências, do BLAST e do peptídeo sinal foram utilizados para preparar o banco de dados do sialotranscriptoma de *T. guianense* (anexo I). Os genes expressos foram classificados em quatro categorias: produtos secretórios (S), celulares (C), produtos de função não conhecida (U) e elementos de transposição (Tabela 5 e Figura 19).

A categoria secretória (S) possui 56,9% dos grupos e 74,7% das sequências com um número médio de 3,1 sequências por grupo. A categoria celular (C) possui 22,9% e 16,2% dos grupos e sequências respectivamente e um número médio de 1,7 sequências por grupo. Uma sequência única foi classificada como elemento de transposição constituindo menos de 0,1% das ESTs ou grupos. Os produtos de função não conhecida (U) possuem 20,1% dos grupos e 9,0% de todas as sequências. Esse grupo recebeu esse nome pois nenhuma atribuição funcional pode ser atribuída para essas sequências. Essa categoria possui um número médio de 1,1 sequências por grupo e a maioria delas consiste de sequências únicas (Tabela 5 e Figura 19).

Os 172 grupos (288 ESTs) da categoria celular (C) foram classificados em 15 subgrupos funcionais (Tabela 7). Os dois subgrupos mais expressos foram a maquinaria de síntese de proteína (71 grupos com 113 ESTs – 39,2%) e o metabolismo energético (28 grupos com 33

ESTs – 11,4%). Nós arbitrariamente incluímos 93 ESTs (32,3%) dentro de um subgrupo da categoria C, devido representarem proteínas altamente conservadas de função não conhecida. A lista completa de todos os 288 grupos junto com as informações de cada um deles pode ser vista na (Tabela 7).

Os produtos secretórios foram classificados de acordo com a classificação funcional em: enzimas, domínios inibidores de protease, famílias de proteínas ubíquas, famílias específicas de insetos e famílias específicas de *Simulium* (essa última possui cerca de 63,6% de todos os produtos secretórios). Um total de 101 novas sequências das quais 82 codificam para proteínas secretoras foram agrupadas na tabela suplementar 1 e 2 do anexo I. A análise do proteoma de *T. guianense* permitiu confirmar 28 das 32 famílias de proteínas encontradas nesse sialotranscriptoma (Tabela 8 e Figura 20). Entre as famílias de proteínas encontradas no sialotranscriptomas de *T. guianense* estão:

### **5.3.1 Enzimas**

Vários transcritos presentes no sialotranscriptoma de *T. guianense* codificam proteínas com sequências similares para enzimas secretadas, como: glicosidases, serino proteases, hialuronidases, apirase, adenosina deaminase (ADA) e destabilase. Essas enzimas podem estar associadas com a alimentação sanguínea, de açúcar ou ambos.

#### **5.3.1.1 Glicosidases**

O sialotranscriptoma de *T. guianense* apresentou dois genes incompletos que codificam para glicosidases (Sg-214 e Sg-296). As glicosidases de *T. guianense* possuem sequências com 79% de identidade para outras glicosidases de simulídeos. A análise filogenética das glicosidases de *T. guianense* com seus homólogos mais próximos encontrados no banco de dados de proteínas do NCBI agrupam as duas proteínas de *T. guianense* em diferentes clados com forte suporte parcimônico (Figura 21). O grupo Sg-296 agrupa-se com as enzimas de *Simulium*, com uma enzima de flebotomíneo e com outras de drosófilas como indicado no clado I (Figura 21). Por outro lado, o grupo Sg-214 agrupa-se com um segundo grupo de enzimas de *Simulium* e com 76% de suporte parcimônico para mosquitos Culicine e enzimas de ceratopogonídeos como demonstrado no clado II (Figura 21). Cada grupo de sequências de mosquitos, simulídeos e *Culicoides* agrupam-se dentro de subclados específicos com forte suporte parcimônico. Um terceiro clado de enzimas somente de mosquitos (incluindo anofelinos e culicíneos) emergiu sem forte suporte parcimônico no clado II. As sequências da proteína glicosidase (Sg-214 e Sg-296) encontradas no sialotranscriptoma de *T. guianense* foram confirmadas pela análise proteômica

dentro das frações 16 e 17, respectivamente, um pouco acima do padrão de 49 kDa (Tabela 8 e Figura 20).

#### 5.3.1.2 Serino proteases

O sialotranscriptoma de *T. guianense* permitiu a identificação de transcritos que codificam para três serino proteases secretadas que variam entre 25,4 para 27 kDa, que pode ser derivada de três genes polimórficos. O alinhamento de membros representativos desses 3 produtos gênicos de *T. guianense* (Sg-416, Sg-138 e Sg-244) com seus homólogos mais próximos (somente sequências de Diptera foram incluídas) produziram um filograma indicativo de um clado (clado I) que tem relação e suporte parcimonônico forte somente com sequências de mosquitos e drosófilas (Figure 22) e um clado adicional específico para *Simulium* que contém dois sub-clados. Cada sub-clado marcado como *Simulium* I, II e III (Figura 22) contém uma enzima de cada espécie de Simuliidae conhecida. Uma sequência adicional de *S. nigrimanum* também foi encontrada nesse clado específico de *Simulium* indicando que um quarto gene pode ser expresso nesse inseto. As proteínas serino proteases encontradas no sialotranscriptoma de *T. guianense* foram confirmadas pela análise proteômica nas frações F24 e F25 localizada próximo do padrão de 28 kDa sendo consistente com o peso maduro predito dessas proteínas (25 para 27 kDa) (Tabela 8 e Figura 20).

#### 5.3.1.3 Hialuronidases

O sialotranscriptoma de *T. guianense* revelou uma sequência completa (com 5 ESTs) codificando para uma proteína com 37,8 MW e pI 9,2 reconhecendo o domínio pfam01630 denominado “Glyco\_hydro\_56, Hyaluronidase” com um e-value de  $1e^{-61}$ . Quatorze peptídeos tripticos obtidos por MS/MS tiveram similaridade para a proteína hialuronidase dentro da fração 20, um pouco acima do padrão de 38 kDa sendo consistente com o peso maduro predito para essa proteína (Tabela 8 e Figura 20). O alinhamento da hialuronidase salivar de *T. guianense* revelou poucos aminoácidos conservados e o filograma não revelou história evolutiva da hialuronidase salivar de fêmea de *T. guianense* com nenhum homólogo (Figura 23).

#### 5.3.1.4 Apirase

Quatorze das 1.324 ESTs codificam para apirases salivares no sialotranscriptoma de *T. guianense* (tabela suplementar 1 e 2 – anexo I). O alinhamento das apirases de *T. guianense* com as sequências de seus homólogos simulídeos e de vertebrados indicam a ausência do domínio carboxiterminal (Figura 24) sugerindo que essas enzimas em simulídeos são secretadas.

Cinquenta e oito peptídeos trípticos foram deduzidos por MS/MS com similaridade para a proteína apirase (Sg-354) originada dentro da fração 16, localizada um pouco abaixo do padrão de 62 kDa (Tabela 8 e Figura 20). O filograma não revelou história evolutiva das apirases salivares de fêmea de *T. guianense* com nenhum homólogo (Figura 24).

#### **5.3.1.5 Adenosina deaminase e Destabilase**

Uma única sequência incompleta de adenosina deaminase foi encontrada no sialotranscriptoma de *T. guianense*. Uma única sequência incompleta com 86% de identidade para a destabilase de *S. nigrimanum* foi encontrada no sialotranscriptoma de *T. guianense*.

### **5.3.2 Domínios inibidores de proteases**

#### **5.3.2.1 Serpinas**

Uma única sequência incompleta de serpina (Serino Protease Inhibitor) foi encontrada no sialotranscriptoma de *T. guianense* (Sg-500) com 60% identidade para um inibidor de serino protease dos homólogos *A. aegypti* e *Anopheles gambiae*.

#### **5.3.2.2 Inibidor de protease com domínio Kunitz**

O sialotranscriptoma de *T. guianense* contém uma típica proteína Kunitz deduzida de 3 ESTs (Sg-395). Essa proteína possui similaridade blastp para seu homólogo *S. nigrimanum* (72% de identidade) e para proteínas de *S. vittatum* (58%). Dois peptídeos trípticos com similaridade para a proteína com domínio Kunitz (Sg-395) foram obtidos por MS/MS dentro da fração 30, localizada abaixo do padrão de 6 kDa (Tabela 8 e Figura 20). O alinhamento da proteína Sg-395 com seus homólogos revelou poucos aminoácidos conservados em sua sequência e o filograma não revelou história evolutiva das proteínas salivares com domínio Kunitz de fêmea de *T. guianense* com nenhum homólogo (Figura 25).

### **5.3.3 Famílias de proteínas ubíquas**

#### **5.3.3.1 Produtos relacionados a imunidade**

Nesse grupo as sequências completas codificam para dois peptídeos antimicrobianos ubíquos: lisozima e cecropina. O sialotranscriptoma de *T. guianense* revelou 25 ESTs codificando para membros da família lisozima com vários possíveis alelos do mesmo gene. A

lisozima salivar de *T. guianense* possui 79% de identidade para a lisozima de *S. nigrimanum* e 52% para mosquitos (Tabela suplementar 1 e 2 – Anexo I). Sete peptídeos trípticos foram deduzidos por MS/MS com similaridade para a proteína lisozima (Sg-263) dentro da fração 30, localizada um pouco abaixo do marcador de 6 kDa (Tabela 8 e Figura 20). O alinhamento demonstrou vários aminoácidos conservados na sequência das lisozimas salivares de fêmeas de *T. guianense* com seus homólogos e o filograma revelou subclados específicos para as lisozimas de *Simulium* e mosquitos sem história evolutiva entre esses dípteros (Figura 26).

As cecropinas foram identificadas nesse sialotranscriptoma como pequenas proteínas básicas de 3kDa, secretadas e ricas em aminoácidos alifáticos, principalmente VAL com alto nível de conservação na região carboxiterminal. Dez ESTs codificam para duas proteínas muito próximas, possivelmente alelos e com similaridades para outras cecropinas de *Simulium* e mosquitos com aminoácidos altamente conservados visto em suas sequências no alinhamento com filograma revelando clados específicos para as cecropinas de mosquitos e *Simulium* (Figura 27). Os peptídeos Sg-368 e Sg-369 foram confirmados no proteoma dentro da fração F31, localizada acima do padrão de 3 kDa, consistente com o peso maduro dessa proteína (Tabela 8 e Figura 20).

#### 5.3.3.2 Família Antígeno-5

O sialotranscriptoma de *T. guianense* revelou dois grupos codificando membros da superfamília CAP. O filograma resultante do alinhamento das duas proteínas de *T. guianense* com seus 25 melhores homólogos do banco de dados de proteínas não-redundantes do NCBI (excluindo proteínas de drosófilas) revelaram um clado de Nematocera com suporte parcimônico de 70% e com suporte parcimônico forte para os sub-clados específicos das famílias de Nematocera (marcados como Salivares na Figura 28). Uma segunda proteína (Sg-453) de *T. guianense* esta localizada fora do clado salivar. Trinta peptídeos trípticos foram deduzidos por MS/MS com similaridade para uma proteína Antígeno-5 (Sg-457) dentro da fração 24, localizada um pouco acima do marcador de 28 kDa, próximo dos 30kDa dessa proteína madura (Tabela 8 e Figura 20).

#### 5.3.3.3 Família Yellow

O sialotranscriptoma de *T. guianense* revelou somente uma EST para proteína Yellow com 64% de identidade para *C. quinquesfasciatus*.

#### **5.3.3.4 Família com domínio ML**

Um transcrito contendo domínios ML foi encontrado em *T. guianense*. O blast para o banco de dados de proteína não redundantes do NCBI sugere similaridade para Niemann-Pick Type C-2 putative de *A. aegypti* (também conhecido como proteína secretora epididimal E1) e similaridades para o alérgeno MPA2 de *Nasonia vitripennis* (Himenoptera).

#### **5.3.3.5 Lipocalinas**

Aqui, nós encontramos duas ESTs codificando para diferentes lipocalinas (Sg-671 e Sg-568). A sequência de aminoácido deduzida de ambos os transcritos tem similaridade para alipoproteína humana do banco de dados do Swissprot.

### **5.3.4 Famílias de proteínas específicas de insetos**

#### **5.3.4.1 Família Aegyptina**

O sialotranscriptoma de *T. guianense* revelou 23 ESTs da classe S codificando para o que parece ser alelos de um único gene e similar para proteínas como Aegyptina com 60% identidade para Aegyptinas de sialotranscriptomas de simulídeos e 35% para os homólogos mosquitos. O alinhamento revelou regiões ricas em GE principalmente no meio da sequência das proteínas e o filograma demonstrou as sequências de *Simulium* e mosquitos em clados específicos (Figura 29). Trinta e sete peptídeos trípticos foram obtidos por MS/MS com similaridades para a proteína Aegyptina (Sg-276) dentro da fração 24, localizada um pouco acima do marcador de 28 kDa, consistente com o peso maduro predito dessa proteína (28 kDa) (Tabela 8 e Figura 20).

#### **5.3.4.2 Família de proteína secretada em Diptera e conservada em insetos**

Cinco proteínas encontradas no sialotranscriptoma de *T. guianense* tiveram similaridade para proteínas secretadas de insetos. O tamanho dessas proteínas variam de 37 para 57 kDa de peso maduro e ponto isoelétrico variando de 6,1 para 9,6. Essas proteínas possuem altas similaridades para várias famílias de Diptera (Culicidae, Ceratopogonidae, Drosophilidae), Himenoptera (Pteromalidae e Formicidae), porém apresentam baixo nível de aminoácidos conservados com filograma sugerindo pelo menos quatro diferentes genes em *T. guianense* (Figura 30). Vários peptídeos trípticos foram obtidos por MS/MS com similaridades para proteínas secretadas em Diptera e conservada em insetos nas frações 21, 17, 14, e 13 (Tabela 8 e

Figura 20). O padrão inicial do Psi blast (PHI-blast) é G-x-[MLI]-x(6)-[WF]-x(7,12)-[KNE]-x-[IMFL]-x(37,40)-[VIL]-x-[YF]-x(3)-[QKR]-x(14)-[IL]-x(5,6)-[NDE]-x(5)-[ILV]-[AS].

#### 5.3.4.3 Família de proteína laminina

O sialotranscriptoma de *T. guianense* revelou uma proteína incompleta (Sg-431) com quatro ESTs codificando para o homólogo *S. vittatum* (76% identidade) e para uma proteína de *Culicoides nubeculosus* com 32% identidade. Dois peptídeos trípticos foram obtidos por MS/MS tendo similaridade para a proteína laminina dentro da fração 14, um pouco acima do padrão de 62 kDa (Tabela 8 e Figura 20).

#### 5.3.4.4 Superfamília D7/OBP

A família de proteína D7 são reconhecidas como membros da superfamília OBP e são proteínas que se ligam a odorantes. As sequências OBP/D7 de *T. guianense* foram agrupadas em 3 subfamílias conforme descritas abaixo:

##### 5.3.4.4.1 Família D7 Longa

Duas proteínas com dois domínios OBP foram reconhecidas no sialotranscriptoma de *T. guianense* com similaridades significantes somente para outras proteínas de *Simulium*. As D7 longas de *Simulium* revelaram poucos aminoácidos conservados no alinhamento e clados distintos no filograma para sequências de mosquitos e flebotomíneos (Figura 31). As proteínas D7 longas tiveram peptídeos trípticos deduzidos por MS/MS dentro da fração F23 e F26 próximo do marcador de 28 kDa, consistente com o peso maduro predito dessas proteínas (28 kDa e 24 kDa) (Tabela 8 e Figura 20).

##### 5.3.4.4.2 Família D7 16 kDa

Em *T. guianense* foram encontradas duas proteínas contendo um domínio OBP. O grupo Sg-331 produz similaridades significantes somente para outras proteínas de *Simulium*, enquanto o Sg-350 adicionalmente recupera o domínio OBP de *C. quinquefasciatus*. O alinhamento das D7 16 kDa com seus homólogos *S. nigrimanum* e *S. vittatum* não revelam história evolutiva dessas proteínas (Figura 32). Peptídeos trípticos foram deduzidos por MS/MS com similaridades para proteínas D7 16 kDa nas frações 29 e 30, um pouco acima de 6 kDa (Tabela 8 e Figura 20).



#### 5.3.4.4.3 Família D7 ultra-curtas (10-12 kDa)

Essa família foi a mais expressa dentro da superfamília D7/OBP com 80 ESTs (Tabela suplementar 1 e 2). Todos os grupos contém peptídeo sinal sugerindo serem proteínas secretadas. Essas informações sugerem a existência de pelo menos quatro genes codificando para proteínas D7 ultra-curtas (Sg-75, Sg-190, Sg-363 e Sg-383) e vários possíveis alelos (Figura 33). Somente Sg-383 foi deduzido por MS/MS dentro da fração 31, um pouco abaixo de 6 kDa (Tabela 8 e Figura 20).

#### 5.3.5 Famílias específicas de *Simulium*

Oitocentos e quarenta e três ESTs das 1.324 ESTs da classe S foram específicas para *Simulium*. Algumas famílias de *S. nigrimanum* foram deorfanizadas e uma nova família que codifica para uma proteína do tipo-kunitoxina foi pela primeira vez encontrada em sialotranscriptoma de insetos. Mais detalhes de cada família foram descritos abaixo.

##### 5.3.5.1 Família SVEP/Marydilan

O sialotranscriptoma de *T. guianense* revelou 190 ESTs com 50 para 70% de identidade para outras SVEP's de *Simulium*. O alinhamento de membros dessa família demonstrou sequências com tamanhos similares mas com poucos aminoácidos conservados. Análises filogenéticas comparativas de todas as proteínas SVEPs depois de 10.000 análises parcimônicas agruparam a maioria dos membros de *S. vittatum* em um clado específico (I) com 80% de suporte parcimônico (Figura 34). A proteína Sg-13 compartilha o clado II com 92% de suporte parcimônico com proteínas de seu homólogo *S. nigrimanum*. As EST's que codificam para Sg-13 ou proteínas muito próximas representam mais do que 50% das sequências codificadoras para membros SVEP nesse sialotranscriptoma. O clado III agrupa somente SVEP de *S. vittatum* e *S. nigrimanum* sem suporte parcimônico. O clado IV revela um possível caso de duplicações gênicas, ou a expressão de um gene muito polimórfico de *T. guianense* (ou ambos). O clado (V) agrupa dois grupos (Sg-344 e Sg-343) de *T. guianense* com seus homólogos *S. nigrimanum* (com 100% suporte parcimônico) que aparece completamente distinto de outras proteínas SVEP. Dessa forma, o filograma indica pelo menos três genes e várias outras recentes duplicações gênicas e/ou alelos que codificam para membros dessa família (clado IV Figura 34). Todas as sequências de proteínas SVEP encontradas no sialotranscriptoma de *T. guianense* foram confirmadas pela análise proteômica dentro da fração 30 localizada um pouco acima do marcador de 14kDa, consistente com o peso maduro predito dessa proteína (14 kDa) (Tabela 8 e Figura 20).

### 5.3.5.2 Família de proteínas ácidas de *Simulium* ricas em H

O sialotranscriptoma de *T. guianense* revelou 9,6% de todas as suas ESTs secretoras codificando para membros dessa família e acima de 70% de identidade para proteínas de seus homólogos *S. nigrimanum*. O alinhamento demonstra que as sequências de *T. guianense* juntas com as proteínas de seus homólogos *S. nigrimanum* contém uma região repetida codificando para resíduos de Pro-Lys-Pro onde o resíduo Lys em *S. vittatum* é substituído por Gln (Figura 35). O filograma dessa família (Figura 35) quando adicionado sequências de mosquitos e *Culicoides* revelam que todas as sequências de *Simulium* tem um ancestral comum com 93% de suporte parcimonioso.

### 5.3.5.3 Mucinas

Dois tipos específicos de mucinas de *Simulium* foram encontradas no sialotranscriptoma de *T. guianense*:

#### 5.3.5.3.1 Mucina de *Simulium*

Nove proteínas (com 40 ESTs) codificam para mucina de *Simulium*. Suas sequências tem altas quantidades de resíduos de Ser+Thr variando de 34,6 para 42,6% e de 40 para 144 locais de galactosilação preditos pelo servidor NetOglyc [12]. O alinhamento das sequências de *T. guianense* e *S. nigrimanum* revelaram extensivas similaridades e identidades ao longo de toda a sequência, mas o filograma claramente distingue clados específicos para *T. guianense* e *S. nigrimanum* (Figura 36). Vários peptídeos tripticos foram deduzidos por MS/MS com similaridade para mucina de *Simulium* dentro da fração 16 (localizada abaixo do marcador de 62 kDa) e da fração 9 (um pouco acima do marcador de 96 kDa) (Tabela 8 e Figura 20).

#### 5.3.5.3.2 Mucinas ácidas similar a família básica de 7-13kDa de *Simulium*

O sialotranscriptoma de *T. guianense* revelou 13 proteínas (com 39 ESTs) com melhores similaridades para membros dessa família, de ambos *S. vittatum* e *S. nigrimanum*, com 50% de identidade. Entretanto, essas sequências codificam para proteínas ácidas e não básicas (pI 4,1 para 4,4) com 0 para 19 locais potenciais de galactosilação. As proteínas de *S. guianense* apresentam um domínio central estendido contendo repetições de Gly-Ser que variam em tamanho entre as proteínas (Figura 37). O filograma produz ramos monoespecíficos indicativos de um único gene polimórfico, ou, alternativamente, de múltiplos genes (Figure 37).

#### **5.3.5.4 Família de proteína do tipo-colágeno de *Simulium***

Esta família revelou 17 grupos contendo 48 ESTs (Tabela Suplementar 1 e 2 – Anexo I). Esses 17 grupos são variantes de três sequências que foram alinhadas com seus homólogos *Simulium* (tendo ~ 40% identidade) para produzir a árvore filogenética demonstrada na Figura 38. O alinhamento revelou aminoácidos relativamente conservados no comprimento da sequência, com alguns espaços devido as longas sequências de *S. vittatum*. O filograma revelou clados monoespecíficos (Figura 38). Vários peptídeos trípticos foram deduzidos por MS/MS dentro da fração 23, um pouco acima do marcador de 28 kDa, e acima do peso maduro predito dessas proteínas (22 kDa) (Tabela 8 e Figura 20).

#### **5.3.5.5 Família Sv 7,8 kDa**

As sequências de *T. guianense* mantém 60% de identidade para proteínas de seus homólogos *S. nigrimanum* e 50% para proteínas de *S. vittatum*, codificando para proteínas básicas (pI 9 para 10,4) com peso maduro variando de 7,1 para 13,4 kDa. O alinhamento revelou baixos níveis de aminoácidos conservados e pelo menos quatro genes para proteínas dessa família em *T. guianense*, marcadas nos clados I-IV da Figura 39. Vários peptídeos trípticos foram deduzidos por MS/MS dentro das frações F27, F30, F32 e F31. Essas frações estão localizadas no gel um pouco acima do marcador de 14 kDa e um pouco acima do marcador de 3 kDa. Esses resultados são consistentes com o peso maduro predito dessas proteínas (7 para 13 kDa) (Tabela 8 e Figura 20).

#### **5.3.5.6 *Simulium* Básica 7-13kDa**

O sialotranscriptoma de *T. guianense* revelou duas proteínas (Sg-420 e Sg-403) com 6 ESTs para essa família codificando proteínas básicas (pI 11) e 8 kDa MW com 59% de similaridade para as proteínas de seus homólogos *S. nigrimanum* e *S. vittatum*, com poucos aminoácidos conservados visto no alinhamento e sem história evolutiva entre as sequências (Figura 40). Peptídeos trípticos foram encontrados por MS/MS nas frações 30 e 31, um pouco abaixo do marcador de 6 kDa (Tabela 8 e Figura 20).

#### **5.3.5.7 Família *Simulium* 4,8kDa**

Cinco transcritos foram adicionados para essa família com sequências que codificam para proteínas ácidas (pI 4.4) com 5 kDa de peso maduro e desprovido de cisteínas. São proteínas

altamente conservadas quando alinhada com seus homólogos e que produz um filograma com subclados específicos para cada espécie de *Simulium* (Figura 41).

#### **5.3.5.8 Família *Simulium* Básica 7,4kDa**

O grupo Sg-422 (com 4 ESTs) codificam para um peptídeo básico de 7 kDa MW e com identidade acima de 50% para as proteínas de seus homólogos *S. vittatum* e *S. nigrimanum*. Essa família de proteína não apresenta nenhum resíduo de cisteína no peptídeo maduro. Quatro peptídeos trípticos originados de Sg-422 foram deduzidos por MS/MS na fração 31, um pouco abaixo do padrão de 6 kDa, consistente com o peso maduro predito dessa proteína (6,8 kDa) (Tabela 8 e Figura 20).

#### **5.3.5.9 *Simulium* Básica 13 kDa**

O grupo Sg-446 possui três ESTs que codificam para essa família com ~ 40% de identidade para outras proteínas de *Simulium*. Elas não possuem similaridades para outras proteínas conhecidas e para nenhum banco de dados de NR, GO, KOG, CDD, PFAM, SMART. Dois peptídeos trípticos foram deduzidos por MS/MS na fração 29, um pouco abaixo do padrão de 14 kDa consistente com o peso maduro predito dessa proteína (13,7 kDa) (Tabela 8 e Figura 20).

#### **5.3.5.10 Família 5-Cys *Simulium***

O sialotranscriptoma de *T. guianense* possui nove ESTs codificando para essa família de proteína. Essas proteínas contêm 5 cisteínas nas suas sequências e revelam alta quantidade de aminoácidos conservados quando alinhada com seus homólogos (Figura 42). Elas se apresentam como proteínas ácidas (pI 5,7) com 14 kDa MW e acima de 60% identidade para homólogos *Simulium*. O filograma revelou as proteínas de *T. guianense* mais próximas filogeneticamente com as proteínas de *S. nigrimanum*. Três peptídeos trípticos foram deduzidos por MS/MS na fração F30, um pouco abaixo do marcador de 6 kDa (Tabela 8 e Figura 20).

#### **5.3.5.11 Família 8-10 Cys W**

O sialotranscriptoma de *T. guianense* possui duas proteínas (11 ESTs) com 9 Cys e 5 ou 6 Trp. O filograma revelou dois clados, um contendo somente proteínas de *S. nigrimanum* e um segundo contendo além de *S. nigrimanum* as proteínas de *T. guianense* (Figura 43). Essas

proteínas foram confirmadas por MS/MS na fração 29, um pouco abaixo do padrão de 14 kDa, próximo do peso maduro predito dessa proteína (16,9 kDa) (Tabela 8 e Figura 20).

#### **5.3.5.12 Família ácida de 28kDa**

O sialotranscriptoma de *T. guianense* revelou três proteínas (Sg-319, Sg-320 e Sg-321) com 12 ESTs para essa família. A corrida NuPAGE MS/MS reconheceu muitos peptídeos trípticos para essas proteínas na fração F23, um pouco acima do marcador de 28 kDa (Tabela 8 e Figura 20), de acordo com o peso maduro predito dessas proteínas (22 para 27 kDa). O alinhamento revelou aminoácidos mais conservados no meio das sequências e produziu clados específicos para *S. nigrimanum* e *T. guianense* (Figura 44).

#### **5.3.5.13 Família básica de 28kDa**

Essa família apresentou um grupo com 9 ESTs (Sg-136) que codificam para proteínas básicas de 28kDa. Essa proteína (Sg-136) apresentou vários peptídeos trípticos deduzidos por MS/MS na fração 24, consistente com a massa predita dessa proteína (Tabela 8 e Figura 20). O alinhamento revelou aminoácidos mais conservados no final das sequências e produziu clados específicos para *S. nigrimanum* e *T. guianense* (Figura 45).

#### **5.3.5.14 Família 19kDa**

Duas proteínas (Sg-303 e Sg-309) com 10 ESTs codificam para proteínas básicas de 16,8 MW e peptídeo sinal nas suas sequências sugerindo serem proteínas secretadas. Trinta peptídeos trípticos foram encontradas por MS/MS na fração 27, localizada um pouco acima do padrão de 14 kDa e consistente com o peso maduro predito dessas proteínas (16,8 kDa) (Tabela 8 e Figura 20). O alinhamento revelou aminoácidos altamente conservados e clados específicos para *S. nigrimanum* e *T. guianense* (Figura 46).

#### **5.3.5.15 Família de proteína básica de 8kDa**

Essa família revelou peptídeos pequenos como o grupo Sg-258 (22 ETS) codificando para proteínas básicas de 8 kDa MW com 70% de identidade para uma proteína anteriormente descrita no sialotranscriptoma de *S. nigrimanum*. Cinco peptídeos trípticos foram deduzidos por MS/MS na fração 31, localizada um pouco abaixo do marcador de 6 kDa (Tabela 8 e Figura 20).

#### **5.3.5.16 Família Sn básica de 4,4kDa**

Um grupo (Sg-438) foi encontrado com duas ESTs codificando para um pequeno peptídeo de 4,4kDa com similaridade para seus homólogos *S. nigrimanum*.

#### **5.3.5.17 Novo peptídeo similar para kunitoxina**

O sialotranscriptoma de *T. guianense* apresentou dois grupos (Sg-375 e Sg-409) codificando para um novo peptídeo similar (32% identidade) para o peptídeo kunitoxina de cobra [103]. Essa proteína é rica em Cys mas não contém domínio Kunitz. Esses dois grupos juntos agrupam 8 ESTs codificando para um peptídeo básico secretado com 8 para 9kDa. A corrida PAGE/MS/MS relatou quatro peptídeos tripticos para essa proteína dentro da fração 31, coincidentemente com a banda mais bem corada entre os padrões de 3 a 6 kDa (Tabela 8 e Figura 20).

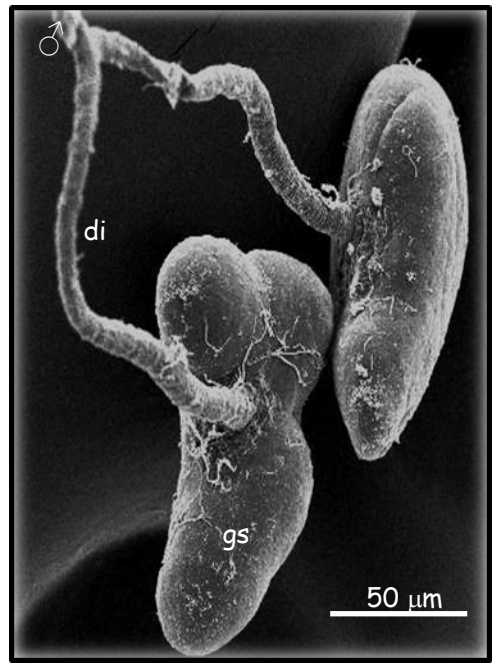
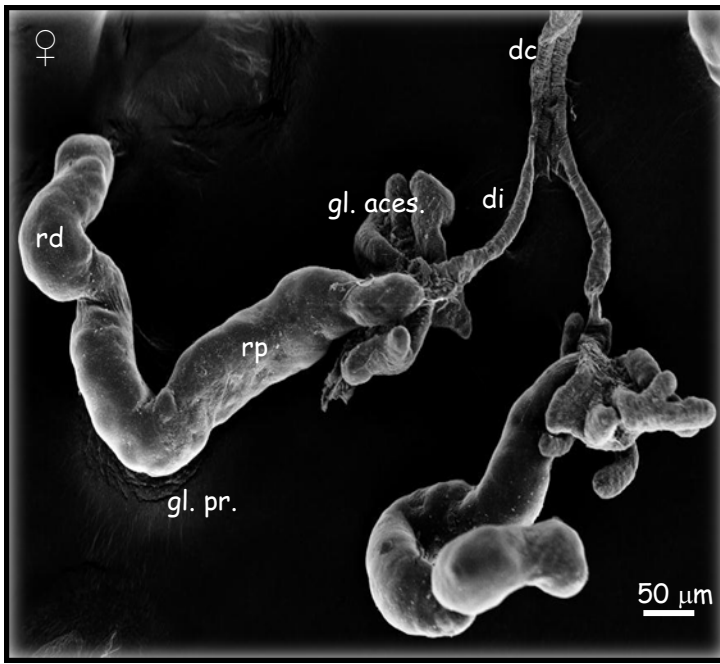
### **5.4 Trabalho de campo**

A cachoeira do Travessão, localizada no rio Jauaperi, apresenta fundo rochoso sendo endêmico para plantas aquáticas da família Podostemaceae. As fases imaturas (ovos, larvas e pupas) foram encontrados aderidas nos ramos da Podostemaceae, principalmente na plantas localizadas em áreas de maior velocidade de água corrente e oxigênio disponível. Duas espécies foram encontradas compartilhando esses criadouros (*T. guianense* e *P. iracouboense*). A grande quantidade de imaturos observada nos criadouros é totalmente desigual a quantidade de adultos que picam próximo ao criadouro. Os imaturos de simulídeos foram coletados durante a subida das águas do rio Jauaperi porque nesse período as podostemaceae expõe a parte vegetativa que se desenvolve debaixo da água (onde larvas e pupas utilizam como substrato), enquanto na vazante as florações aparecem aderidas na plataforma rochosa como finas hastes florais.

Somente substratos contendo pupas maduras de simulídeos foram removidas do criadouro, acondicionadas e transportadas em sacos plásticos até o laboratório. As sacolas de emergência continham pequenos furos para evitar a condensação na parede e consequente adesão e morte dos adultos nas gotas de água durante a emergência. Observamos um pico de emergência decrescente do primeiro ao terceiro dia após a coleta, com uma natalidade de 1:1 de machos e fêmeas. Uma forte atração pela luz foi observada durante e após a emergência. Das duas espécies que emergiram das pupas, *P. iracouboense* e *T. guianense*, somente a segunda foi estudada nesse trabalho. As fêmeas de *T. guianense* são antropofílicas e suas picadas, no campo, ocorreram

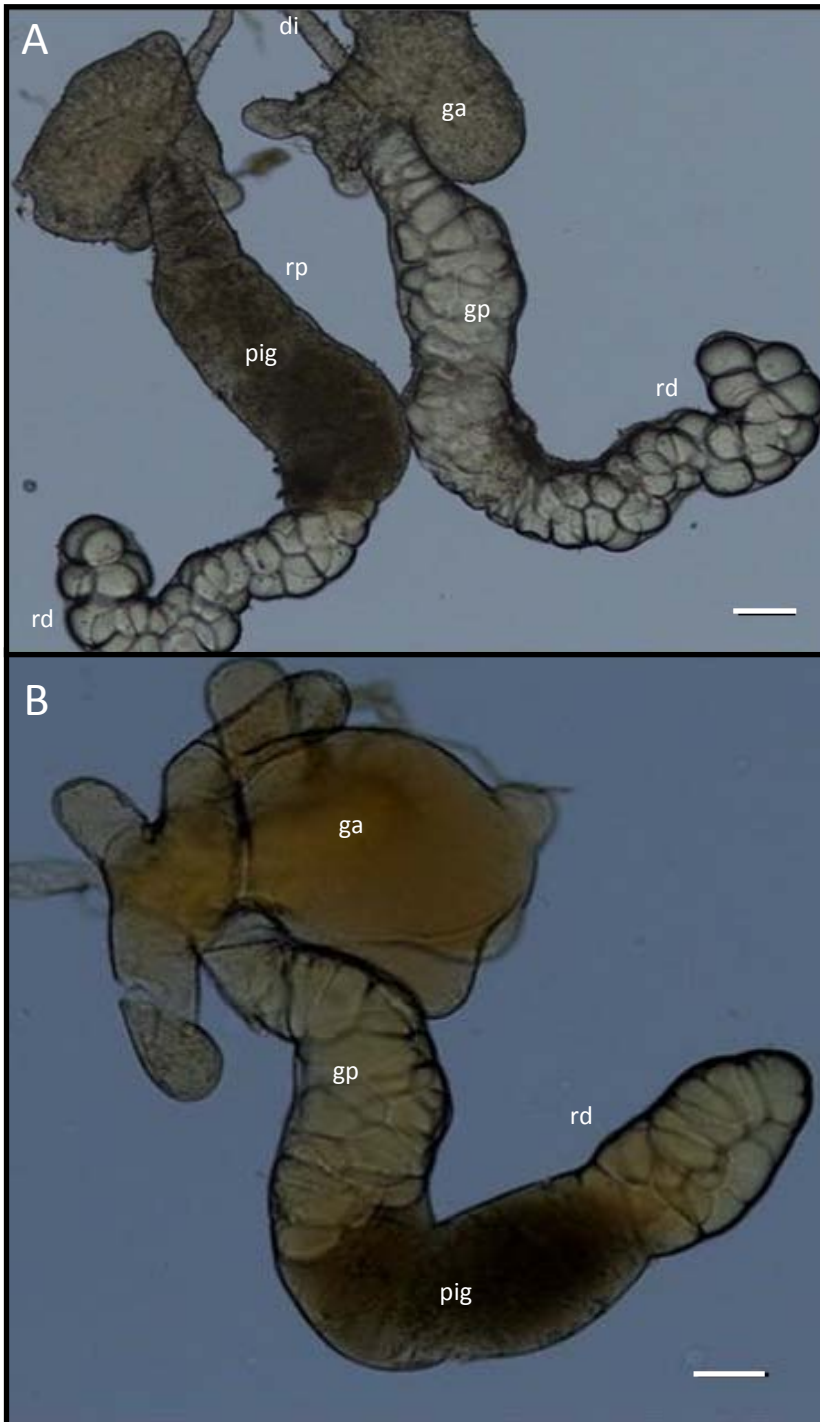
quase que exclusivamente nos membros inferiores. As picadas causam eritema, exantema, prurido, inchaço e forte anestesia (observação pessoal).

A alimentação por açúcar foi imediatamente procurada pela fêmea após a emergência, enquanto a alimentação por fonte sanguínea ocorreu a partir do segundo para terceiro dia após a emergência, com preferência para humanos e com um tempo médio de repasto sanguíneo completo variando de cinco para 12 minutos. Após alimentação sanguínea observamos a excreção de gotas de fluido claro na parede dos tubos aonde as fêmeas foram acondicionadas. Essa rápida eliminação de fluidos deve ser essencial para reduzir o peso e facilitar o vôo da fêmea alimentada.





**Figura 5** - Dimorfismo sexual em glândula salivar de *T. guianense* (fêmeas e machos) observadas em microscopia eletrônica de varredura (MEV). Glândula salivar de fêmea (Figura 5A) e a glândula salivar de macho (Figura 5B). Abreviações: C: constrição; d: ducto individual; dc: ducto comum; ga: glândula acessória; gms: glândula salivar de macho; gp: glândula principal; rd: região distal e rp: região proximal.

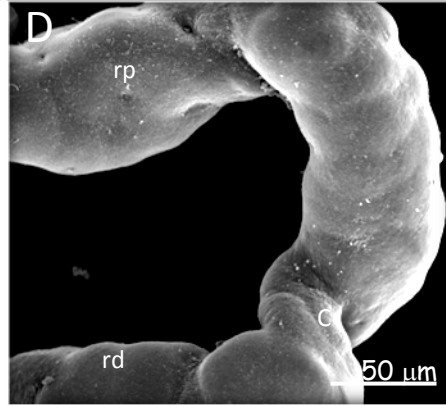
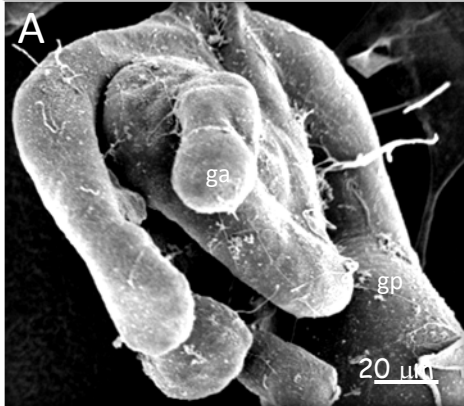


**Figura 6** - Glândula salivar de fêmea de *T. guianense* com dois dias de idade com intensa área pigmentada (pig) em diferentes regiões da glândula. As imagens foram obtidas no microscópio óptico de contraste de fase. As barras indicam 50 µm. Abreviações: di: ducto individual; ga: glândula acessória; gp: glândula principal; rd: região distal; rp: região proximal; pig: pigmentação. Na fígura 6A pode ser visto toda a glândula e em 6B somente um lado da glândula.

glândula acessória

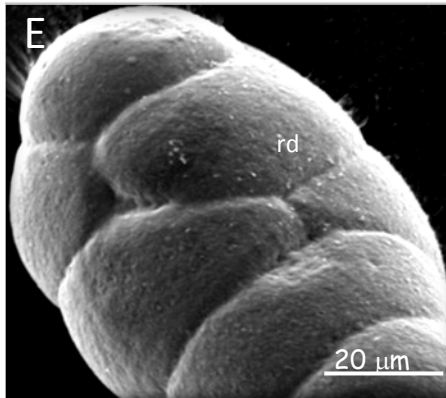
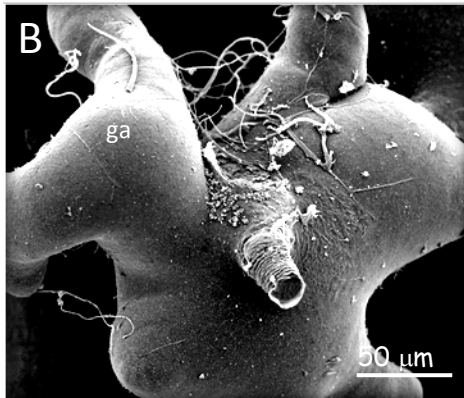
glândula principal

24h



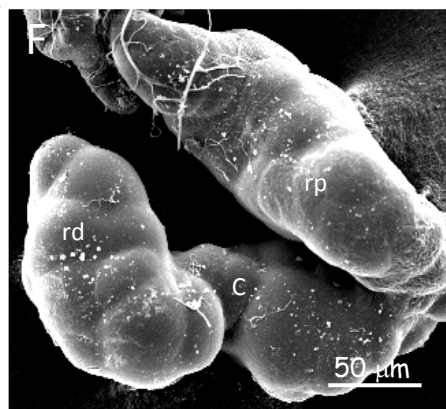
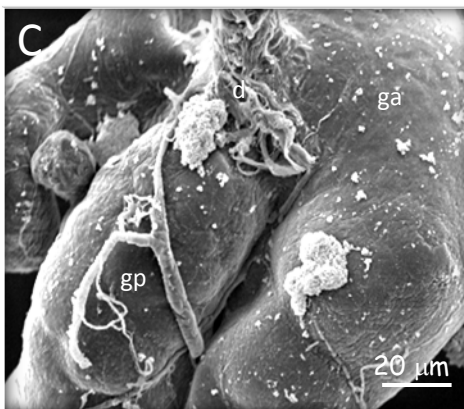
24h

48h



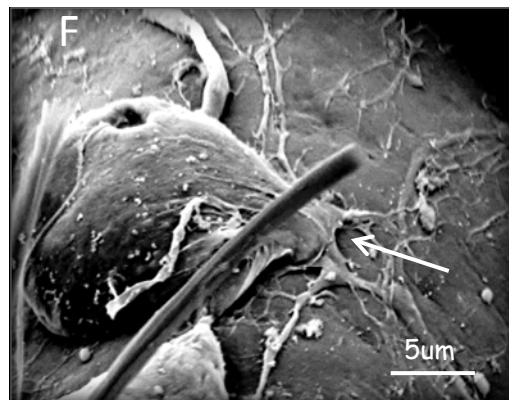
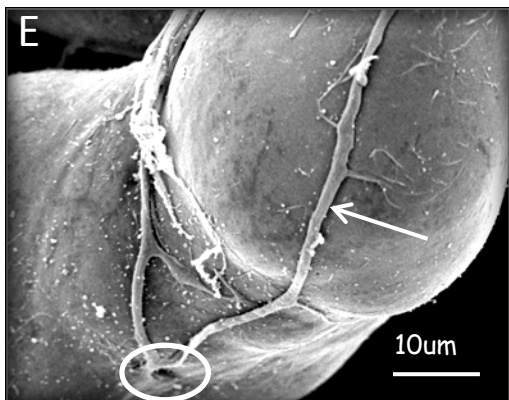
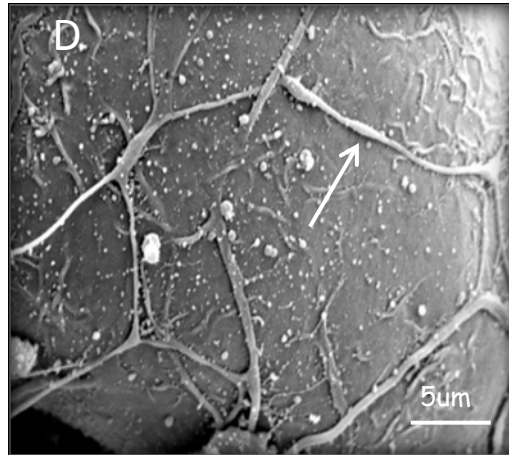
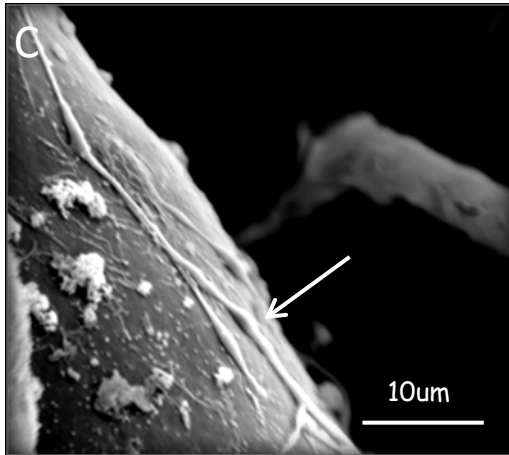
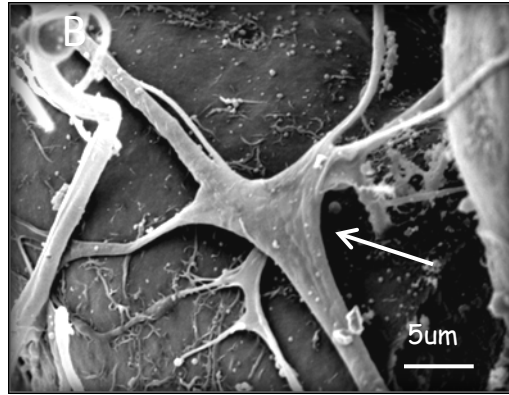
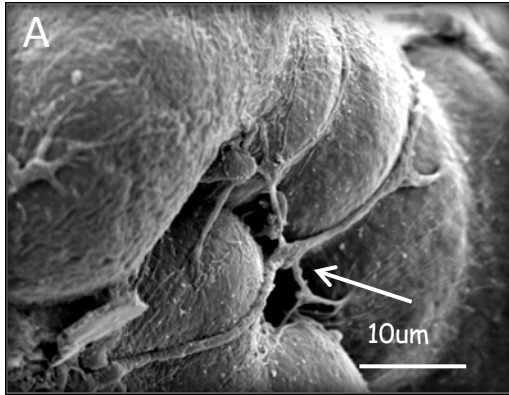
48h

72h

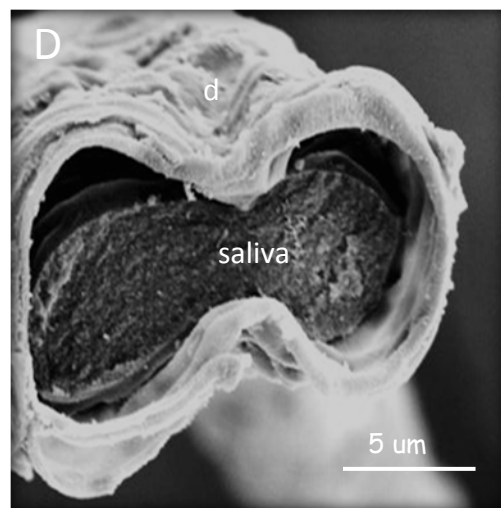
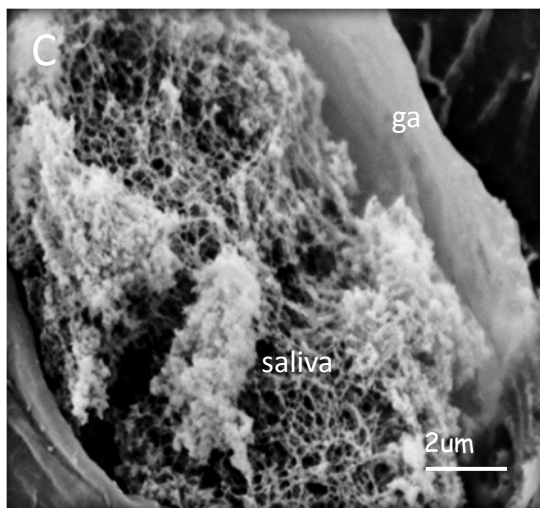
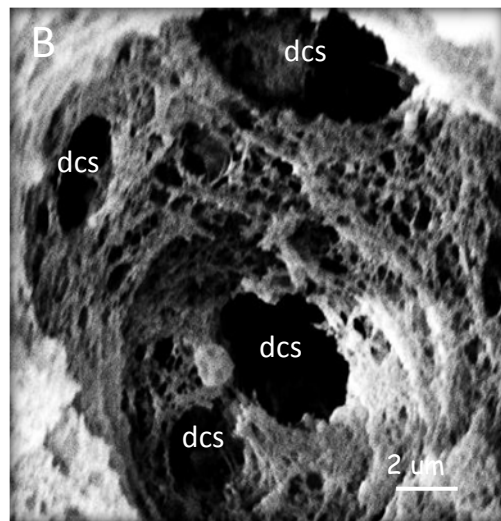
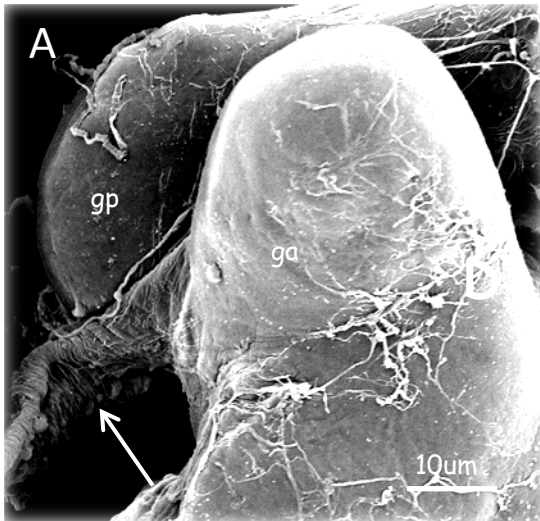


72h

**Figura 7** - Topografias da glândula salivar de fêmeas de *T. guianense* em diferentes idades de envelhecimento glandular obtidas em microscópio eletrônico de varredura (MEV). A - C: Glândula acessória com 24, 48 e 72h respectivamente e D-F: Glândula principal com 24, 48 e 72h respectivamente. Abreviações: C: constrição; d: ducto; ga: glândula acessória; gp: glândula principal; rd: região distal e rp: região proximal. Nota-se: Aspecto irregular da glândula acessória (Figuras 7A-B), constrição pronunciada na região distal (Figuras 7D e 7F), entumescimento da glândula acessória e na região distal da glândula principal (Figuras 7B e 7E-F).

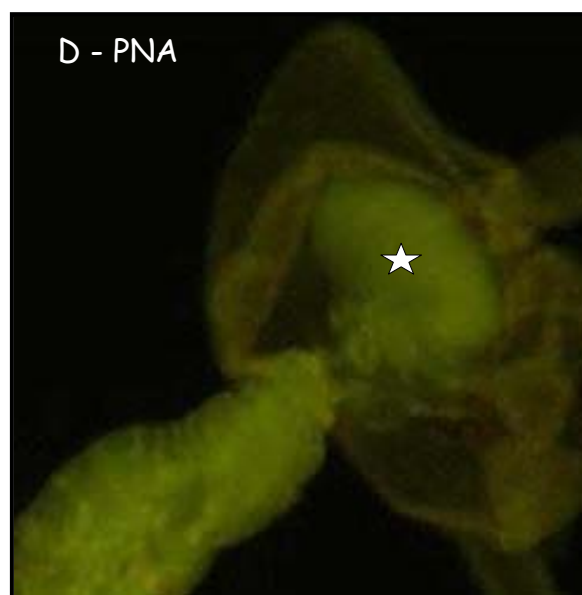
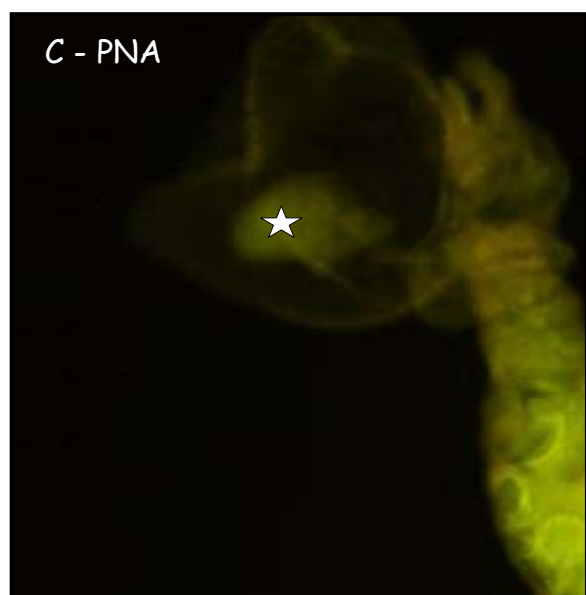
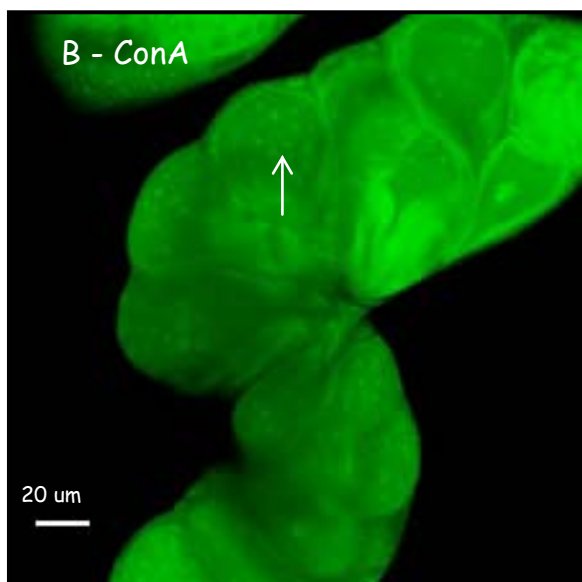
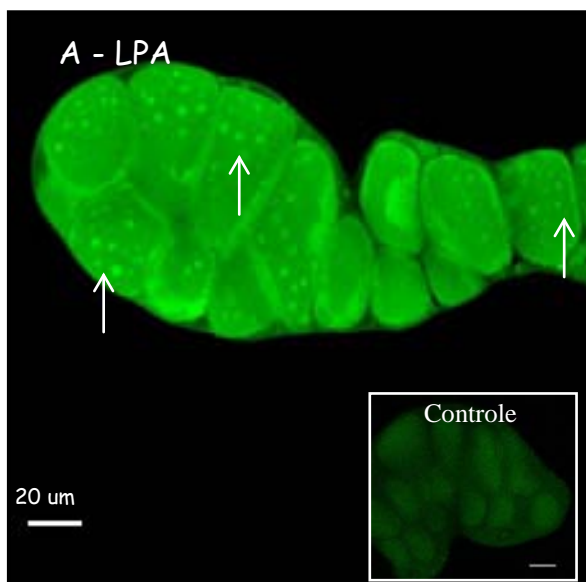


**Figura 8** - Detalhes de fibras nervosas na glândula salivar de fêmeas de *T. guianense*. As imagens foram obtidas em microscópio eletrônico de varredura (MEV). Fibras nervosas no epitélio da região proximal da glândula principal (Figuras 8A, E) e no epitélio da glândula acessória (Figuras B-D, F). As fibras nervosas estão indicadas com uma seta. O círculo indica os furos na região proximal da glândula principal.

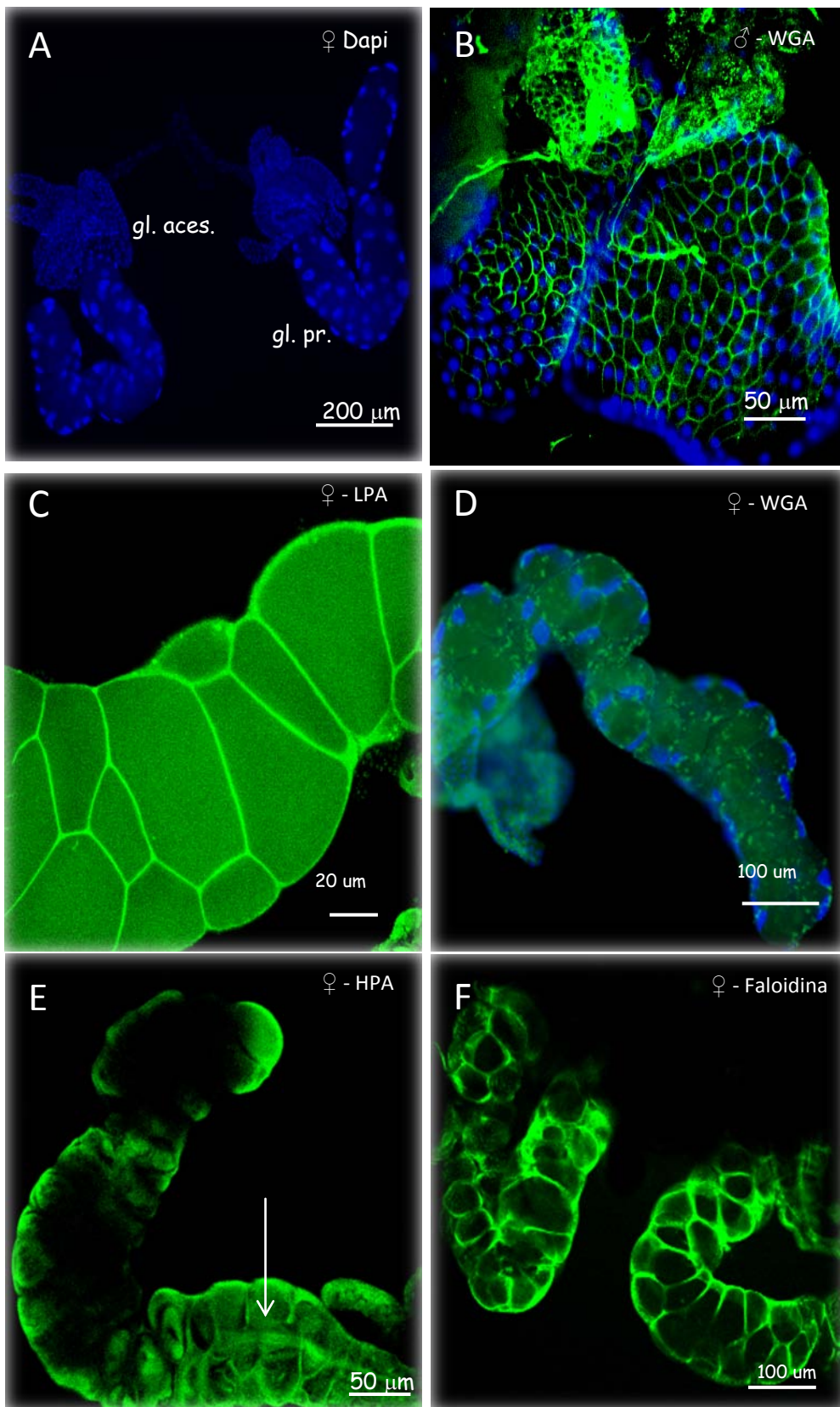




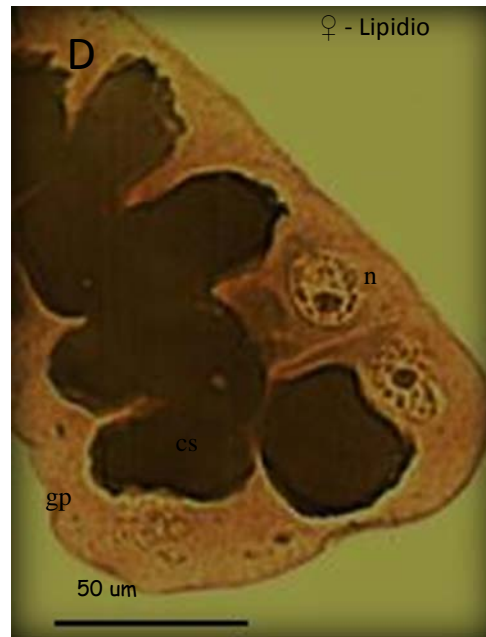
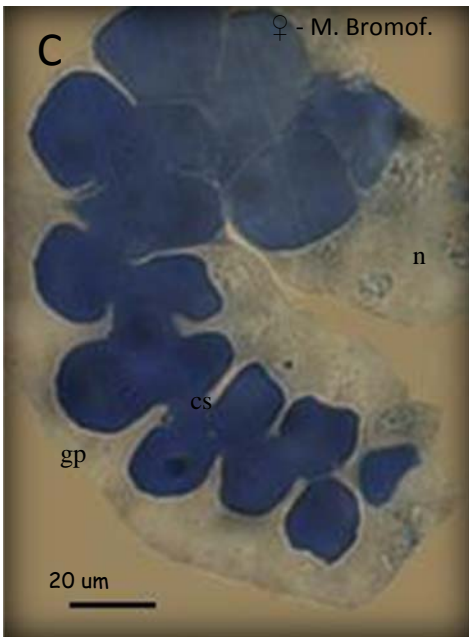
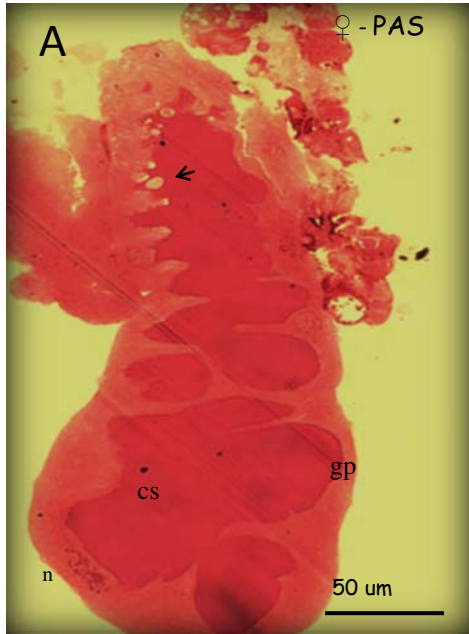
**Figura 9** - Aspecto geral da saliva de fêmea de *T. guianense* em diferentes partes da glândula salivar. Detalhe da junção dos ductos da glândula principal com a glândula acessória (Figura 9A), região interna da glândula principal fraturada com presença de aberturas (possivelmente ductos das cavidades secretoras – dcs, Figura 9B), Secreção salivar da glândula acessória (Figura 9C) e secreção dentro do ducto comum (Figura 9D). Abreviações: dcs: ducto da cavidade secretora; ga: glândula acessória; gp: glândula principal. Seta: detalhe da junção dos ductos das glândulas.



**Figura 10** - Marcação da secreção salivar de fêmeas de *T. guianense* com o uso de lectinas fluorescentes em glândulas com dois dias de idade. A: Lectina LPA, B: Lectina ConA, C – D: Lectina PNA. As imagens foram obtidas em microscópio laser confocal (figuras 10A - B) e microscópio de fluorescência (figuras 10 C - D). As setas e asterísticos indicam marcação de secreção. Controle: glândula salivar sem tratamento com lectina.

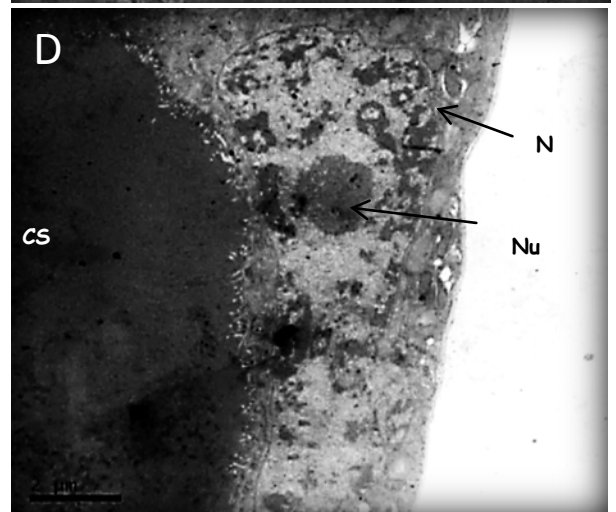
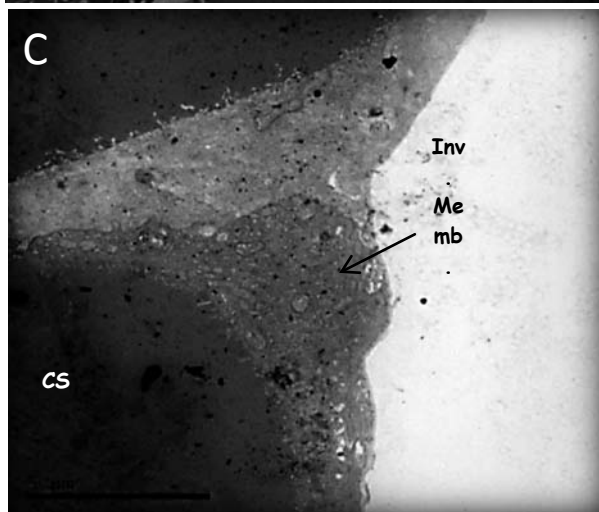
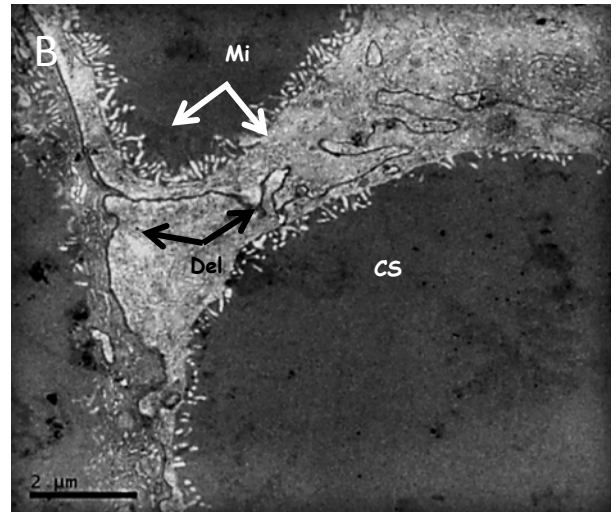
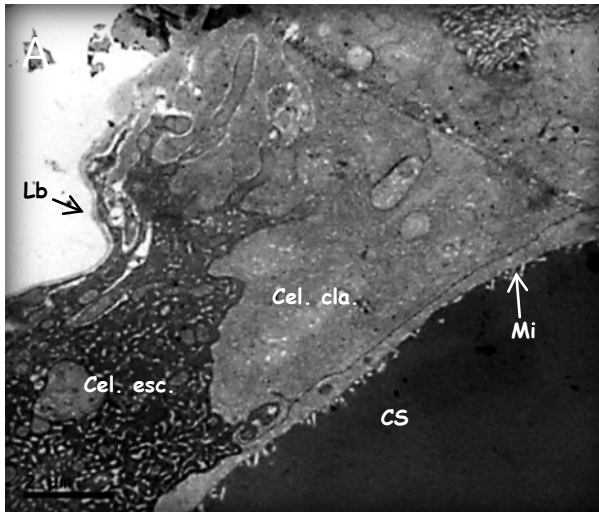


**Figura 11** - Micrografias da glândula salivar de fêmeas de *T. guianense* em microscopia laser confocal. A: Marcação nuclear da glândula de fêmea; B: Marcação com lectina WGA em glândula salivar de macho; C: Marcação de LPA na glândula principal de fêmea; D: Marcação com lectina WGA em glândula salivar de fêmea; E: Marcação com lectina HPA em glândula salivar de fêmea mostrando o ducto interno na região anterior (seta); F: Marcação com faloidina-FITC em glândula salivar de fêmea.



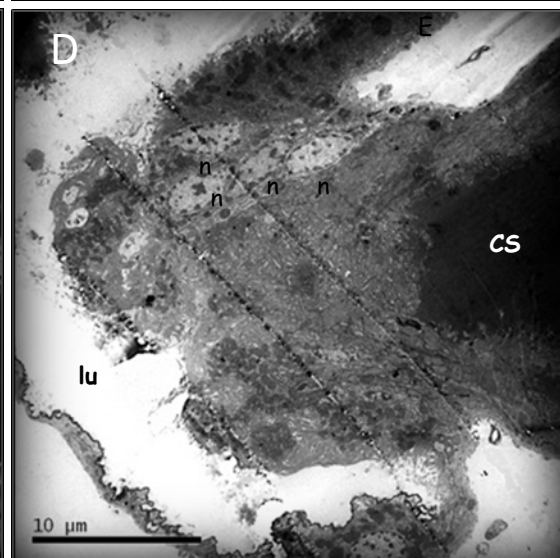
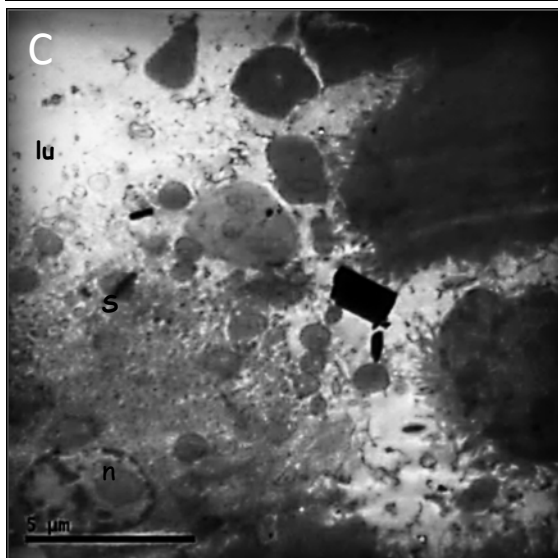
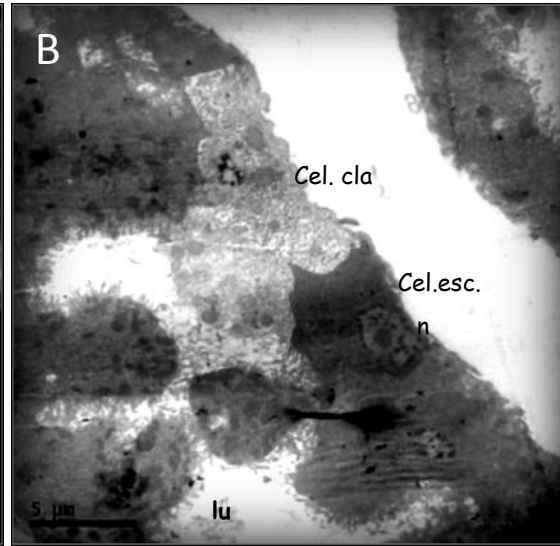
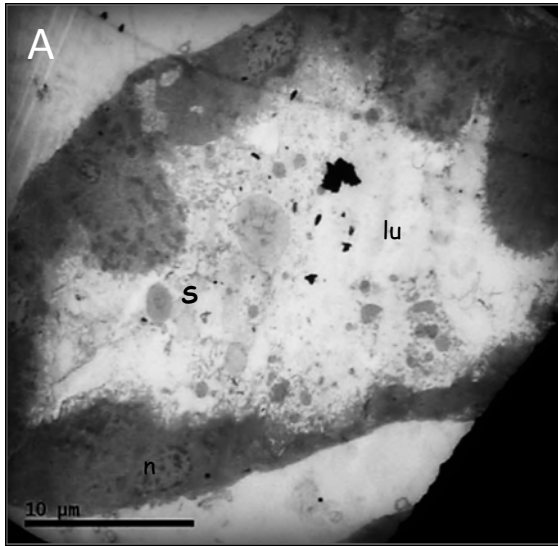
**Figura 12** - Histologia da glândula salivar de fêmeas de *T. guianense*. Secções de 2 µm de glândula salivar de fêmeas corados com PAS (Figura 12A), azul de toluidina (Figura 12B), mercúrio de bromofenol (Figura 12C) e ósmio (Figura 12D). Abreviações: d: ducto; cs: cavidade secretora; lu: lúmen; n: núcleo; gp: glândula principal. A seta indica vesículas secretoras. Nota-se: Em fêmeas, uma saliva rica em substâncias acidófilas, osmiófilas, proteicas e glicídica.







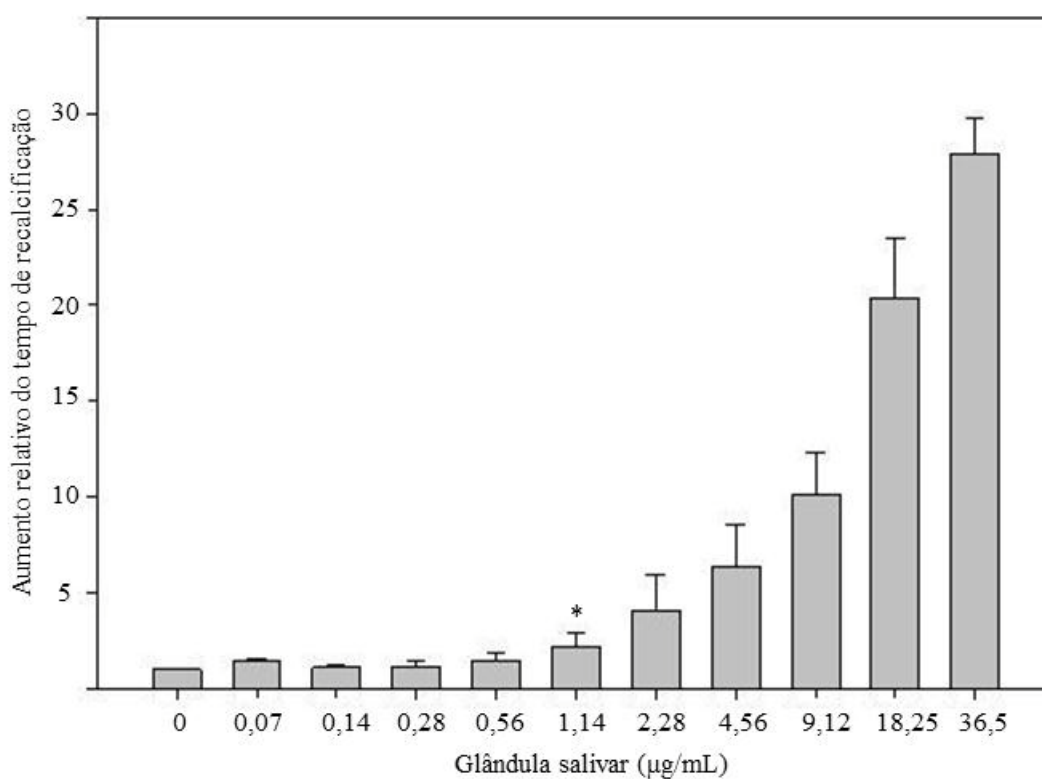
**Figura 13** - Ultraestrutura da glândula principal de fêmeas de *T. guianense*. As imagens foram obtidas em microscópio eletrônico de transmissão (MET). Nota-se: Presença de dois tipos celulares (célula clara e escura) (Figura 13A), cavidades secretoras elétrondensa com microvilosidades circundando essas cavidades (Figura 13 A-D), delimitações celulares revelando pouco espaço celular (Figura 13B), invaginações de membrana (Figura 13C), núcleo basal e comprimido contra a membrana da célula e nucléolo aparente (Figura 13D).



**Figura 14** - Ultraestrutura da glândula acessória de fêmeas de *T. guianense*. As imagens foram obtidas em microscópio eletrônico de transmissão (MET). Nota-se: Lúmen elétron-lucente com presença de secreção (Figuras 14A e C), dois tipos celulares (célula clara e escura) (Figura 14B), região de intersecção da glândula principal com a glândula acessória com presença de muitas pequenas células, grande quantidade de microvilosidades na região de contato com o lúmen da glândula acessória (Figura 14D).

**Tabela 2** - Quantificação de proteína de homogenados de glândulas salivares (HGS) de fêmeas de dois dias de idade de *T. guianense*

| Amostra              | µg de proteína/par glândula |
|----------------------|-----------------------------|
| 1                    | 2.44                        |
| 2                    | 2.84                        |
| 3                    | 2.80                        |
| 4                    | 2.28                        |
| 5                    | 2.72                        |
| <b>Média</b>         | 2.60                        |
| <b>Desvio Padrão</b> | ±0.2                        |



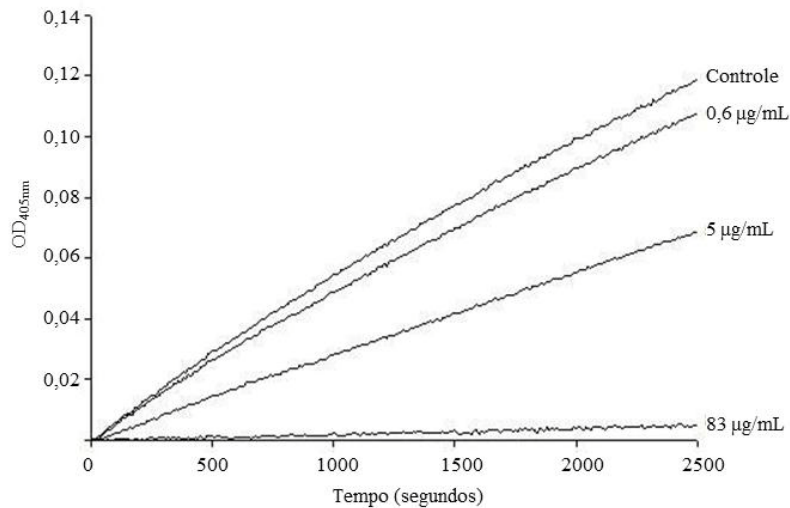
**Figura 15** – Atividade anticoagulante de homogenados de glândulas salivares (HGS) de fêmeas de dois dias de idade de *T. guianense*. O tempo de recalcificação do plasma humano normal foi medido na ausência (controle) e na presença de HGS (contendo diferentes concentrações de proteína). O ensaio foi feito em triplicatas. Os resultados foram expressos como o aumento relativo do tempo de recalcificação (o aumento da turbidex da amostra teste em relação ao controle).

**Tabela 3** - Ensaio de tempo de tromboplastina parcial ativada (TTPa) na ausência (controle) e na presença de diferentes concentrações de homogenados de glândulas salivares (HGS) de fêmeas de dois de idade de *T. guianense*

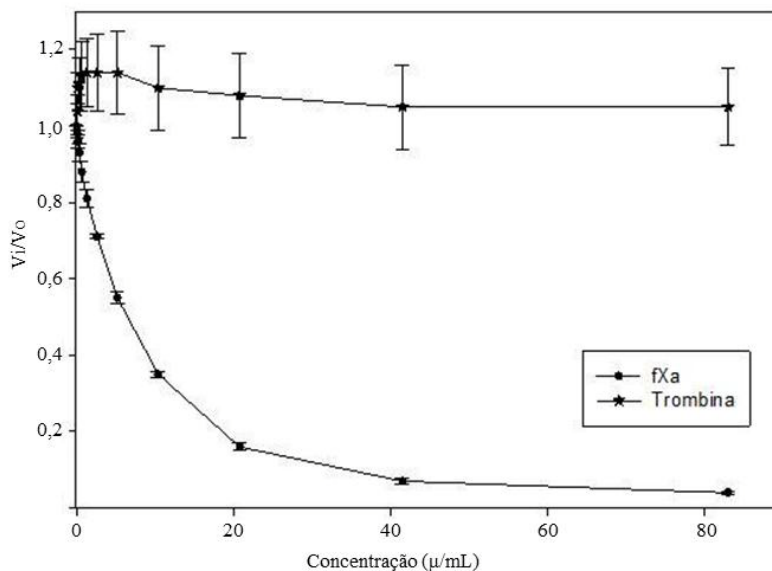
| µg/mL (HGS) | TTPa (seg) | STDEV |
|-------------|------------|-------|
| 0           | 36,6       | 0,2   |
| 2,1         | 37,9       | 0,6   |
| 6,3         | 43,3       | 1,3   |
| 18,9        | 49,6       | 4,4   |
| 31,5        | 64,3       | 7,1   |
| 63          | 73,2       | 11,2  |

**Tabela 4** - Ensaio de tempo de tromboplastina (TP) na ausência (controle) e na presença de diferentes concentrações de homogenados de glândulas salivares (HGS) de fêmeas de dois dias de idade de *T. guianense*

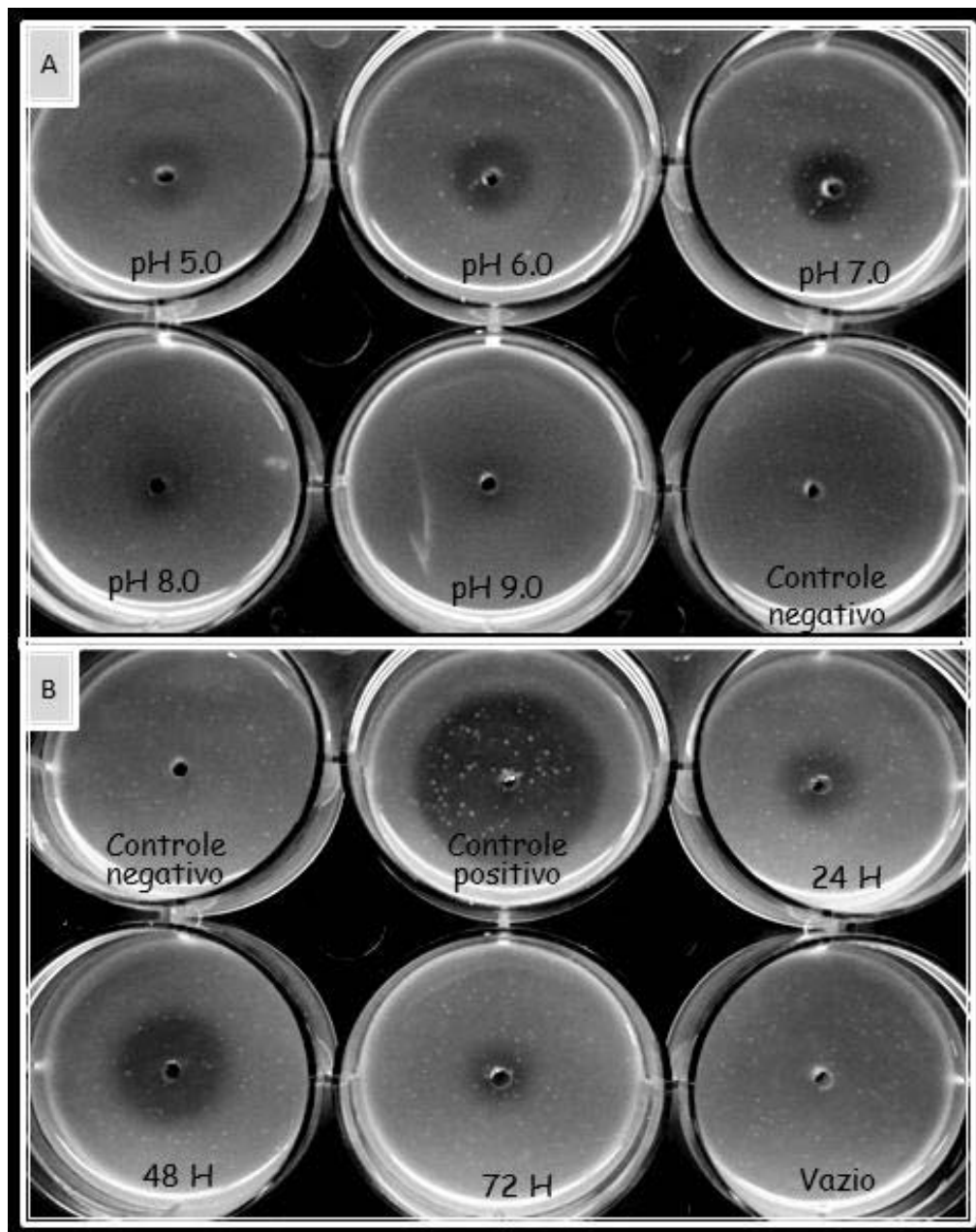
| µg/mL (HGS) | TP (seg) | STDEV |
|-------------|----------|-------|
| 0           | 14,5     | 0,4   |
| 3,2         | 14,2     | 0,0   |
| 9,6         | 15,8     | 0,9   |
| 28,8        | 17,5     | 0,8   |
| 48          | 20,4     | 0,6   |
| 96          | 25,0     | 0,5   |



**Figura 16** – Inibição da atividade catalítica de 0.5 nM de fXa por homogenados de glândula salivar de *T. guianense*. A hidrólise do substrato cromogênico pelo fXa foi medido na ausência (controle) e na presença de HGS (contendo diferentes concentrações de proteína). A velocidade máxima da reação ( $V_{max}$ ) foi obtida em OD 405nm. O ensaio foi realizado em triplicata. O valor de  $V_i/V_o$  foi calculado como a relação da  $V_{max}$  da amostra pela  $V_{max}$  do controle, onde o valor  $V_i/V_o$  do controle é igual a 1.



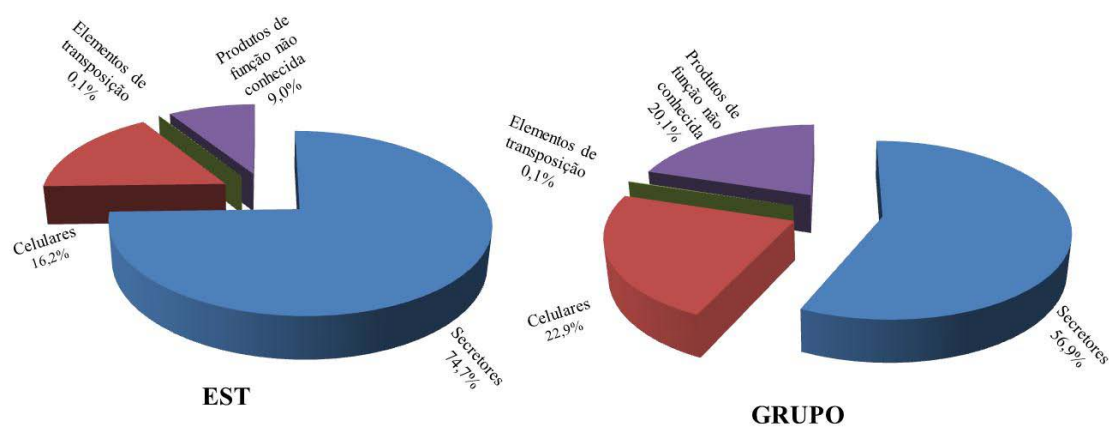
**Figura 17** – Ensaio de atividade catalítica de 0.5 nM de fXa e 0.1 nM de trombina por homogenados de glândula salivar de *T. guianense*. A hidrólise do substrato cromogênico S-2222 (pelo fXa) e S-2238 (pela trombina) foi medido na ausência (controle) e na presença de HGS (contendo diferentes concentrações de proteína) e a velocidade máxima da reação ( $V_{max}$ ) foi obtida em OD 405nm. O ensaio foi conduzido em triplicata. O valor de  $V_i/V_o$  foi calculado como a relação da  $V_{max}$  da amostra pela  $V_{max}$  do controle, onde o valor  $V_i/V_o$  do controle é 1.



**Figura 18** - Atividade de lisozima em homogenados de glândulas salivares (HGS) de fêmeas de *T. guianense*. A atividade de lisozima foi medida em placas de agarose 1% contendo 1 mg/mL de *M. lysodeikticus*, em diferentes faixas de pH (5,0 – 9,0) (Figura A) e em diferentes idades de envelhecimento glandular (24 a 72h) em pH 7,0 (Figura B). As placas foram incubadas por 16 horas a 37°C. Como controle positivo foi utilizado 2 µg de lisozima comercial (Sigma) e como controle negativo foi utilizado H<sub>2</sub>O.

**Tabela 5** - Classificação funcional de transcritos salivares originados da glândula salivar de *T. guianense*

| Classe                           | Número de ESTs | Número de Grupos | EST's/Grupo |
|----------------------------------|----------------|------------------|-------------|
| Produtos secretorários           | 1324           | 428              | 3,1         |
| Celulares                        | 288            | 172              | 1,7         |
| Elementos de transposição        | 1              | 1                | 1,0         |
| Produtos de função não conhecida | 159            | 151              | 1,1         |
| Total                            | 1772           | 752              |             |



**Figura 19** - Classificação das classes funcionais das etiquetas de sequências expressas (EST) ou grupos assemblados (GRUPO) derivados da biblioteca de cDNA de fêmeas de *T. guianense*

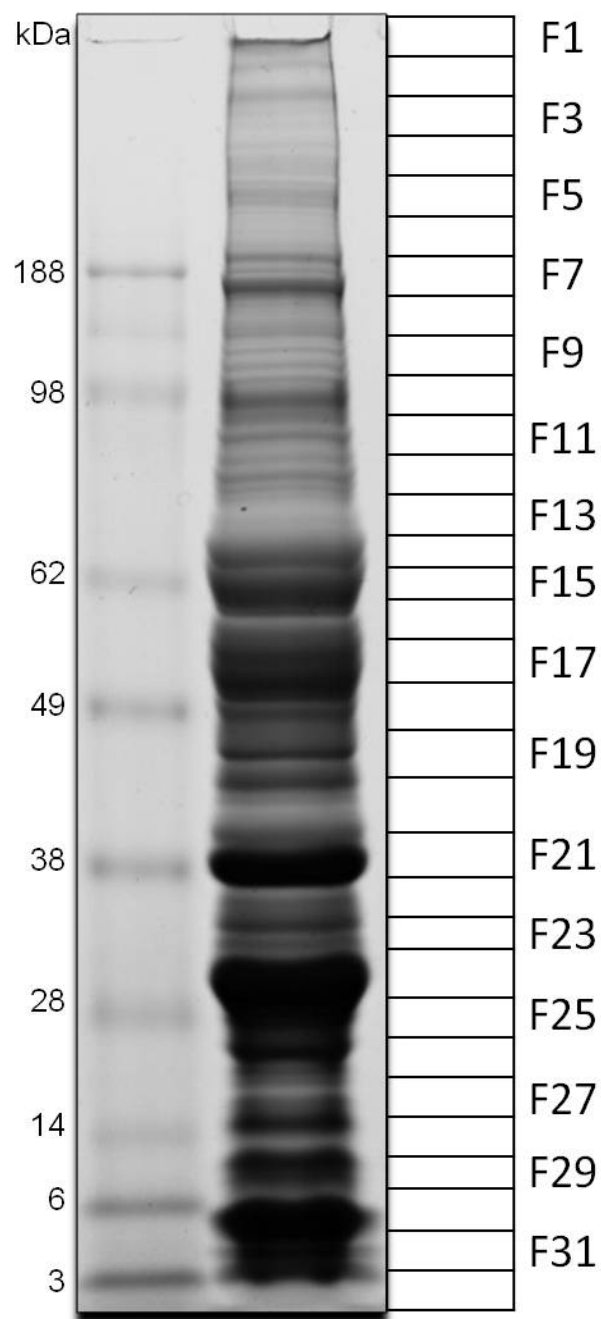


**Tabela 6** - Classificação funcional dos transcritos secretados originados das glândulas salivares de *T. guianense*

| <b>Família</b>                                 | <b>Número de ESTs</b> | <b>Grupo</b> | <b>ESTs/Grupo</b> |
|--|-----------------------|--------------|-------------------|
| <b>Famílias de proteínas ubíquas</b>           |                       |              |                   |
| Antígeno-5                                     | 5                     | 2            | 2,5               |
| Yellow   | 1                     | 1            | 1,0               |
| Domínios ML                                    | 1                     | 1            | 1,0               |
| Lipocalinas                                    | 2                     | 2            | 1,0               |
| Produtos relacionados a imunidade              | 36                    | 12           | 3,0               |
| <b>Domínios inibidores de proteases</b>        |                       |              |                   |
| Serpina  | 1                     | 1            | 1,0               |
| Domínio Kunitz contendo polipeptídeos          | 19                    | 9            | 2,1               |
| <b>Enzimas</b>                                 |                       |              |                   |
| Tripsina                                       | 69                    | 28           | 2,5               |
| Hialuronidase                                  | 5                     | 1            | 5,0               |
| Apirase  | 14                    | 4            | 3,5               |
| Amilase  | 61                    | 14           | 4,4               |
| Adenosina deaminase                            | 1                     | 1            | 1,0               |
| Destabilase                                    | 1                     | 1            | 1,0               |
| <b>Famílias específicas de insetos</b>         |                       |              |                   |
| Aegyptina                                      | 23                    | 11           | 2,1               |
| D7 (superfamília OBP)                          | 242                   | 72           | 3,4               |
| <b>Famílias específicas de <i>Simulium</i></b> |                       |              |                   |
| SVEP   | 190                   | 56           | 3,4               |
| Outras famílias de <i>Simulium</i>             | 427                   | 146          | 2,9               |
| Mucinas PolyQ                                  | 51                    | 16           | 3,2               |
| Repetições GH                                  | 127                   | 33           | 3,8               |
| Similar a colágeno                             | 48                    | 17           | 2,8               |
| <b>Total</b>                                   | <b>1324</b>           | <b>428</b>   |                   |

**Tabela 7** - Classificação funcional de transcritos celulares originados de glândulas salivares de *T. guianense*

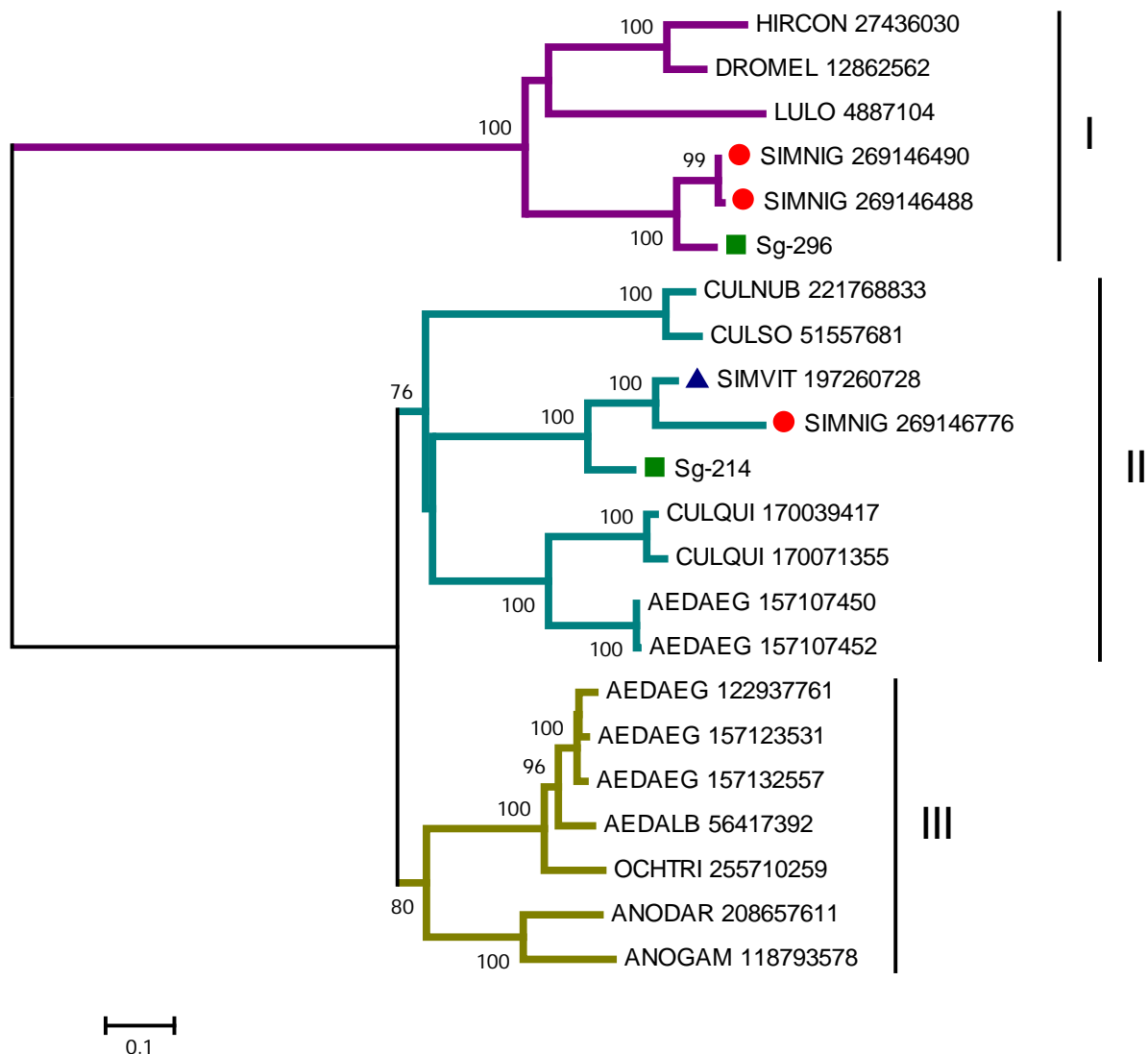
| <b>Função</b>                         | <b>Número de ESTs</b> | <b>Grupo</b> | <b>ESTs/Grupo</b> |
|---------------------------------------|-----------------------|--------------|-------------------|
| Síntese de proteína                   | 113                   | 71           | 1,6               |
| Metabolismo energético                | 33                    | 28           | 1,2               |
| Tradução de sinal                     | 18                    | 14           | 1,3               |
| Transportadores e estocagem           | 5                     | 5            | 1,0               |
| Maquinaria do proteosoma              | 1                     | 1            | 1,0               |
| Maquinaria de modificação de proteína | 6                     | 6            | 1,0               |
| Maquinaria de exportação de proteína  | 4                     | 4            | 1,0               |
| Regulação nuclear                     | 2                     | 2            | 1,0               |
| Metabolismo do nucleotídeo            | 2                     | 2            | 1,0               |
| Maquinaria de transcrição             | 2                     | 2            | 1,0               |
| Metabolismo do carboidrato            | 3                     | 3            | 1,0               |
| Metabolismo do aminoácido             | 1                     | 1            | 1,0               |
| Metabolismo de detoxificação          | 3                     | 3            | 1,0               |
| Citoesqueleto                         | 2                     | 2            | 1,0               |
| Conservadas de função não conhecida   | 93                    | 28           | 3,3               |
| <b>Total</b>                          | <b>288</b>            | <b>172</b>   |                   |



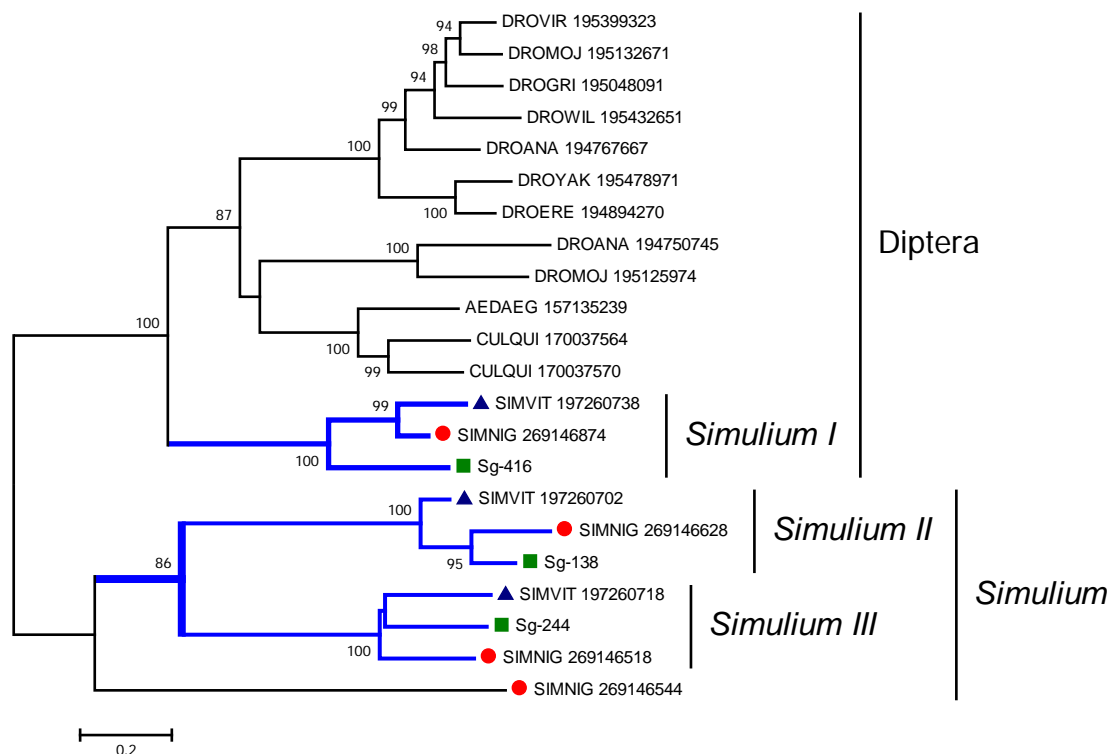
**Figura 20** - Gel de eletroforese 1D de homogenados de glândulas salivares de *T. guianense*. O número da esquerda indica o peso molecular do marcador de proteína (kDa), demonstrado na linha esquerda. O gel da linha direita demonstra a separação das proteínas da glândula salivar. O molde na direita (F1-32) representa as frações do gel submetidas para digestão trípica e identificação MS/MS.

**Tabela 8:** Proteínas secretadas deduzidas da análise do sialotranscriptoma e identificação de expressão por análise proteômica

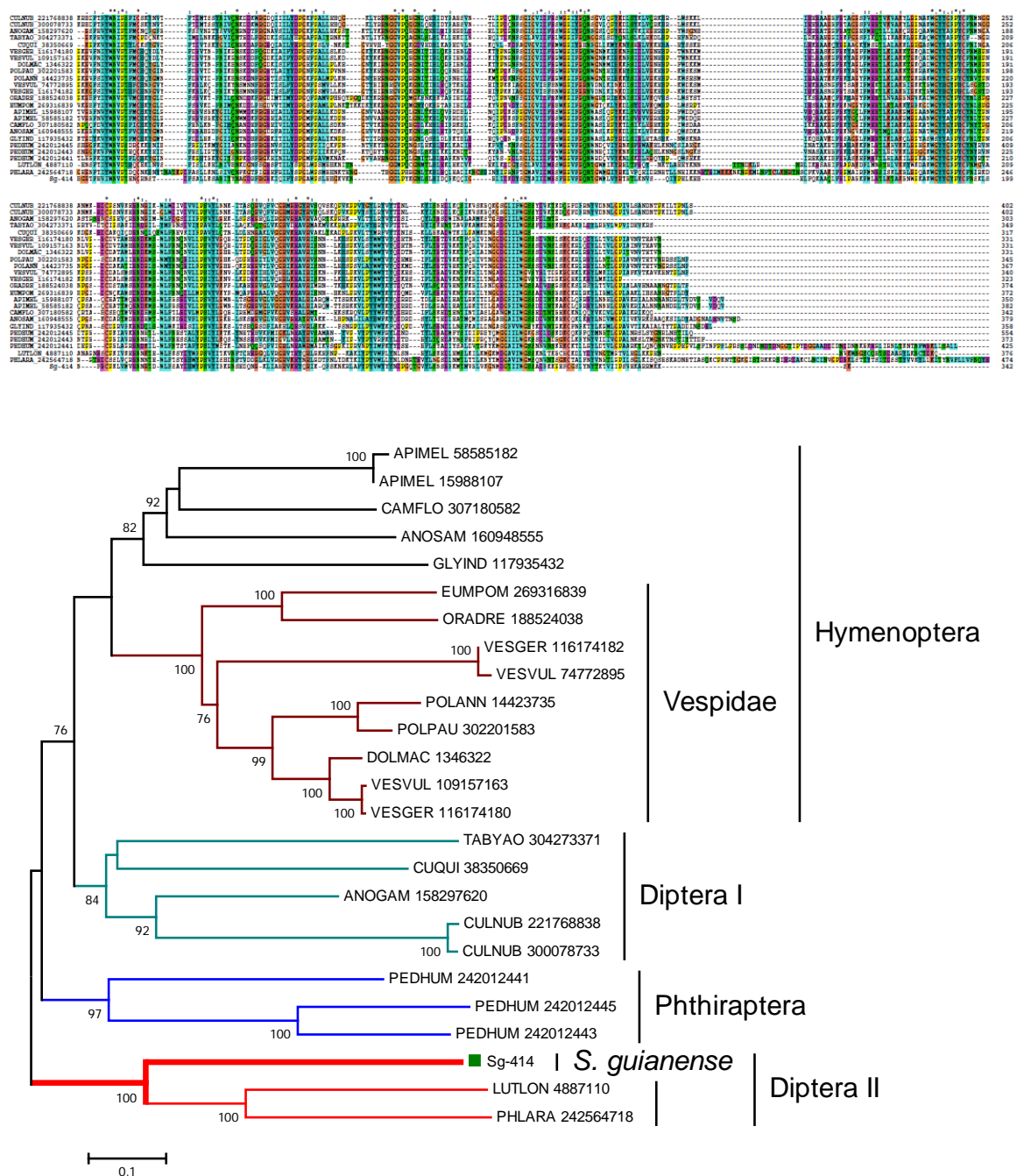
| Descrição  | Nome da proteína   Fração → número de peptídeos   |
|--|---|
| Similar a laminina                                 | Sg-431 F14→2  |
| Amilase/maltase                                    | Sg-214 F16→116, Sg-296 F17→32   |
| Apirase  | Sg-354 F16→58   |
| Mucina de <i>Simulium</i>                          | Sg-126 F16→12, Sg-129 F16→12, Sg-121 F16→12, Sg-120 F16→11, Sg-127 F16→11, Sg-117 F16→10, Sg-128 F16→10, Sg-119 F16→9, Sg-125 F9→7  |
| Hialuronidase                                      | Sg-414 F20→14   |
| Proteína secretada de Diptera conservada de inseto | Sg-215 F21→134, Sg-216 F21→133, Sg-292 F17→49, Sg-347 F14→36, Sg-256 F13→28   |
| Similar a colágeno de <i>Simulium</i>              | Sg-152 F23→87, Sg-149 F23→84  |
| Ácida de 28 kDa                                    | Sg-320 F23→45, Sg-321 F23→41, Sg-319 F23→23   |
| D7 longa   | Sg-261 F26→42, Sg-220 F23→24, Sg-218 F23→14   |
| Serino proteases                                   | Sg-244 F24→48, Sg-138 F25→28, Sg-416 F24→6  |
| Antígeno-5   | Sg-457 F24→30   |
| Aegyptina  | Sg-276 F24→37   |
| <i>Simulium</i> básica de 28 kDa                   | Sg-136 F24→34   |
| Deorfanizada de 19 kDa                             | Sg-303 F27→17, Sg-309 F27→9   |
| Sv 7,8 kDa   | Sg-356 F27→11, Sg-372 F30→9, Sg-227 F32→4, Sg-205 F31→3   |
| D7 16 kDa  | Sg-331 F29→14, Sg-350 F30→9,  |
| <i>Simulium</i> básica de 13 kDa                   | Sg-446 F29→2  |
| Deorfanizada Sn 8-10 Cys W                         | Sg-340 F29→22, Sg-324 F29→12  |
| Lisozima   | Sg-263 F30→7  |
| Domínio Kunitz                                     | Sg-395 F30→2  |
| SVEP   | Sg-1 F30→164, Sg-8 F30→149, Sg-92 F30→53, Sg-102 F30→43, Sg-94 F30→34, Sg-93 F30→33, Sg-100 F30→32, Sg-95 F30→26, Sg-90 F30→25, Sg-99 F30→25, Sg-103 F30→22, Sg-101 F30→20, Sg-344 F30→14 |
| 5-Cys <i>Simulium</i>                              | Sg-282 F30→3  |
| Básica 7-13 kDa de <i>Simulium</i>                 | Sg-403 F30→8, Sg-420 F31→3  |
| Cecropina  | Sg-368 F31→2, Sg-369 F31→3  |
| D7 ultra-curtas                                    | Sg-383 F31→3  |
| <i>Simulium</i> básica 7,4 kDa                     | Sg-422 F31→4  |
| Deorfanizadas básicas de 8 kDa                     | Sg-258 F31→5  |
| Similar para Kunitoxina                            | Sg-375 F31→4  |



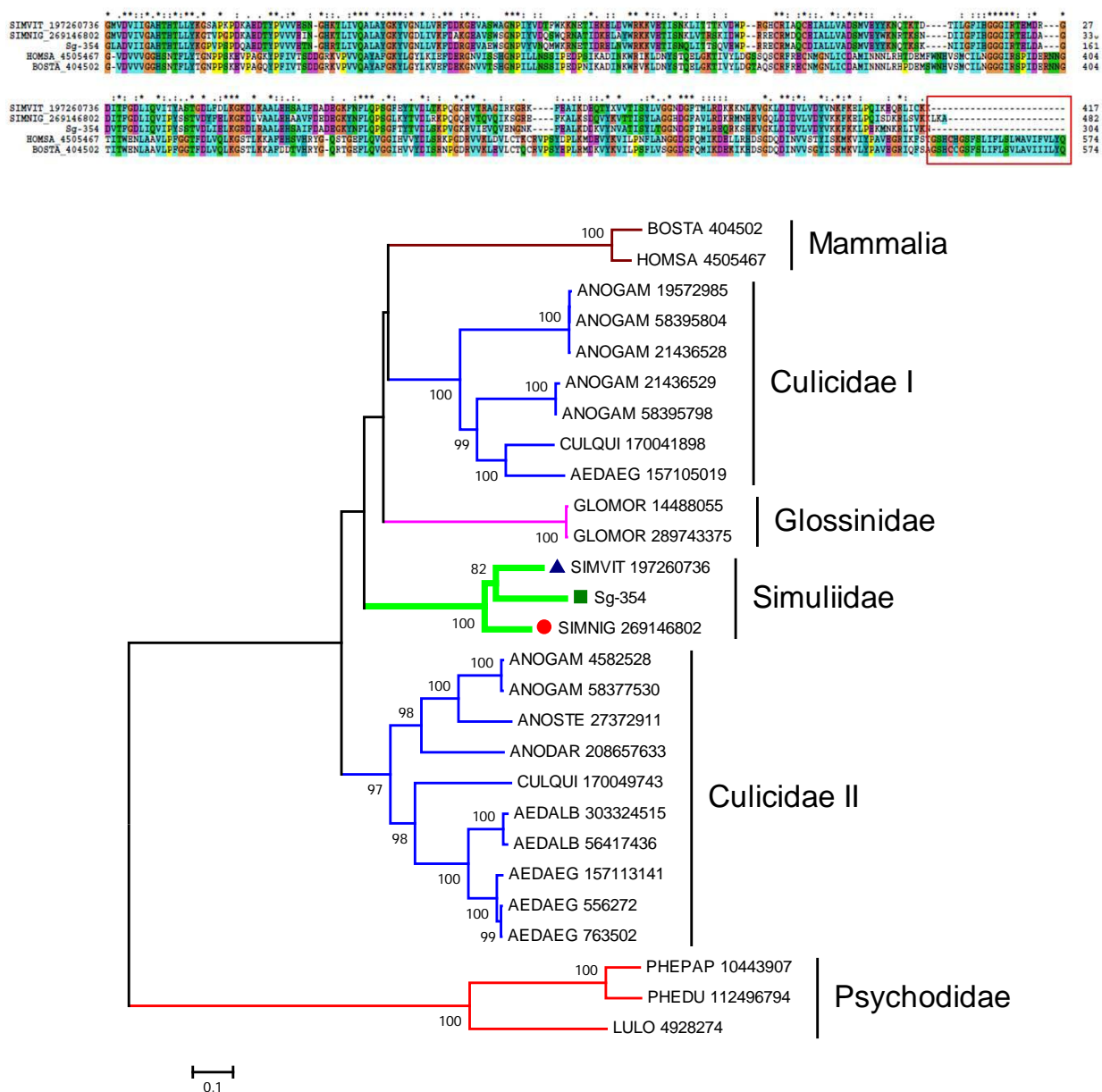
**Figura 21** - Família de proteína amilase/maltase de *T. guianense*. Filograma derivado do alinhamento de duas proteínas de *T. guianense* (indicado por um quadrado e começando com Sg-), com seus melhores homólogos no Centro Nacional para Informação em Biotecnologia (NCBI) contendo seus números de acessos. Glicosidases de seus homólogos *S. vittatum* e *S. nigrimanum* estão indicados por um triângulo e um círculo, respectivamente. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 10% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



**Figura 22** - Família serino proteases de *T. guianense*. Filograma derivado do alinhamento de três proteínas de *T. guianense* e seus melhores homólogos no Centro Nacional para Informação em Biotecnologia (NCBI). As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. As serino proteases de seus homólogos *S. nigrimanum* (SIMNIG) estão indicadas por um círculo ou por um triângulo para *S. vittatum* (SIMVIT), com seus respectivos números de acessos. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 20% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.

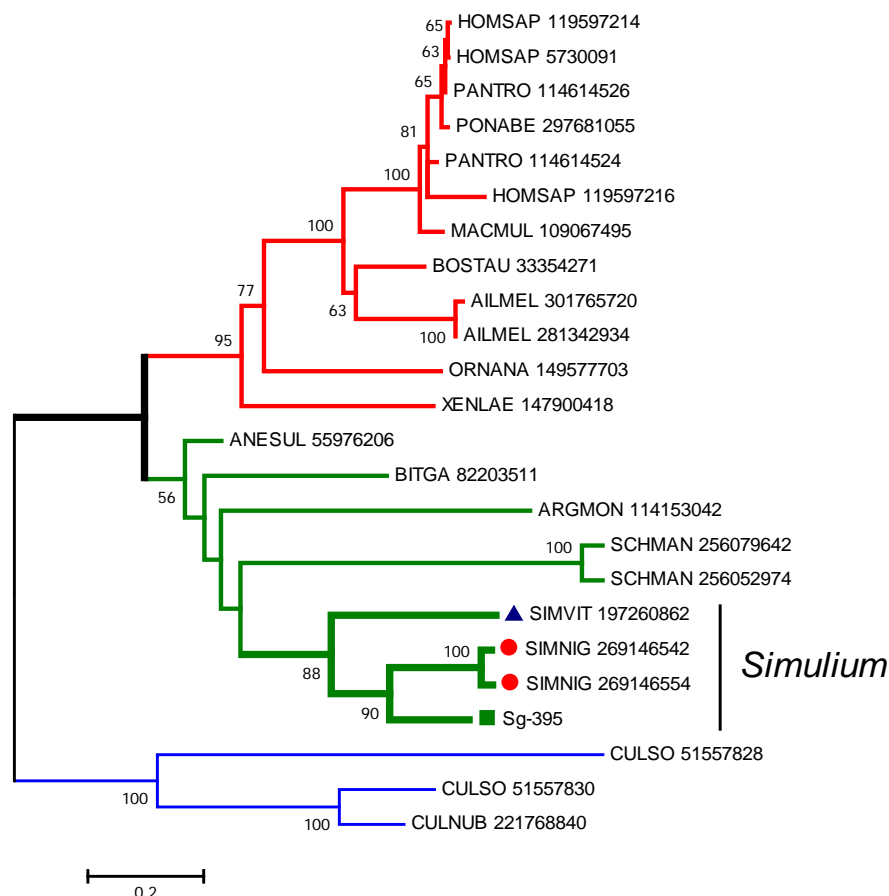


**Figura 23** - Família hialuronidase de *T. guianense*. No topo da figura está demonstrado o alinhamento Clustal da proteína hialuronidase derivada do sialotranscriptoma de *T. guianense* e seus melhores homólogos no Centro Nacional para Informação em Biotecnologia (NCBI). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 10% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendrograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



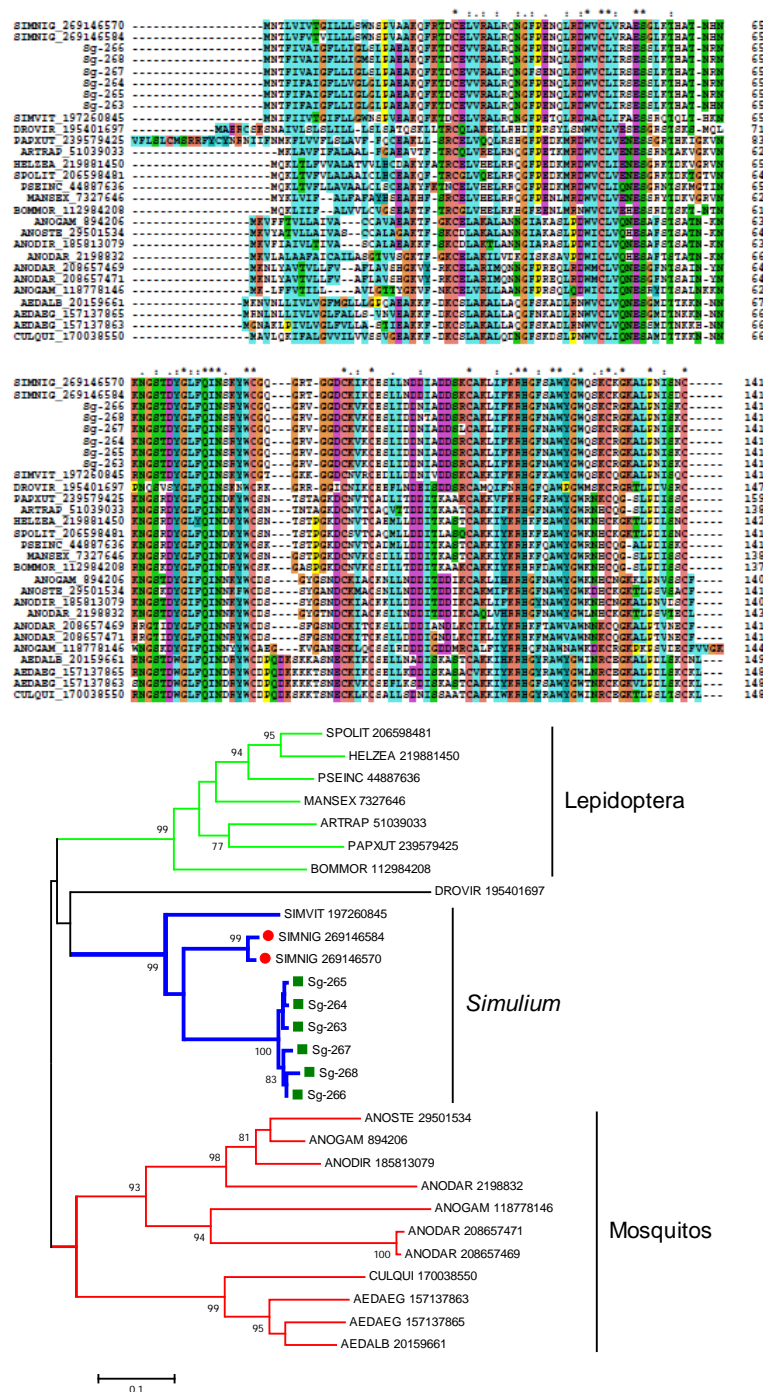
**Figura 24** - Família apirase de *T. guianense*. O alinhamento de uma proteína apirase de cada sialotranscriptoma de simulídeo comparada com enzimas de mamíferos. A proteína de *T. guianense* aparece como Sg-354 e seus homólogos aparecem como SIMVIT\_ (*S. vittatum*) e SIMNIG\_ (*S. nigrimanum*), com seus números de acesso no NCBI. O box demonstra a deleção do domínio carboxiterminal onde o glicolípídeo é ancorado permitindo sua secreção nas apirases salivares de simulídeos. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações da árvore filogenética indicam a porcentagem de suporte da análise parcimonômica após 10.000 vezes, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com o programa Mega.

SIMNIG 269146554 MK-----IADNGRIFEVCLDTMS-----ADGCLPNDLGLCALAYFYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 110  
 SIMNIG 269146542 MK-----IADNGRIFEVCLDTMS-----ADGCLPNDLGLCALAYFYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 110  
 Sg-395 MK-----IADNGRIFEVCLDTMS-----ADGCLPNDLGLCALAYFYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 106  
 SIMNIG 197260862 MN-----LPIGAPFSLTQSLA-----GKNVGNLPFEDVCSALFFFTYRPAICKEPFGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 102  
 SCHMAN 256079642 MP-----SPCLAGILLICVQVASYRKH-----SD-----CLLWERNICGALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 140  
 SCHMAN 256052974 MP-----SPCLAGILLICVQVASYRKH-----SD-----CLLWERNICGALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 140  
 ANESUL 55976206 MK-----LPIGAPFSLTQSLA-----GKNVGNLPFEDVCSALFFFTYRPAICKEPFGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 50  
 ARGMON 114153042 MK-----LPIGAPFSLTQSLA-----GKNVGNLPFEDVCSALFFFTYRPAICKEPFGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 111  
 BITGA 82203511 MDS-----GILLLLGLLILWAEUTPYSGKK-----RPDCCYLPATDSCGMPFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 90  
 BOMGAP 119597214 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 BOMGAP 5730091 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 PANTRO 114614526 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 PONABE 297681055 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 PANTRO 114614524 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 BOMGAP 119597216 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 MACMUL 109067495 MDS-----ARPLGSGILLIFLRAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 BOSTAU 33354271 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 145  
 AILMEL 301765720 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 145  
 AILMEL 281342934 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 146  
 XENLAR 147900418 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 139  
 ORNANA 149577703 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 148  
 CULSO 221768840 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 150  
 CULSO 51557828 MDS-----VPLIWMILGSLILVTAALGDAASEPT-----GNASICLLPLUTGFCALLFFFTYVDSANPCMPYGGCGGAGNPFKRAKACIKRNSVLEK-----TRERKKY-----NFKRQKPLPQVNLDDK----- 146



**Figura 25** - Família Kunitz de *T. guianense*. Alinhamento e o filograma de uma proteína Kunitz de *T. guianense* (indicada por Sg-) com seus melhores homólogos no NCBI. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações da árvore filogenética indicam a porcentagem de suporte da análise parcimônica após 10.000 vezes, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com o programa Mega.



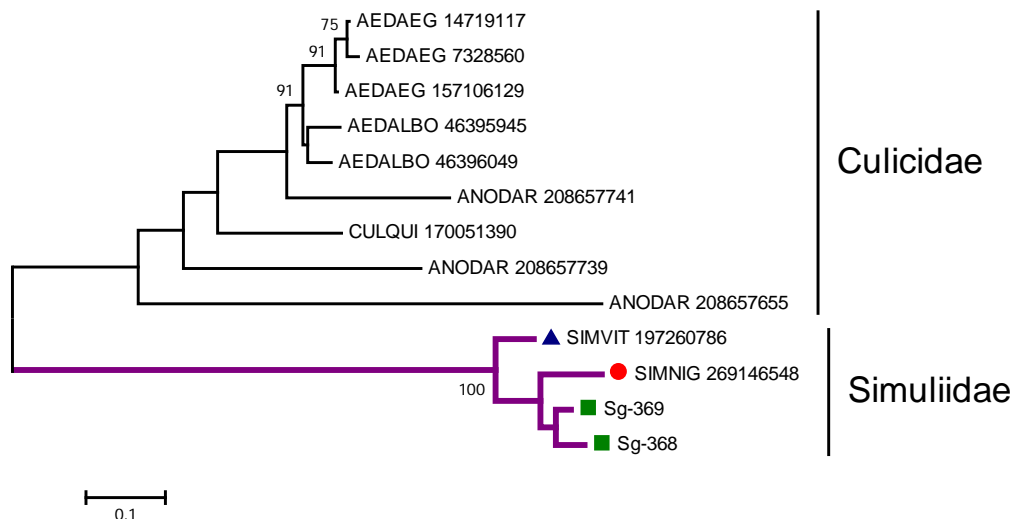


**Figura 26** - Família lisozima de *T. guianense*. Alinhamento e dendograma de lisozimas salivares de *T. guianense* (indicada por Sg-) com seus melhores homólogos no NCBI. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações da árvore filogenética indicam a porcentagem de suporte da análise parcimoniosa após 10.000 vezes, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com o programa Mega.

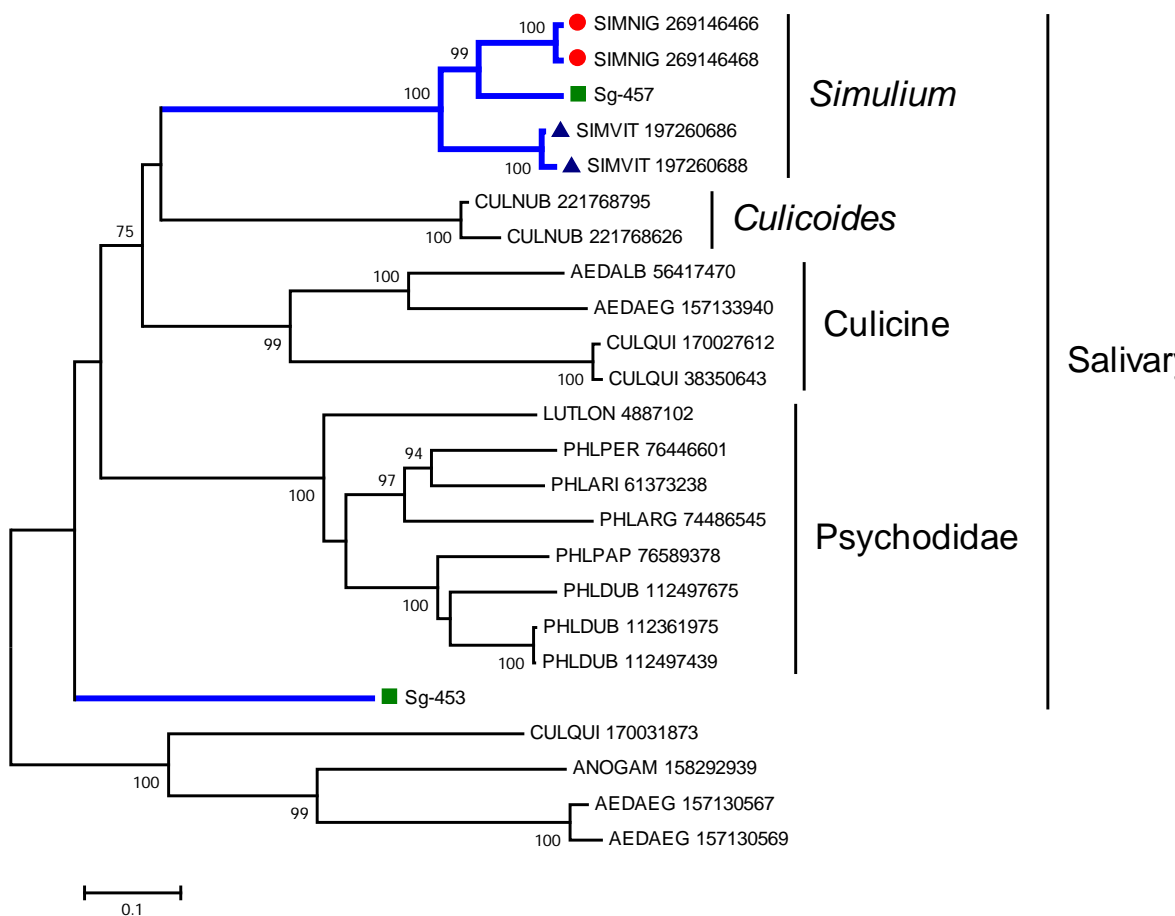
```

*::  ::.  :  *  :*  ::  ::::  .::  .::.*  *:  .:  .:
AEDALBO_46395945 MNFAK-VFVLVMAVLLLVG-Q-SEAGGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALGK- 59
AEDALBO_46396049 MNFTK-IFVLIAMAALLLVG-Q-SEAGGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALGK- 59
AEDAEG_14719117 MNFTK-LFLLIAVAVLLLTG-Q-SEAGGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALGK- 59
AEDAEG_7328560 MNFTK-LFLLIAVAVLLLTG-Q-SEAGGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALRK- 59
AEDAEG_157106129 MNFTK-LFLLIAMAVLLLTG-Q-SEAGGLKKLGKKLEGAGKRVFNAAEKALPVVAGAKALGK- 59
ANODAR_208657741 MNFNKILFLAVVLAALLLVG-Q-TDAGFLKKLGKKLEGAGKRVFNAAEKALPVAVGVKALGK- 60
CULQUI_170051390 MNFNK-LFVIVLLAALAFFG-Q-AGAGGLKKFGKKLEGVGRVFKASEKALPVVIGFKALGK- 59
ANODAR_208657739 MNFTK-LFILVAIVLVIAIGI-PADAAARWKFGRRLERLGRNVFKAACKALPVIAGYKALG-- 60
ANODAR_208657655 MQLKVILLVALVLMALFGG-E-TEARRFRKFLKKVEGAGRRITNAAGKGLPVVAGVKGIIIG- 60
Sg-369 MNFKT-IFLIVAMILVADFG----QAD--ARFKDKLKHGAKKALDVASKAAPIVAAGAAIARG 56
Sg-368 MNFKT-IFVIVAMILVAVFG----QAD--ARFKDKLKHGAKKALDVASKAAPIVAAGAAIARG 56
SIMNIG_269146548 MNFKT-IFVLAIVLVAVFG----QAD--AKFKDKLKHGAKKALDVASKAAPIVAAGAAIARG 56
SIMVIT_197260786 MNFKG-LFFVFAILLVVIFG----QAD--AKFKDKLKSAGKALDVASKAAPIVAAGAAIARG 56

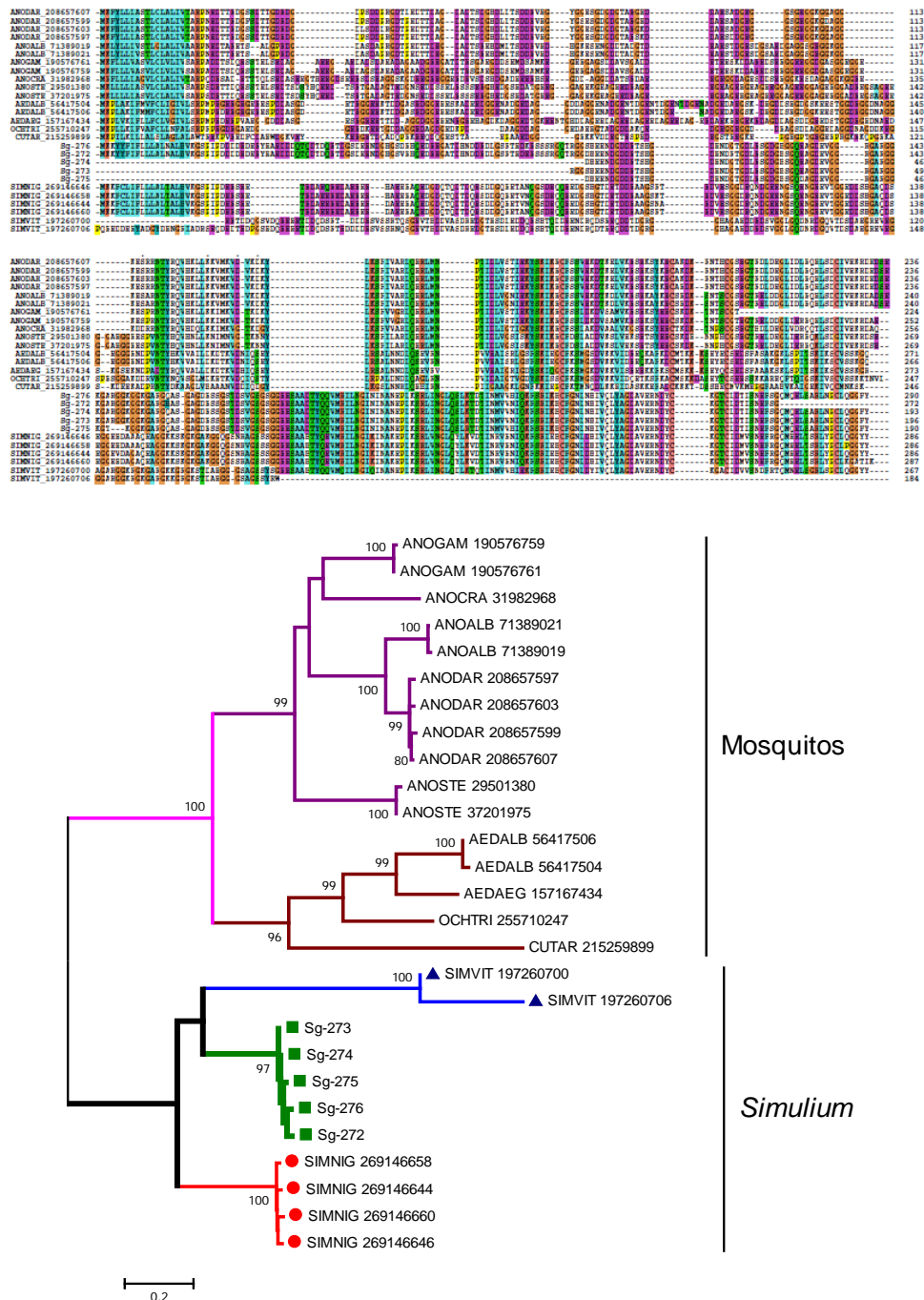
```



**Figura 27** - Família cecropina de *T. guianense*. Alinhamento e dendograma de duas proteínas cecropinas de *T. guianense* (indicada por Sg-). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.

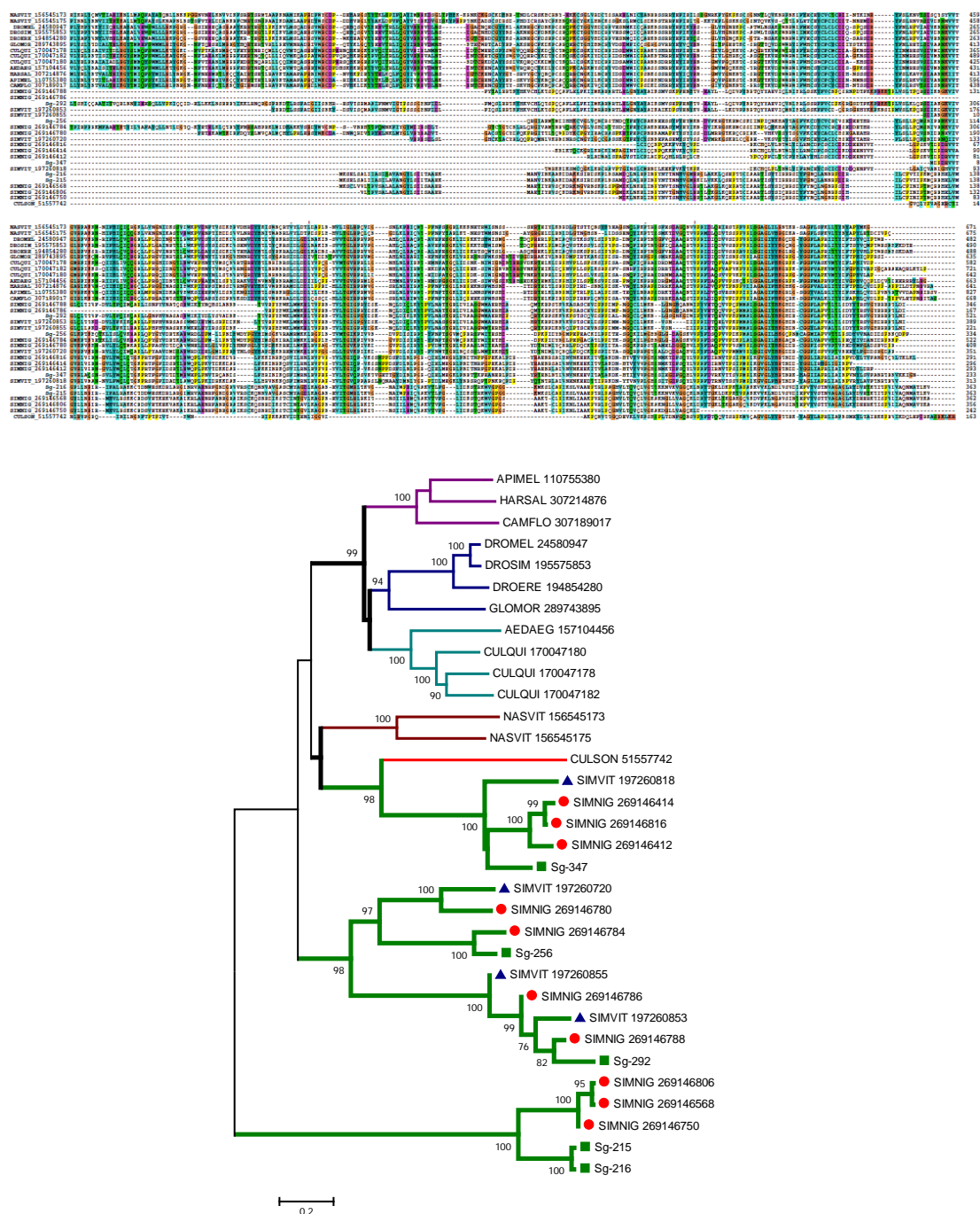


**Figura 28** - Família antígeno-5 de *T. guianense*. O filograma de duas proteínas antígeno-5 derivadas do sialotranscriptoma de *T. guianense* com seus melhores homólogos no NCBI. As proteínas de *T. guianense* (Sg-457 e Sg-453) estão indicadas por um quadrado e seus homólogos *S. vittatum* por um triângulo e *S. nigrimanum* por um círculo. OS números de acesso no NCBI estão incluídos. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.



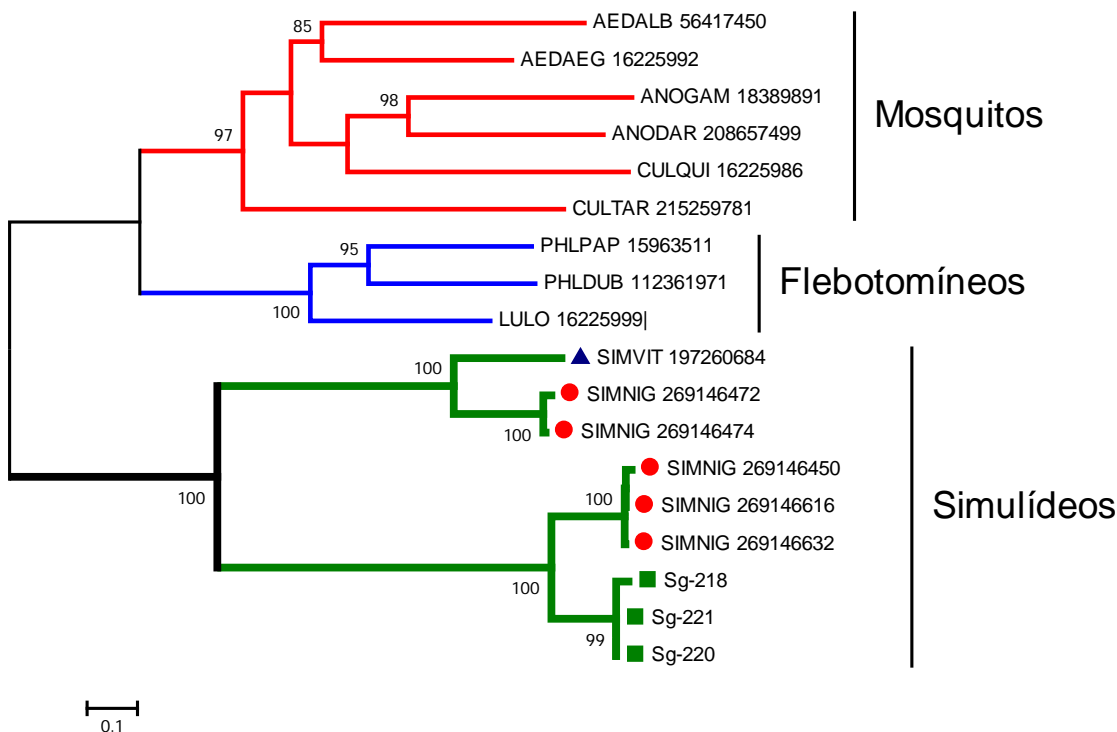
**Figura 29** - Família aegyptina de *T. guianense*. Alinhamento e dendograma das proteínas Aegyptina de *T. guianense* (indicada por Sg-). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 20% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o filograma foi feito com um pacote do programa Mega.





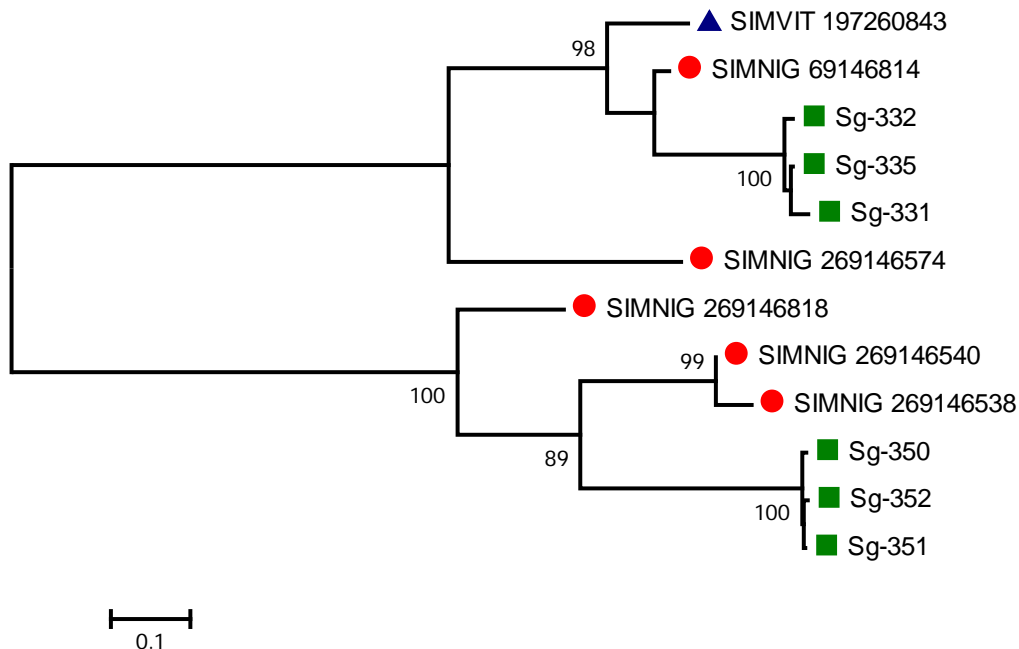
**Figura 30** - Família de proteína secretada em Diptera e conservadas em insetos de *T. guianense*. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimônica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 20% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o filograma feito com um pacote do programa Mega.

AEDALB 16225992 MPPRFKILSSDIAALVLAAPMCNE--DPKPPFIRORNEHWFAKN--FAALQMLNKLKESDQAOVCKVLRKIPPS-----PREFPKVVMNRNPNFLNMRKFKHGTSTGFPKSSGSEVPAAPFKVGRHSGAILPFGSSFK 165  
 AEDALB 56417450 -----NHILLVAVVSSIDALVFAKPPPMPPFIRORNEHWFAKN--FAALQMLNKLKESDQAOVCKVLRKIPPS-----PREFPKVVMNRNPNFLNMRKFKHGTSTGFPKSSGSEVPAAPFKVGRHSGAILPFGSSFK 166  
 ANODAR 208657499 ---MILVALSLICVWVTPCQWATVPTTPTORNEHWFAKN--FAALQMLNKLKESDQAOVCKVLRKIPPS-----PREFPKVVMNRNPNFLNMRKFKHGTSTGFPKSSGSEVPAAPFKVGRHSGAILPFGSSFK 167  
 ANOGAM 18389891 ---MVPFVLLPILGIR--VQTCAPVAGDPPANVYERCHDGLDGGPR--KCHFTWFWLGR--NAVHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 168  
 CULQUI 16225986 ---MVPFVLLPILGIR--VQTCAPVAGDPPANVYERCHDGLDGGPR--KCHFTWFWLGR--NAVHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 169  
 CULQUI 16225986 ---MVPFVLLPILGIR--VQTCAPVAGDPPANVYERCHDGLDGGPR--KCHFTWFWLGR--NAVHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 170  
 PHLDUB 112361971 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 171  
 PHLPAP 15963511 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 172  
 LULO 16225999 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 173  
 Sg-220 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 174  
 Sg-221 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 175  
 Sg-218 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 176  
 SIMNIG 26914616 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 177  
 SIMNIG 26914632 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 178  
 SIMNIG 269146472 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 179  
 SIMNIG 269146474 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 180  
 SIMVIT 197260684 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 181  
 CULUAR 215259781 ---MNPVAVGLVPLTDEAYVPPNADTWMPTTQSE--TSTAE--SVLMDWKLGR--NNHNVKCVLALGMD--EHTPTPTVPPVREARVPTVGRVNEQALGSSNAGSG--KAPVPAATVFNHVSRSRTHGQVAKI 182

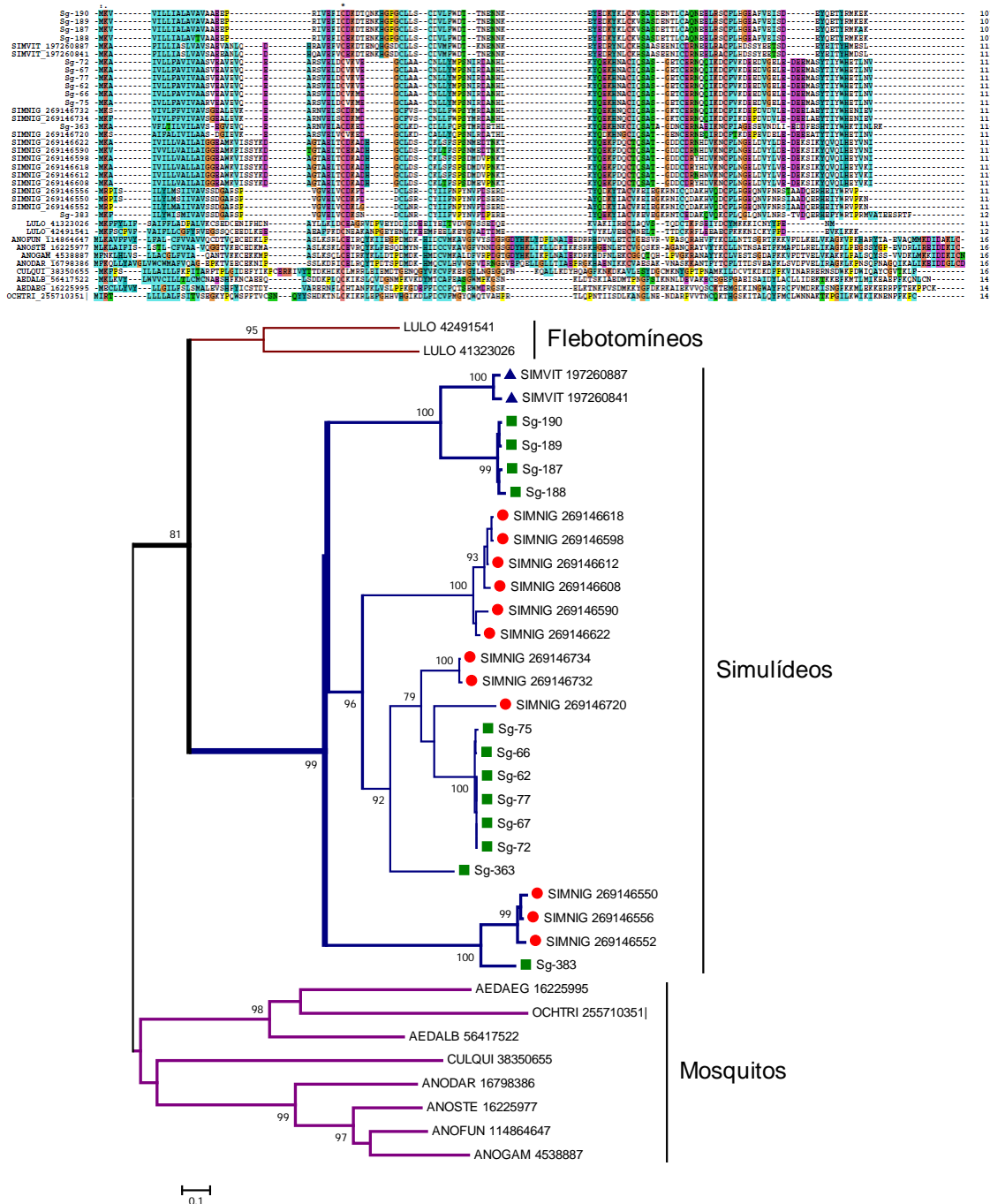


**Figura 31** - Família D7 longa de *T. guianense*. Alinhamento e dendograma das proteínas D7 longas de *T. guianense* (indicada por um quadrado e começando com Sg-). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimonômica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.

Sg-331 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKQ 166  
 Sg-335 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 168  
 Sg-332 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 166  
 SIMNIG 69146814 -----LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 162  
 SIMVIT 197260843 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 167  
 SIMNIG 269146574 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 163  
 Sg-352 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 161  
 Sg-351 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 161  
 Sg-350 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 161  
 SIMNIG 269146540 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 166  
 SIMNIG 269146538 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 166  
 SIMNIG 269146818 ---LIVLLIVGSLTNGHAGSGDAAGAAAVTTCVRINPGQVAMPSLVVGVVEGKCAFCDIKAIIVRVVHTVDQVHTWTFELINPSAAGSRQKQPIVWLDQPSALISVYKQRYTPQFPAAGKPKKCKDTPKALLLVVGCILPPTKVKR 103

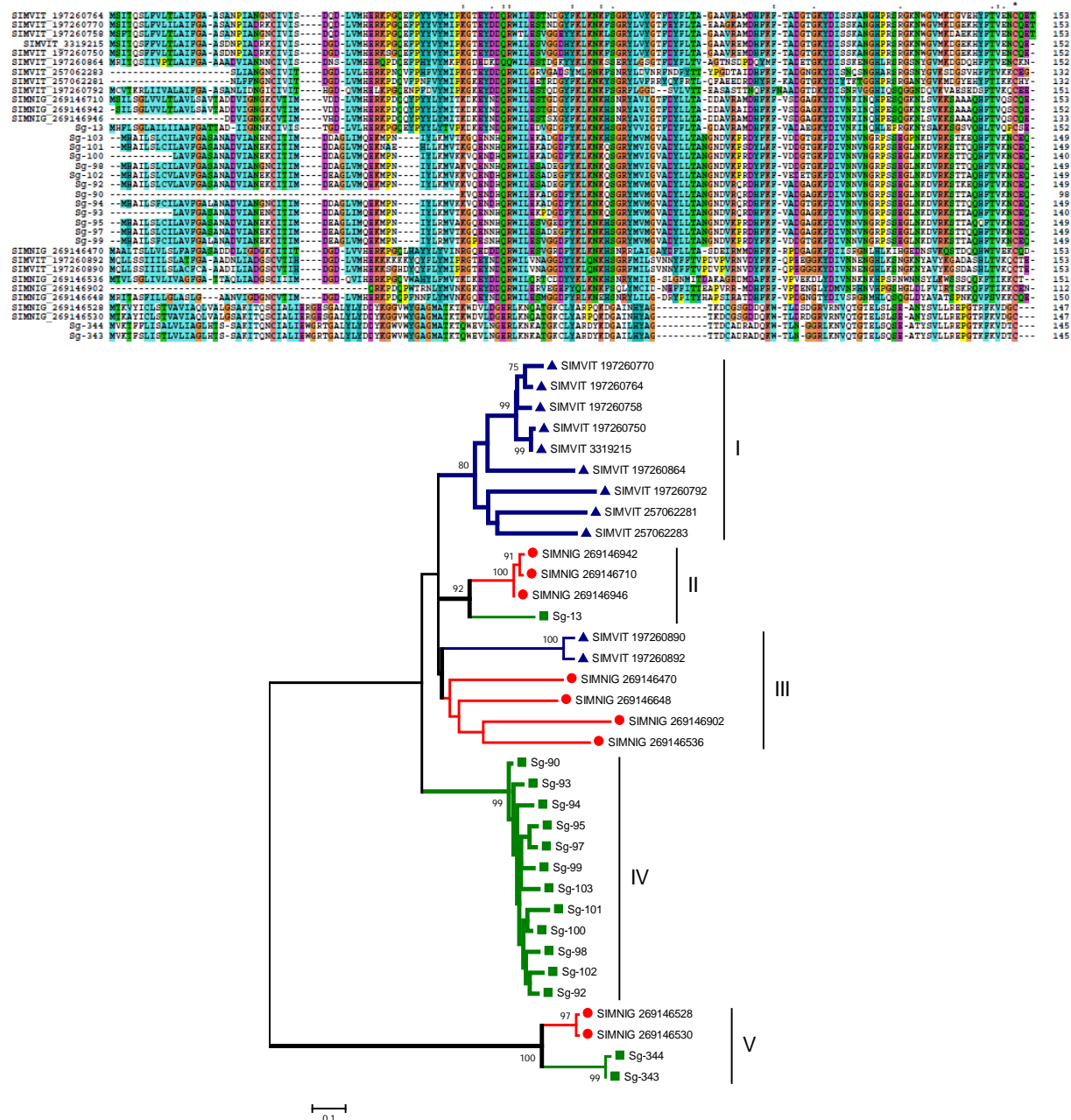


**Figura 32** - Família D7 16kDa de *T. guianense*. Alinhamento e dendograma das proteínas D7 16 kDa de *T. guianense* (indicada por Sg-). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com o pacote do programa Mega.

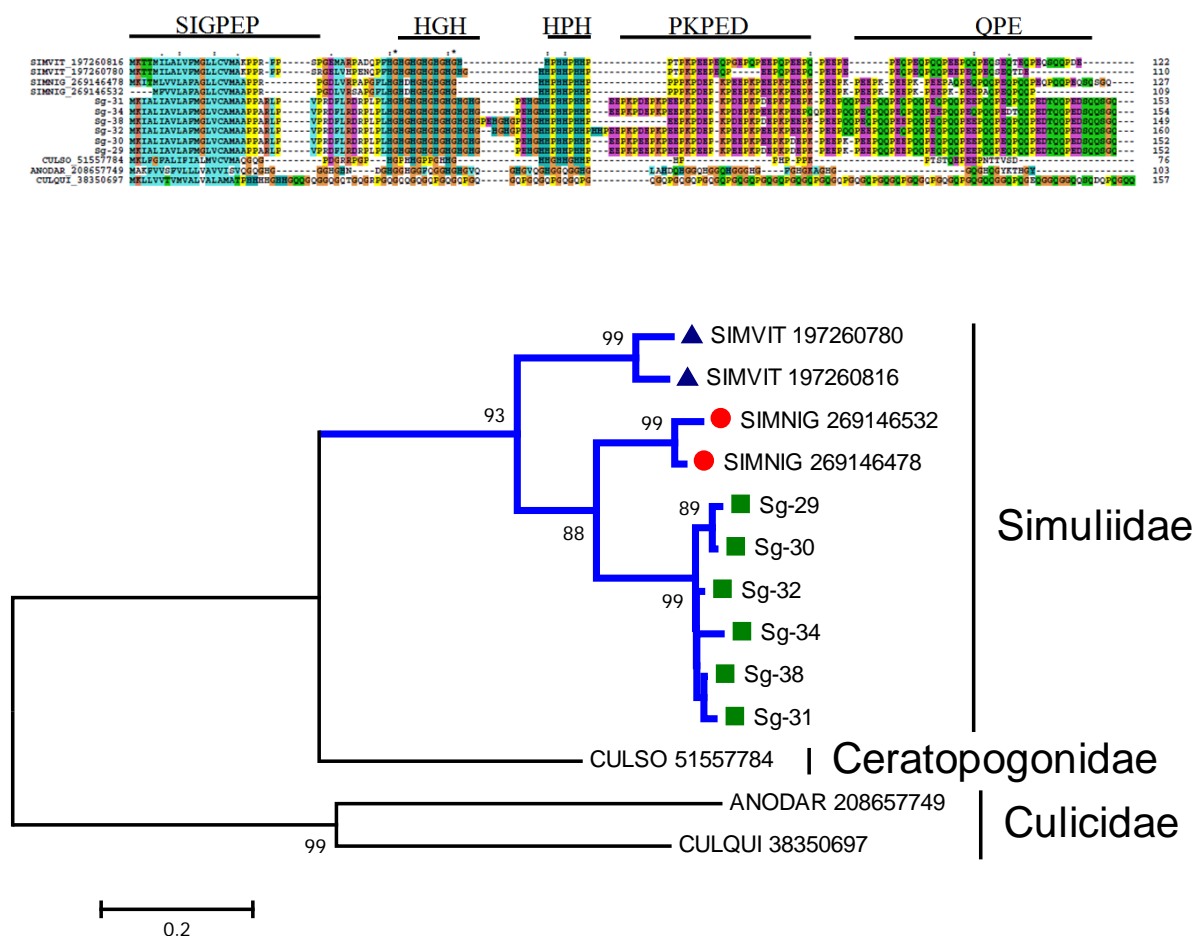


**Figura 33** - Família D7 super curtas de *T. guianense*. Alinhamento e dendograma das proteínas D7 super pequenas de *T. guianense* (indicada por Sg-). Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimonômica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.

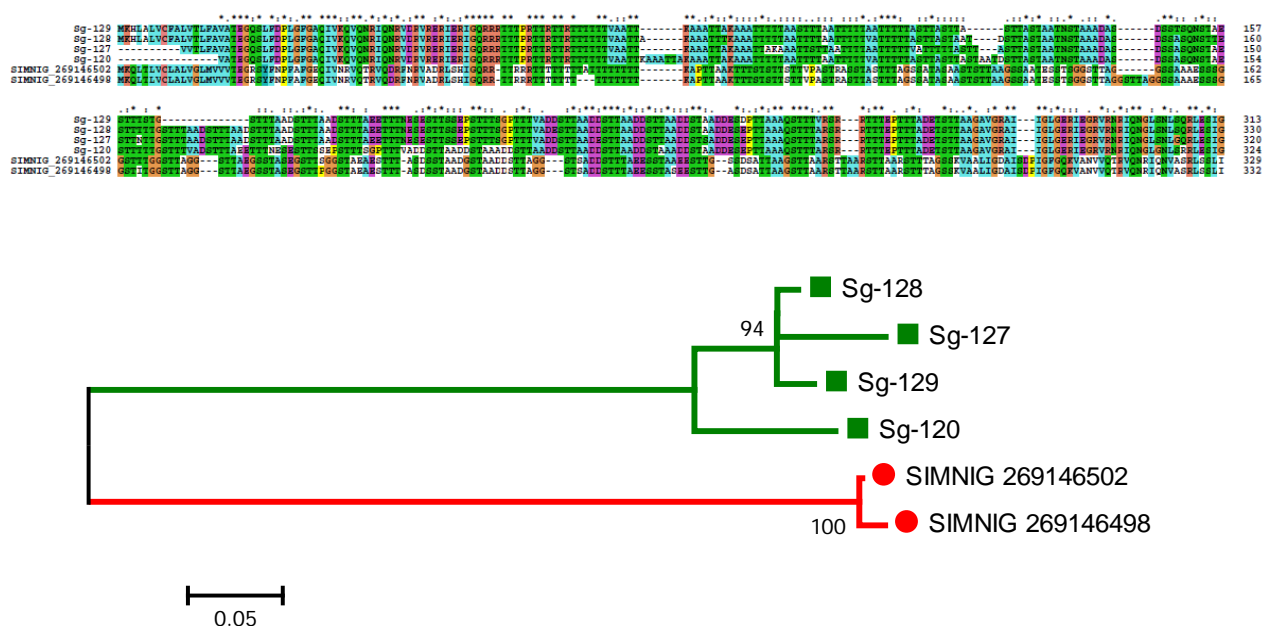




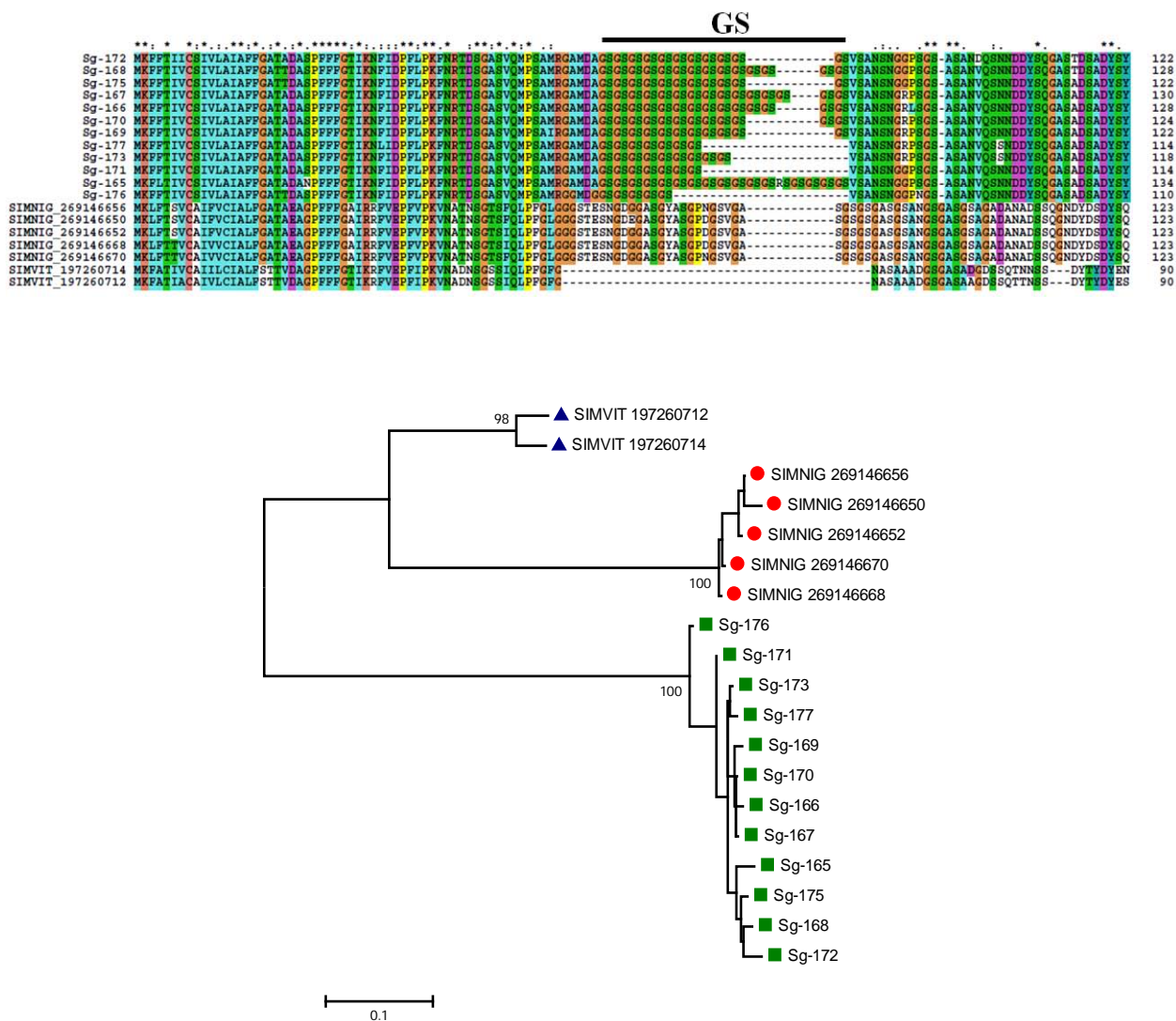
**Figura 34** - A proteína eritrema de *Simulium vittatum* (SVEP) superfamília de *Simulium*. No topo da figura está demonstrado o alinhamento Clustal de todas as proteínas SVEP derivadas do sialotranscriptoma de simulídeos. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonico acima de 75%. A barra abaixo representa 10% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendrograma foi feito com o pacote do Mega depois de 10.000 análises parcimonicas.



**Figura 35** - Proteínas de *T. guianense* ricas em His. O alinhamento Clustal está demonstrado no topo da figura com seus respectivos domínios indicados como: região com repetição de HG, (2): região com repetição HPH, (3) região com repetição HKPED e (4) região com repetição QPE. O alinhamento das proteínas salivares HP de *Culicoides sonorensis* (CULSO\_), *Anopheles darlingi* (AD\_) e *Culex quinquesfasciatus* (CQ\_) foram adicionadas com seus números de acesso no NCBI. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-, e seus homólogos *S. vittatum* por um triângulo e *S. nigrimanum* por um círculo. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 20% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



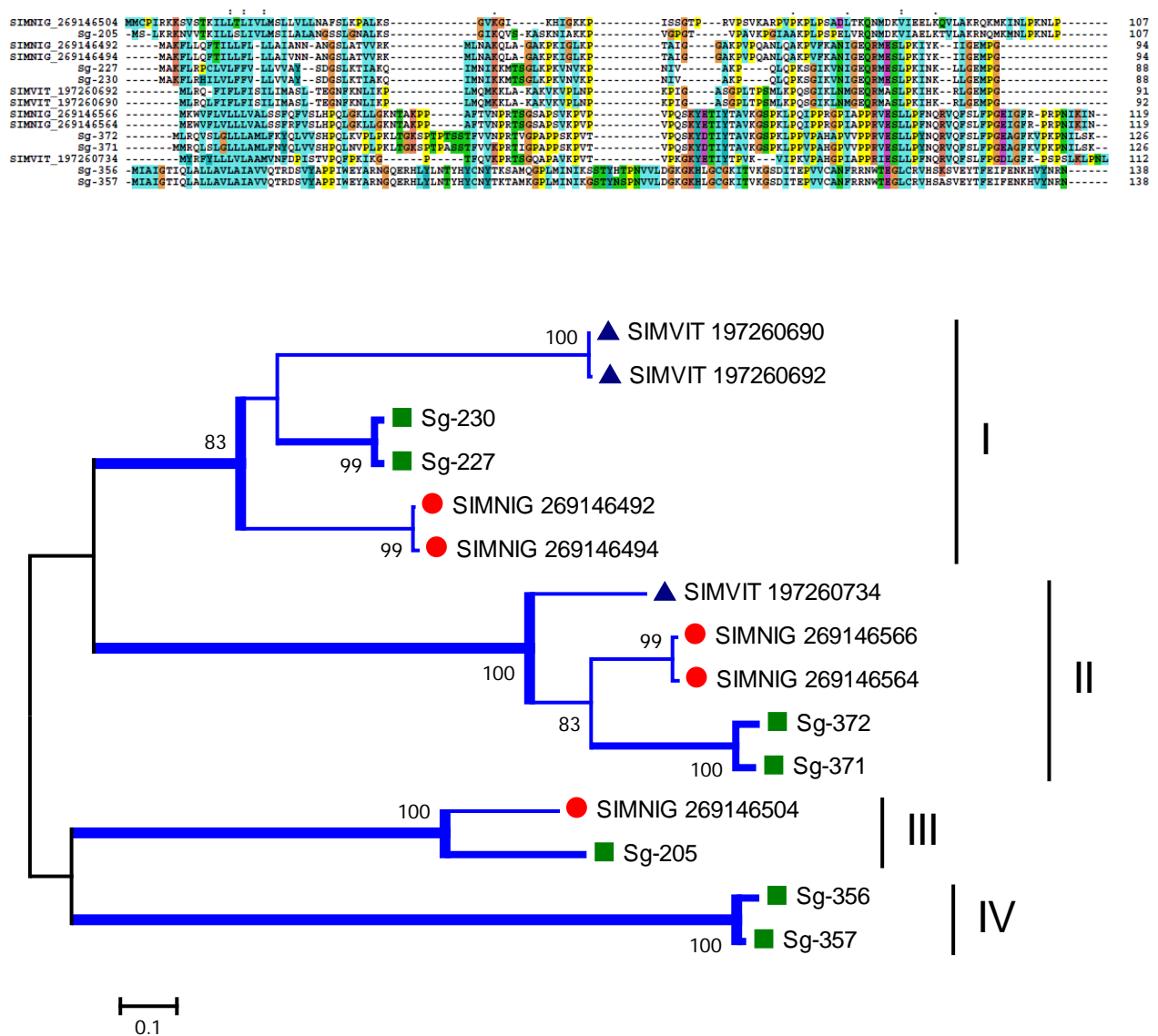
**Figura 36** - Família mucina de *Simulium* de *T. guianense*. O alinhamento é demonstrado no topo da figura e os símbolos acima indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 5% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



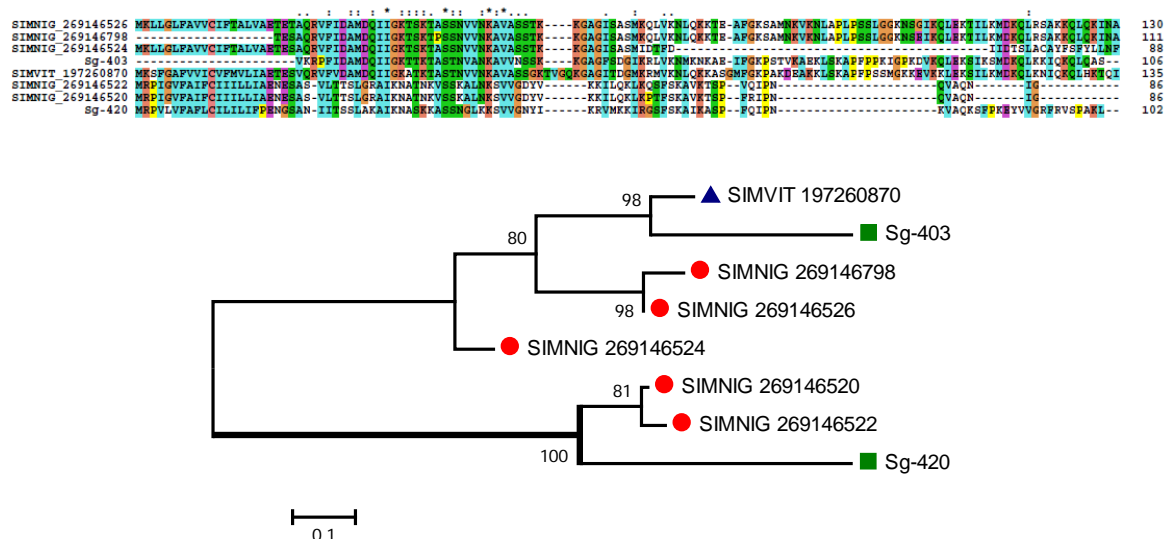
**Figura 37** - Análise filogenética de proteínas mucinas ácidas similares a família básica 7-13 *Simulium*, derivadas do sialotranscriptoma de *T. guianense*. O alinhamento é demonstrado no topo da figura. *T. guianense* está indicado por um quadrado e seus homólogos *S. nigrimanum* por um círculo e *S. vittatum* por um triângulo, os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 10% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



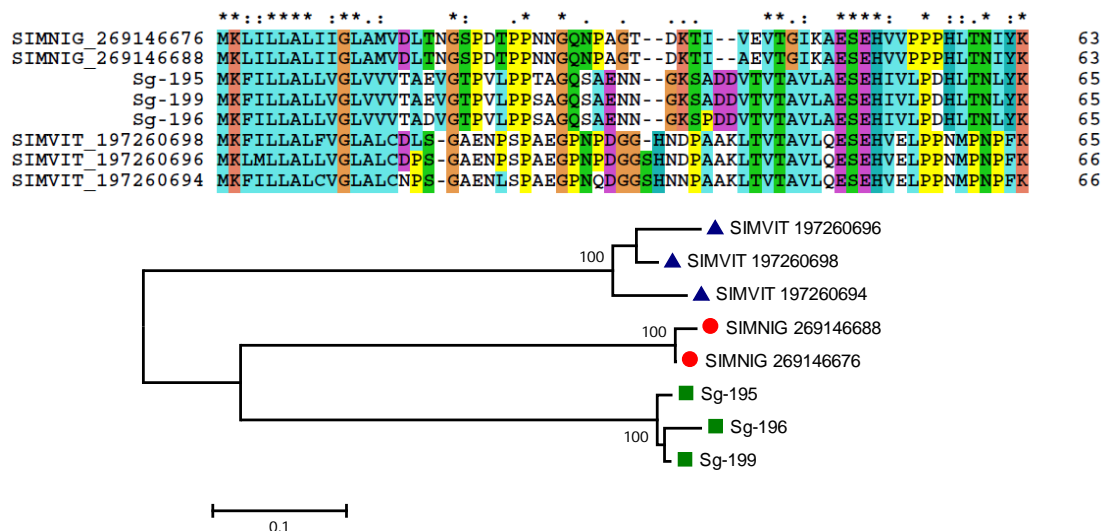




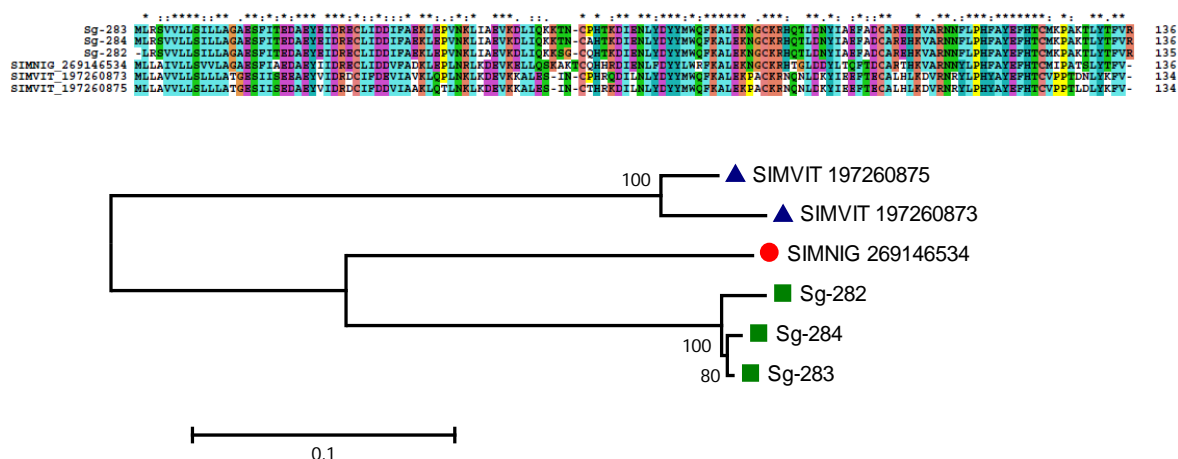
**Figura 39** - Família de proteína Sv 7,8 kDa de *T. guianense*. O alinhamento está no topo da figura. *T. guianense* está indicado por um quadrado e seus homólogos *S. nigrimanum* por um círculo e *S. vittatum* por um triângulo. os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. As proteínas de *T. guianense* estão indicadas por um quadrado e começando com Sg-. Os números nas bifurcações das árvores indicam a porcentagem de suporte parcimonômico acima de 75%. A barra abaixo representa 10% de substituição de aminoácido. As sequências das proteínas foram alinhadas com o auxílio do programa ClustalW e o dendograma foi feito com o pacote do Mega depois de 10.000 análises parcimonômicas.



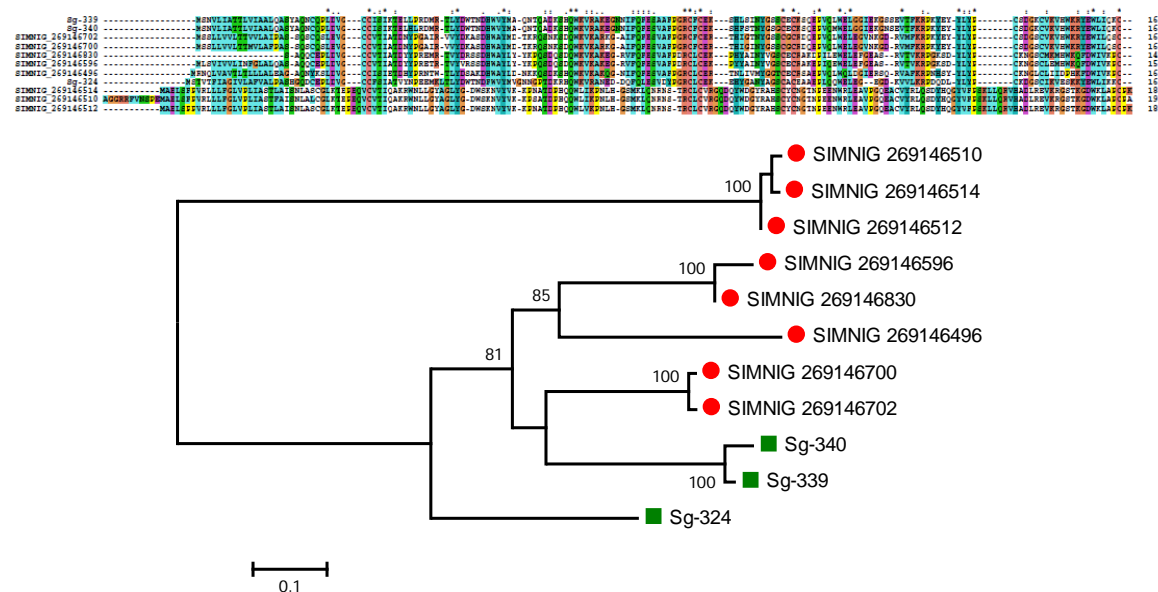
**Figura 40** - Alinhamento e dendograma da proteína básica 7-13kDa de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimônica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.



**Figura 41** - Alinhamento e dendograma da proteína 4,8kDa de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimônica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.

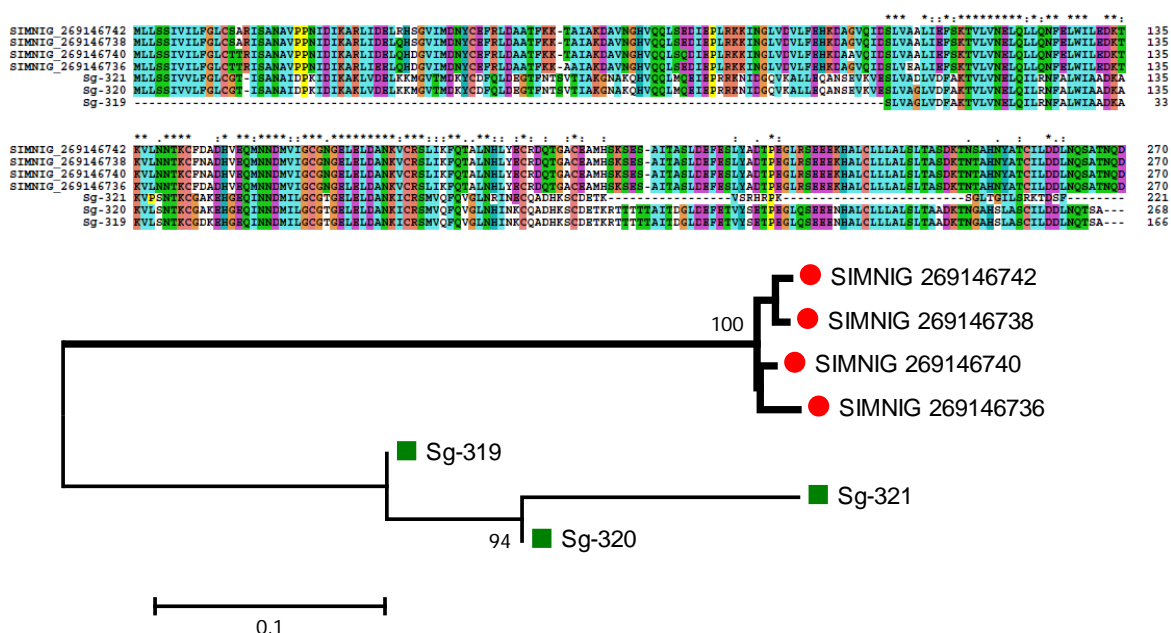


**Figura 42** - Alinhamento e dendrograma da proteína 5-Cys de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimônica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendrograma foi feito com um pacote do programa Mega.

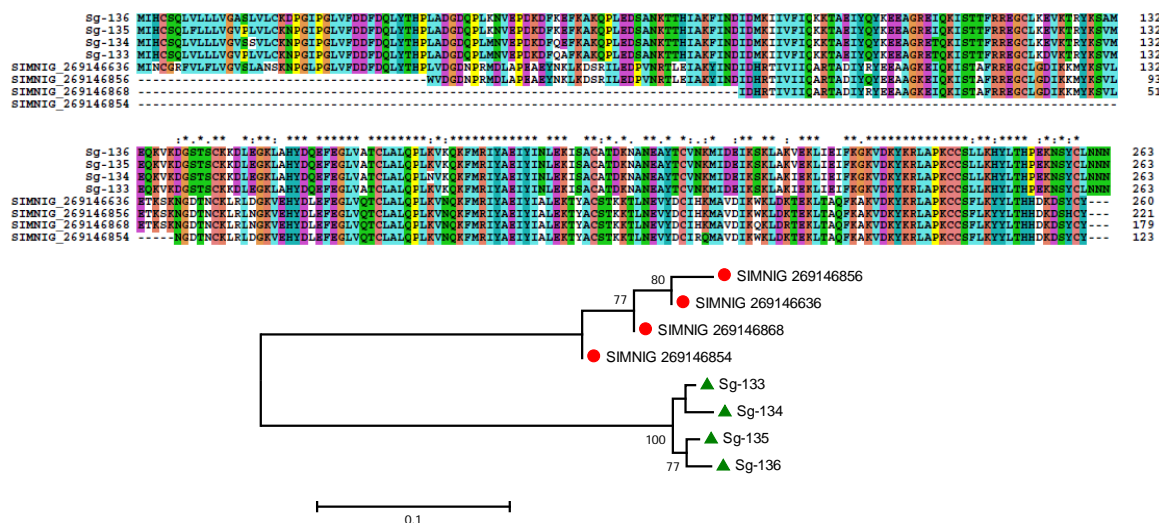


**Figura 43** - Família Sn 8-10 Cys W de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais poucos conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimônica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendrograma foi feito com um pacote do programa Mega.

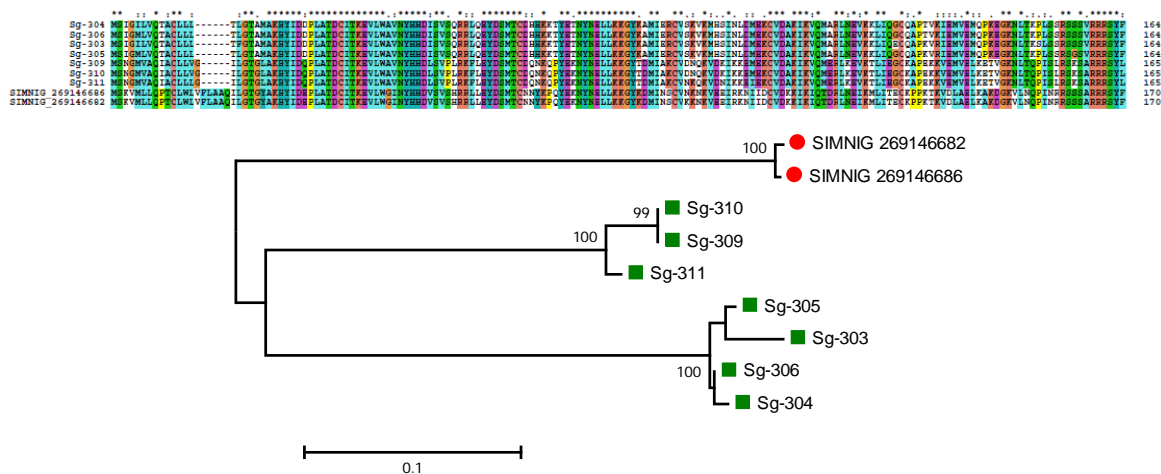




**Figura 44** - Família proteína ácida 28kDa de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.



**Figura 45** - Família de proteína básica 28kDa de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimoniosa após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.



**Figura 46.** Família de proteína 19kDa de *T. guianense*. Os símbolos acima do alinhamento indicam: (\*) locais idênticos; (:) locais conservados e (.) locais pouco conservados. Os números nas bifurcações das árvores indicam a porcentagem de suporte da análise parcimonômica após 10.000 análises, somente valores acima de 75% são demonstrados. A barra representa 10% de substituição de aminoácidos. As sequências de proteínas foram alinhadas pelo programa Clustal e o dendograma foi feito com um pacote do programa Mega.

## 6 DISCUSSÃO

As glândulas salivares de hematófagos apresentam uma grande variedade de formas com peculiaridades observadas dentro de uma mesma família ou até mesmo dentro de um mesmo gênero, como visto para mosquitos (Orr *et al.*, 1961), simulídeos (Gosbee *et al.*, 1969; Wächtler *et al.*, 1971), barbeiros (Barth, 1954; Lacombe, 1999), cimecídeos (Serrão *et al.*, 2008), flebotomíneos (Pinto, 2009) e carrapatos (Balashov, 1979). A diversidade morfológica de glândulas de hematófagos está relacionada com as distintas substâncias sintetizadas por cada grupo.

Dois tipos principais de glândulas foram descritas para hematófagos: tipo lobular com presença de cavidades secretoras para o armazenamento da saliva (em mosquitos) (Orr *et al.*, 1961; Wright, 1969; James & Rossignol, 1991) e tipo sacular com lúmen para estocar o material secretado (em barbeiros, cimecídeos e flebotomíneos) (Barth, 1954; Serrão *et al.*, 2008; Pinto, 2009). Glândulas lobulares contendo reservatório salivar são conhecidas em simulídeos (Gosbee *et al.*, 1969) e em ceratopogonídeos (Perez de Leon *et al.*, 1994). Em simulídeos a presença de reservatório salivar foi discutida em Wächter *et al.* (1971) como sendo uma glândula acessória. Nossas observações nos permitem suportar essa última afirmação. O estudo da morfologia, histologia e da histoquímica da glândula salivar tem ajudado a revelar a composição e a complexidade dos produtos secretados pela glândula principal e acessória nos ceratopogonídeos *Culicoides varripennis* (Perez de Leon *et al.*, 1994) e *Culicoides nubeculosus* (Megahed, 1956) e nos simulídeos *Boophthora erythrocephala* e *Odagmia ornata ornata* (Wächtler *et al.*, 1971).

O estudo da histologia de glândulas salivares de mosquitos dos gêneros *Aedes*, *Culex* e *Anopheles* tem mostrado que diferentes partes da glândula produzem um ou mais produtos secretores, tais como: mucopolissacarídeo pelo lobo central, o complexo carboidrato-proteína pelo corpo do lobo lateral e de mucinas neutras e proteínas ácidas no pescoço ou nas porções basais dos lobos laterais (Orr *et al.*, 1961; Wright, 1969; Janzen & Wright, 1971; Barrow *et al.*, 1975; Beckett, 1977). Essa mesma proposição foi sustentada pelo estudo do perfil protéico da saliva, onde distintas composições salivares são produzidas de acordo com a região glândular em mosquitos ou por diferentes glândulas em simulídeos (Poehling, 1979). Da mesma forma, a hibridação *in situ* dos transcritos expressos em glândulas salivares de *A. aegypti* revelaram perfis únicos de localização para um ou mais lobos glandulares bem como um novo perfil para distintas regiões subglandulares (Juhn *et al.*, 2011).

Em simulídeos pode ser conhecido sobre a biologia da glândula salivar principalmente de espécies neotropicais. Um dos primeiros estudos da glândula salivar de simulídeos retratou 14

espécies do Canadá sugerindo este órgão como útil na identificação de simulídeos (Bennett, 1963). Neste trabalho a estrutura sacular foi descrita como um reservatório salivar com um ducto salivar partindo dessa estrutura. Baseado nas imagens geradas por Bennett (1963) é evidente acreditar que toda a saliva produzida pelo lobo glandular tenha que passar por dentro do reservatório antes de ser liberada no ducto. Talvez tenha sido esse o motivo da sua designação como reservatório não-glandular visto que aspectos histológicos não foram relatados neste trabalho.

Em *T. guianense* observamos a glândula salivar como um órgão único contendo uma glândula principal e uma glândula acessória de cada lado (direito e esquerdo). Essas glândulas somente se conectam na altura dos ductos externos sugerindo que a mistura dos produtos secretados somente ocorram nessa região. Os ductos de ambas as glândulas se encontram formando um único ducto que se encontra com o do lado oposto formando o ducto comum (por onde a saliva é liberada). Com base em nossas observações morfológicas nós concordamos com as observações propostas por (Wächtler *et al.*, 1971) sobre a presença de duas glândulas salivares, principalmente porque: i) o ducto salivar não parte “exclusivamente” da glândula acessória, e sim, é formando pela fusão de dois pequenos ductos que partem de cada glândula; ii) a composição da saliva difere entre as glândulas, com cavidades secretoras altamente elétron-densa e com o lúmen da glândula acessória elétron-lucente. Na histologia observamos cavidades secretoras com maior marcação de proteína, lipídio e polissacarídeo do que no lúmen da glândula acessória (dados demonstrados somente para a glândula principal); iii) similaridade da morfologia das células da glândula acessória com as da glândula salivar de macho. Além disso, observações pessoais (dados não demonstrados) revelaram um perfil eletroforético dos homogenados salivares contendo proteínas de alto peso molecular na glândula acessória enquanto a glândula principal apresenta proteínas de baixo peso molecular. Dados similares foram observados por Wächter *et al.*, 1971, nos sugerindo que a glândula principal esteja associada a secreção de produtos salivares associados à síntese de sangue enquanto a glândula acessória secrete produtos para a síntese de açúcares. Para uma conclusão mais precisa sugerimos o estudo do perfil protéico seja efetuado para a glândula salivar de machos.

Para hematófagos com presença de glândulas lobulares como mosquitos, é conhecido que mudanças morfológicas ocorrem durante o envelhecimento glandular. Nesses insetos, logo após a emergência, a glândula é pequena, fina e desenrolada e somente com 36 h as secreções são observadas preenchendo as cavidades secretoras, com núcleos bem periféricos seguido de uma fase de redução do núcleo e do volume citoplasmático no terceiro dia (Orr *et al.*, 1961). Além

disso, o perfil protéico das glândulas salivares em algumas espécies de mosquitos não revelam variações qualitativas durante os dez primeiros dias de vida sugerindo que a diferença na quantidade de proteínas de cada idade é um reflexo do aumento de quantidade de proteínas na glândula de mosquitos mais velhos (Racioppi & Spielman, 1987; Poehling, 1979; Al-Ahdal *et al.*, 1990; Brennan *et al.*, 2000; Moreira *et al.*, 2001).

Trabalhos anteriores mostraram que as fêmeas de *T. guianense* apresentam desde o primeiro dia de vida todos os produtos protéicos na glândula salivar sem nenhuma variação até o terceiro dia após a emergência (Chagas *et al.*, 2010). Observações pessoais indicam que as fêmeas dessa espécie possuem uma preferência pela alimentação sanguínea somente a partir do segundo dia após a emergência não se conhecendo a explicação para esse fato visto que desde o primeiro momento essas fêmeas já teriam secretado todos os produtos salivares. As fêmeas de *P. papatasi* necessitam de três dias de envelhecimento glandular para sintetizar todos os seus compostos, explicando o porquê de fêmeas jovens não serem vistas picando nos três primeiros dias após a emergência (Volf *et al.*, 2000). Logo após o repasto sanguíneo ocorre um esgotamento glandular seguido de uma fase de ressíntese de um para dois dias em mosquitos (Orr *et al.*, 1961; Soliman *et al.*, 1999), e um quase completo esgotamento glandular em simulídeos com ressíntese até o quarto dia (Gosbee *et al.*, 1969). Observações pessoais para fêmeas de *T. guianense* revelaram que após o repasto sanguíneo as proteínas salivares são reduzidas quanto a sua expressão, entretanto, nenhuma é indicativa como específica para o repasto sanguíneo.

Segundo os conceitos de regressão glandular tem sido postulado que em mosquitos (Barrow *et al.*, 1975; Beckett, 1988; 1990) e ceratopogonídeos (Perez de Leon *et al.*, 1994) a diminuição do tamanho glandular é resultado da degeneração dos tecidos como resultado da reabsorção dos produtos secretores que não foram utilizados pelas fêmeas devido à ausência da hematofagia. Em simulídeos nada se conhece sobre esse aspecto.

Em muitos insetos não se conhece o estímulo para a secreção podendo ser produzida e secretada ao mesmo tempo. Sabe-se que em glândulas com presença de reservatório a secreção é regulada podendo ser do tipo endócrina e neuronal (Chapman, 1998; House & Ginsborg, 1985). Em mosquitos não se conhece bem sobre a habilidade da fêmea em controlar a alimentação. Acredita-se que possa existir um controle diferencial dirigido por estimulação nervosa para a região proximal e estimulação hormonal para região distal (James & Rossignol, 1991). Em *A. aegypti*, a glândula salivar parece ser estimulada pelo neurotransmissor serotonina com estímulos originados do nervo ventricular do sistema nervoso estomogástrico (Novak *et al.*, 1995). Em

simulídeos não se conhece o estímulo para salivação, entretanto, em fêmeas de *T. guianense* nós encontramos uma grande quantidade de fibras nervosas no epitélio da glândula acessória e na região anterior da glândula principal justamente na região aonde observamos a presença de um ducto interno. É possível que o estímulo para salivação nessa espécie seja realizada por essas fibras nervosas.

A glândula principal de *T. guianense* apresenta epitélio ligeiramente túrgido logo após a emergência com maior evidência após o segundo dia. Nesse momento, observamos na histologia e no MET uma grande quantidade de vesículas secretoras, mitocôndrias, redução do volume citoplasmático, núcleos basais comprimidos contra a membrana celular, núcleo com nucléolo aparente e cromatina condensada, além de cavidades secretoras extracelulares elétron-densas circundadas por uma grande quantidade de microvilosidades e invaginações de membrana provavelmente associadas com liberação da saliva produzida pela célula. Essa saliva revelou ser rica em substâncias protéicas, lipídicas e glicídicas. Todas essas características são típicas de célula com intensa atividade secretora.

Nessa mesma idade a glândula acessória aparenta completar sua síntese com aumento do volume celular revelado pelo intenso intumescimento glandular. As células dessa glândula são pequenas com citoplasma rico em mitocôndrias e núcleo pequeno com nucléolo aparente e cromatina condensada. Uma grande quantidade de microvilosidades foram observadas voltadas para o lúmen que apresenta produtos secretores de caráter elétron-lucente, e histologicamente revelou menor marcação protéica, lipídica e glicídica, sugerindo uma secreção diferente da sintetizada pela glândula principal. Após o quarto dia observamos uma diminuição do tamanho da glândula e um maior enrugamento das duas glândulas em fêmeas que não realizaram o repasto sanguíneo.

Em simulídeos não tem sido relatado a forma como a saliva chega no ducto externo. Em mosquitos, sabe-se que anofelinos possuem um pequeno ducto somente na região anterior enquanto culicíneos apresentam um ducto cuticular ao longo de toda a região lobular (James & Rossignol, 1991). Em fêmeas de *T. guianense* o ducto interno somente está presente na região anterior da glândula similarmente a anofelinos, enquanto a glândula acessória apresenta um grande lúmen contendo um ducto que se conecta com o da glândula principal formando o ducto individual que se irá com o oposto liberando assim a saliva.

Com o auxílio da microscopia laser confocal nós confirmamos a ausência de um ducto interno ao longo de toda a glândula principal e observamos a presença de pequenos ductos interligando as cavidades secretoras ao longo da glândula. Paralelamente produzimos fraturas nas

glândulas e observamos sua anatomia interna em MEV, que revelaram: i) a presença de várias aberturas que aparentam ser os ductos que saem de cada cavidade secretora; ii) presença de vesículas na glândula principal indicando secreção merócrina e/ou apócrina; iii) saliva com consistência frouxa na glândula principal e acessória e iv) saliva com consistência compacta no ducto individual e comum. Esses dados nos sugerem que a saliva secretada por fêmeas de *T. guianense* seja carregada através desses pequenos ductos até a região proximal da glândula, e que, a característica diferenciada da secreção seja resultado da mistura de diferentes produtos secretores na altura do ducto. A presença de secreção apócrina e/ou merócrina indica que a secreção salivar possa conter proteínas tipicamente intracelulares ou de membrana, além das que são oriundas de proteínas com peptídio sinal.

Uma ferramenta importante no estudo de glândula salivares é o emprego de lectinas fluorescentes. Nos últimos 20 anos, o uso de lectinas tem sido empregado na detecção de carboidratos na superfície celular de glândulas salivares de insetos como *A. aegypti* (Perrone *et al.*, 1986), *Anopheles stephensi* e *Anopheles albimanus* (Mohamed *et al.*, 1991), *Glossina* spp (Okolo *et al.*, 1990) e espécies do gênero *Phebotomus* e *Lutzomyia* (Volf *et al.*, 2000). Em *A. aegypti*, a lectina Con A foi primeiramente observada como útil para o reconhecimento do lobo medial (que representa o local de penetração de esporozoítos de *Plasmodium gallinaceum*) (Perrone *et al.*, 1986). Posteriormente, dados similares foram relatados para o complexo *Anopheles gambiae* (Molyneux *et al.*, 1990), fortalecendo a hipótese de correlação de marcação de lectina com taxas de penetração do esporozoíto nas glândulas. Em flebotomíneos o estudo de lectinas em glândula salivar revelou a localização de populações celulares em glândulas de *Lutzomyia migonei*, assim como, revelou subpopulações que produzem o vasodilatador Maxadilan em glândulas salivares de *L. longipalpis* (Lerner & Shoemaker, 1992).

No nosso trabalho o emprego de lectinas foi importante para evidenciar aspectos da glândula como: a secreção, os pigmentos da membrana, a delimitação das cavidades secretoras, o ducto interno na região proximal da glândula principal e a delimitação celular em glândulas de machos. A forma de ácinos do citoesqueleto glandular também foi evidenciada por actina e os núcleos foram revelados pelo marcador nuclear DAPI. O tamanho e quantidade dos núcleos da glândula principal e acessória mostram similaridade da última com a glândula de macho.

Previamente tem sido proposto o uso da glândula salivar de simulídeos como uma ferramenta útil na identificação de espécies de simulídeos (Bennett, 1963). Essa proposta foi baseada nas características: tamanho da glândula, padrão e localização da pigmentação, forma da glândula acessória, comprimento da região distal e proximal, tamanho das células e dos núcleos,

ausência e/ou presença da constrição, assim como sua característica e localização em relação à região distal. Caracteres como a cor da pigmentação e a presença/localização da constrição foram sugeridos como úteis na identificação de espécies ornitófilas e mastofílicas (Bennett, 1963). Essa proposta aparentemente parece ser consistente, entretanto, nós observamos que alguns caracteres propostos podem variar com o envelhecimento glandular, como o tamanho da glândula e padrão/localização da pigmentação.

As principais colorações conhecidas pela pigmentação da membrana glandular em simulídeo é marrom-ouro e preto-cinza. Entretanto, uma novo padrão de pigmento foi observado em *Gomphostilbia asakoe* que revelou pigmentos marrom-avermelhado (Jariyapan *et al.*, 2006). Em *T. guianense* observamos pigmentos marrom-ouro, que pela proposta de Bennett nos indicariam essa espécie como ornitófila, entretanto, é sabidamente conhecido que fêmeas de *T. guianense* são extremamente antropofílicas.

Acreditamos que a variação da pigmentação está envolvida com o envelhecimento glandular, visto que em *T. guianense* a pigmentação foi observada mais concentrada em glândulas de indivíduos mais velhos. Exemplares examinados após a alimentação sanguínea revelaram um perfil mais uniforme na pigmentação. Observações pessoais revelaram que a glândula acessória difere entre espécies antropofílicas e zoofílicas. Na primeira, a glândula acessória possui forma sacular semelhante a pequenos lobos fundidos, enquanto que na espécie zoofílica ela apresenta-se completamente sacular.

Postula-se que o comportamento da alimentação sanguínea tenha evoluído independente várias vezes nos hematófagos resultando numa diversidade de coquetéis salivares para atuar em harmonia ou não com seus hospedeiros vertebrados (Ribeiro, 1995; Ribeiro & Francischetti, 2003). No momento da alimentação sanguínea o inseto precisa vencer o tripé hemostasia, inflamação e imunidade dos hospedeiros vertebrados. Sabe-se que os hematófagos elaboraram pelo menos um anticoagulante, um antiplaquetário e um vasodilatador na saliva (Law *et al.*, 1992; Ribeiro, 1987; 1995; Ribeiro & Francischetti, 2003).

Graças ao avanço de técnicas de biologia molecular vários compostos salivares vem sendo identificados e caracterizados resultando no conhecimento atual sobre o mecanismo, função das proteínas salivares de vários de hematófagos de importância médica (Ribeiro & Francischetti, 2003; Mans & Francischetti, 2010).

Os anticoagulantes estão entre os anti-hemostáticos de hematofágos mais conhecidos, e seu estudo tem revelado diferentes classes e até mesmo mais de uma classe dentro da mesma família de artrópodes (Koh & Kini, 2009) como visto em gêneros de mosquitos (Stark & James,



1996; 1998; Francischetti *et al.*, 1999; Valenzuela *et al.*, 1999), barbeiros (Friedrich *et al.*, 1993; Noeske-Jungblut *et al.*, 1995; Van de Loch *et al.*, 1995; Pereira *et al.*, 1996; Fuentes-Prior *et al.*, 1997; Mende *et al.*, 1999; Campos *et al.*, 2002) e carrapatos (Waxman *et al.*, 1990; Gaspar *et al.*, 1996; Wei *et al.*, 1998; Horn *et al.*, 2000; Ciprandi *et al.*, 2006; Macedo-Ribeiro *et al.*, 2008).

Em simulídeos tem sido observado diferentes estratégias anti-hemostáticas que variam de acordo com presença/ausência ou alta/baixa atividade biológica, algumas das quais são sugestivas com o status vetorial do inseto para a oncocercose (Abebe *et al.*, 1994, Cupp & Cupp, 1997). O nosso trabalho reafirmou os achados de Chagas *et al.*, 2010 sobre a presença de anticoagulantes em homogenados salivares de fêmeas de *T. guianense*. Observamos que a glândula salivar contém o dobro da quantidade de proteína em relação às outras espécies de simulídeos previamente descritas na literatura (Abebe *et al.*, 1994). Nossos dados mostraram que 0,4 de HGS é capaz de prolongar em 6 vezes o tempo de formação do coágulo em relação ao controle. Além disso, o uso de 3 HGS é responsável por inibir a coagulação pois mais de uma hora.

No intuito de saber qual o alvo dessa inibição nós avaliamos os homogenados salivares em ensaios de rotina clínica para a avaliação da via intrínseca (tempo de tromboplastina parcial ativada) e a avaliação da via extrínseca (tempo de protrombina) resultando na inibição de ambas as vias. Esses resultados sugeriram que os componentes alvos da inibição da coagulação são fatores presentes na via comum. Desse modo, realizamos ensaios cromogênicos para o fXa e trombina confirmando a inibição de fXa, mas não de trombina nos HGS dessa espécie. Nossos resultados corroboram com os achados de Abebe *et al.*, 1994 sobre a presença de inibidores de fXa e não de trombina para espécies antropofílicas e com status vetorial confirmado para oncocercose. Provavelmente a inibição de fXa em *T. guianense* está associada a proteínas da família Kunitz assim como sugerido para ceratopogonídeos (Perez de Leon *et al.*, 1998) e em *Ixodes scapularis* (Francischetti *et al.*, 2002), uma vez que essa família de proteína foi encontrada no sialotranscriptoma de *T. guianense* (detalhes na tabela suplementar 1 e 2 - anexo I).

Proteínas com domínio Kunitz estão associadas com inibidores de proteases e tem sido encontradas no sialotranscriptoma de simulídeos (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010a) e ceratopogonídeos (Campbell *et al.*, 2005; Russell *et al.*, 2009) e ausentes em mosquitos, flebotomíneos e em barbeiros. Proteínas com domínio Kunitz contendo polipeptídeos também são abundantes em sialotranscriptomas de carrapatos (Francischetti *et al.*, 2005; Ribeiro *et al.*, 2006; Alarcon-Chaidez *et al.*, 2007). Esses inibidores podem atuar em pontos específicos da cascata da coagulação, principalmente contra trombina e fXa. Múltiplos domínios Kunitz podem ocorrer em uma única proteína proporcionando dessa forma a inibição de várias proteases na

cascata de amplificação, assim como ocorre com o inibidor da via do fator tissular de vertebrado da via extrínseca da cascata de coagulação (Bajaj *et al.*, 2001; Francischetti *et al.*, 2008)

A atividade de inibição de fXa tem sido relatada para várias espécies de simulídeos mas sua natureza molecular permanece desconhecida (Jacobs *et al.*, 1990). O sialotranscriptoma de *T. guianense* contém uma típica proteína Kunitz deduzida de 3 ESTs (Sg-395) e com maior identidade para seus homólogos *Simulium*. Estudos de caracterização dessa proteína recombinante faz-se necessário para a confirmação da inibição de fXa em HGS de *T. guianense*. Embora HGS tenha inibidor de fXa não podemos descartar a hipótese que outros anticoagulantes estejam presentes na saliva dessa espécie.

O sialotranscriptoma de *T. guianense* possui transcrito para o domínio inibidor de protease – serpina. Serpinas ainda não tinham sido encontradas nos sialomas de *S. vittatum* e *S. nigrimanum* possivelmente porque esse insetos realmente possam estar usando as proteínas com domínio Kunitz como anticoagulantes. As serpinas foram previamente relatadas em sialomas de *A. aegypti*, *Aedes albopictus* e *Ochelotatus triseriatus* e *L. longipalpis*. O gene que codifica para essa proteína gi|3411116 em *Ae. aegypti* representa o principal anticoagulante salivar em mosquitos com especificidade para fXa (Stark & James, 1995; 1996) tendo como homólogo a proteína gi|56417456 em *A. albopictus*.

O surgimento do sialoma permitiu que a saliva de muitos hematófagos fosse estudada, uma vez que a confecção de um transcriptoma requer uma menor quantidade de material biológico (cerca de 20 a 50 glândulas), o que facilita em casos de espécies não colonizadas como *T. guianense*. O sialotranscriptoma de *T. guianense* representa o terceiro para a família Simuliidae e o primeiro para um vetor de *O. volvulus* que resultou no sequenciamento de 1.974 clones com 89% das sequências apresentando resultados de boa qualidade. Vários grupos de sequências relacionadas foram formadas das quais 491 continham apenas uma EST. Nesse sialotranscriptoma observamos uma maior expressão dos produtos secretores em relação aos outros sialotranscriptomas de simulídeos com um incremento de 46% em relação a *S. vittatum* e de 21,4% em relação à *S. nigrimanum*. Esse aumento de expressão também foi observado para os subgrupos celulares relacionados com a síntese protéica e metabolismo energético.

Elemento de transposição foi encontrado no sialotranscriptoma de *T. guianense* e provavelmente reflete transcritos regulatórios. Uma boa proporção desses transcritos podem ser derivados de regiões 3' ou 5' não traduzidas de genes das categorias secretores e celulares. O aumento de expressão de determinadas famílias de proteínas ou de subgrupos celulares também tem sido observado em sialotranscriptomas de outros hematófagos (Francischetti *et al.*, 2002).

Esse aumento pode ser reflexo da rápida evolução de proteínas salivares por duplicação gênica impulsionada pela necessidade de evitar a resposta imune do hospedeiro que poderia inativar uma única proteína conservada e assim, inibir a alimentação sanguínea (Calvo *et al.*, 2006; Arca *et al.*, 2007).

Essa terceira experiência com *T. guianense* confirmou quase todas as famílias de proteínas ubíquas previamente relatadas para a família Simuliidae. Em geral, os sialotranscriptomas de hematófagos tem revelado famílias de proteínas ubíquas algumas das quais são altamente expressas em uma determinada ordem de insetos, como as lipocalinas em barbeiros (Assumpção *et al.*, 2008) ou podem ser específicas para algumas famílias de dípteros, como a Aegyptina e a família de proteína D7 em mosquitos (Calvo *et al.*, 2004; 2007) e simulídeos (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010a). Nos sialotranscriptomas de simulídeos, e em especial em *T. guianense*, observamos que mais de 60% de seus transcritos são específicos para essa família com uma grande quantidade de famílias de proteínas comuns a *Simulium* e outras somente encontradas em espécies antropofílicas.

Um grande número de enzimas foram encontradas no sialotranscriptoma de *T. guianense* como as amilases e maltases. Essas enzimas ajudam na digestão de carboidratos e são comumente encontradas em sialotranscriptomas de Nematocera incluindo mosquitos, ceratopogonídeos, flebotomíneos e simulídeos (Ribeiro *et al.*, 2010b). Essas enzimas podem ser reconhecidas pelo domínio0471 da KOG denominada alfa-amilase. O proteoma do mosquito *An. gambiae* possui 17 membros dessa família um do qual (AGAP002102) é expresso na glândula salivar (Arca *et al.*, 2005). O proteoma de *Ae. aegypti* contém 24 dessas enzimas, onde pelo menos duas são expressas na glândula salivar (Ribeiro *et al.*, 2007), enquanto *C. quinquefasciatus* possui 35 dessas enzimas com duas expressas na glândula salivar. Interessantemente, as enzimas de mosquito vista no filograma de *T. guianense* (clados II e III) tem sido todas previamente descritas em transcriptomas salivares sugerindo uma origem comum dessas enzimas que hidrolizam açúcares nos ancestrais comuns que deram origem a mosquitos, simulídeos e ceratopogonídeos. Esses resultados indicam que as duas sequências de *T. guianense* aparentam ser um produto de duplicação gênica antiga, onde o grupo Sg-296 do clado I é o mais ancestral das enzimas do grupo de Brachycera, enquanto os dois produtos gênicos salivares de mosquitos Culicine aparentam ser derivados de uma duplicação gênica depois da separação dos Culicidae.

O sialotranscriptoma de *T. guianense* também apresentou transcritos para as enzimas serino proteases. Essas enzimas também são ubíquas em sialomas de insetos hematófagos (Ribeiro & Arca, 2009), exceto em flebotomíneos onde somente foi encontrado em sialoma de

*Phlebotomus ariasi* (Oliveira *et al.*, 2006). Essa família tem um importante papel no sistema immune, atuando como ativadores de profenoloxidase, ou digerindo componentes da matrix extracelular, assim como pode atuar com função de elastase, ou hidrolizando componentes da coagulação sanguínea do hospedeiro como fibrinogênio/fibrina, ou ativando plasminogênio (Calvo *et al.*, 2010; Francischetti *et al.*, 2003). Em *Ae. aegypti* foram encontrados transcritos que codificam serino proteases com domínio CUB, indicando especializado reconhecimento do substrato. Todos os sialotranscriptomas de simulídeos já estudados apresentam serino proteases, e no filograma, as serinos proteases de *T. guianense* agrupam sub-clados específicos indicando conservação dos genes salivares expressos em simulídeos.

No sialotranscriptoma de *T. guianense* também foi encontrado a enzima hialuronidase responsável por clivar o ácido hialurônico (AH), que é o componente principal da matrix extracelular de vertebrados. Essa enzima foi primeiramente descrita na saliva de *L. longipalpis* (Charlab *et al.*, 1999) e posteriormente na saliva de várias outras espécies de flebotomíneos do Novo e do Velho Mundo (Cerna *et al.*, 2002; Volfova *et al.*, 2008) e de *S. vittatum* (Ribeiro *et al.*, 2000a). Hialuronidases também tem sido descritas no sialotranscriptoma de *C. quinquesfasciatus* (Ribeiro *et al.*, 2004) e *Glossina morsitans morsitans* (Alves-Silva *et al.*, 2010). Interessantemente, homogenados salivares de *P. papatasi* e *Phlebotomus dubosqui* possuem atividade de hialuronidase mas nenhum transcrito foi encontrado nas bibliotecas de cDNA dessas espécies (Cerna *et al.*, 20002). Transcritos de hialuronidase também foram ausentes em sialotranscriptomas de *S. vittatum* e *S. nigromanum* (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010a). A hialuronidase de *T. guianense* revelou identidades acima de 43% para hialuronidases de *L. longipalpis* e *Phlebotomus arabicus* além de similaridades para outras enzimas de insetos como de *Pediculus humanus* e alguns vespídeos, entretanto essas sequências de não-dípteros foram somente 34% idênticas a nível de aminoácido indicando uma rápida evolução dessa família de proteína.

O nosso trabalho não pesquisou atividade de apirase nos HGS, entretanto, genes codificantes para a proteína apirase foram encontrados no sialotranscriptoma de *T. guianense*. A apirase é uma proteína responsável pela hidrólise de ATP e ADP em AMP e ortofosfatos. Essas proteínas são comumente encontradas em sialotranscriptomas de artrópodes hematófagos onde ela tem sido sugerida como um caso de evolução convergente (Ribeiro *et al.*, 2010a). Como o ADP e ATP são importantes ativadores de plaquetas e neutrófilos, a atividade apirásica remove esses agonistas da hemostase e da inflamação (Ribeiro & Francischetti, 2003). A atividade apirásica tem sido identificada em homogenados salivares de vários hematófagos e os genes

codificadores para esta enzima tem sido confirmados através de seus sialotranscriptomas (Champagne *et al.*, 1995; Valenzuela *et al.*, 1998; 2001; Faundry *et al.*, 2004; 2006; Sun *et al.*, 2006; Andersen *et al.*, 2007). Em mosquitos a expressão dessa enzima tem ajudado a entender um pouco mais sobre a preferência alimentar pois como *Culex* tem aves como principais fontes de sangue e não enfrentam a barreira plaquetária, membros desse gênero revelam baixa ou ausência de expressão dessa enzima (Ribeiro, 2000).

A atividade apirásica foi previamente descrita em homogenados salivares de várias espécies de simulídeos com diferentes níveis de antropofilia ou zoofilia, ciclo gonotrófico e diferentes posições como vetor ou não de filárias *O. volvulus* revelando uma atividade dependente de íons  $\text{Ca}^{+2}$  ou  $\text{Mg}^{+2}$  para sua ativação. Nos simulídeos foi confirmada uma associação positiva da apirase com espécies de posição vetorial comprovada para oncocercose (Cupp *et al.*, 1993; 1995).

Os estudos de sialotranscriptomas de simulídeos vem revelando um aumento na expressão de transcritos da enzima apirase em *S. nigrimanum* (antropofílico e não vetor) quando comparado com *S. vittatum* (zoófilico), sendo estatisticamente diferente  $P=0,00337$  (Ribeiro *et al.*, 2010a). A proteína apirase de *T. guianense* encontrada no sialotranscriptoma não possui a região carboxiterminal para o a âncora de GPI indicando que é uma enzima secretada.

A enzima adenosina deaminase (ADA) foi pela primeira vez encontrada em sialotranscriptoma de *Simulium*. Os transcritos de ADA de *T. guianense* possuem 68% de identidade com a ADA de *A. aegypti*. Nesta espécie os homogenados salivares hidrolizam adenosina para inosina, e posteriormente para hipoxantina mais ribose, com atividade enzimática na saliva e nos homogenados salivares (Ribeiro & Modi, 2001). Em flebotomíneos a ADA recombinante de *Phlebotomus dubosqi* foi demonstrada ter potente atividade (Kato *et al.*, 2007) e em *P. papatasi* não foi encontrado transcritos para ADA em seu sialotranscriptoma, entretanto essa espécie possui adenosina e AMP na sua saliva, o qual atua como o principal vasodilatador salivar (Ribeiro *et al.*, 1999). A adenosina também é um poderoso anti-plaquetário e vasodilatador e a presença de ADA salivar deveria ser considerada não-adaptativa.

Os hematófagos também foram especialistas na elaboração de proteínas para ajudar numa alimentação rica em açúcares. Sabe-se que a esterilidade da alimentação com néctar é duvidosa, e nesse momento, o inseto encontra-se suscetível à qualquer invasão por micro-organismos (Moreira *et al.*, 1998). Diferentes gêneros de mosquitos elaboraram a proteína lisozima para atuar como a primeira linha de defesa durante a alimentação. Essa enzima é expressa em ambos os sexos e acumulada principalmente na região proximal dos lobos laterais e com alguma atividade

no lobo medial. Essa enzima possui atividade ótima em pH 6,0 em *A. darlingi* e com diferentes níveis de atividade, sendo maior em *Aedes* do que *Anopheles* e *Culex* (Rossignol & Lueders, 1986; Pimental & Rossignol, 1990; Moreira-Ferro *et al.*, 1998; 1999). A hibridação *in situ* de transcritos para essa proteína foram observadas na região proximal do lobo lateral (Juhn *et al.*, 2011) e transcritos para lisozima são comuns em mosquitos (Calvo *et al.*, 2004; 2009) e em várias espécies de insetos (Valenzuela *et al.*, 2002; Ribeiro *et al.*, 2010b). Em *Drosophila melanogaster* foram identificados sete genes de lisozimas sendo um salivar e os outros intestinais (Kylsten *et al.*, 1992; Daffre *et al.*, 1994).

No nosso trabalho verificamos a presença de atividade de lisozima em homogenados de glândulas salivares que ocorre em pH ótimo 7,0 diferente da atividade ótima vista em mosquito. Em *T. guianense* encontramos maior atividade de lisozima em glândulas de fêmeas de 48h após a emergência. Nós somente investigamos a presença em glândulas de fêmeas. O sialotranscriptoma de *T. guianense* revelou transcritos para essa enzima com vários possíveis alelos para o mesmo gene. A lisozima salivar de *T. guianense* possui 79% de identidade com seus homólogos *S. vittatum* e *S. nigrimanum* e 52% para mosquitos. Possivelmente, essa enzima desempenhe em simúldeos a mesma função relatada para mosquitos.

A família de proteína antígeno-5 pertence a superfamília CAP (Gibbs *et al.*, 2008) e tem sido associada a defesa de patógenos em plantas, com função de toxinas em cobras e em venenos de lagarto (Nobile *et al.*, 1996; Yamazaki *et al.*, 2002), como um inibidor de agregação plaquetária em tabanídeos (Xu *et al.*, 2008) e como um possível inibidor da ativação da via clássica do complemento em mosca de chifre (Ameri *et al.*, 2008; Wang *et al.*, 2009). A anotação do proteoma de *A. gambiae* revelou 21 proteínas para essa família, duas das quais são expressas nas glândulas salivares (Arca *et al.*, 2005). Similarmente, o proteoma de *A. aegypti* e *C. quinquefasciatus* possuem acima de 30 membros para essa família, dos quais pelo menos 2 são expressos nas suas glândulas salivares (Ribeiro *et al.*, 2004; 2007). As proteínas antígeno-5 de *T. guianense* são filogeneticamente próximas das proteínas de nematóceras com 70% de suporte parcimônico, entretanto, elas formam sub-clados específicos para *Simulium*, *Culicoides*, Culicine e Psychodidae. Embora os sialotranscriptomas de *Simulium* vem revelando proteínas que codificam antígeno-5 mas nenhum deles tem sido estudados molecularmente.

O sialotranscriptoma de *T. guianense* também revelou transcritos para a família Yellow, domínios ML e lipocalinas. A proteína Yellow recebeu esse nome devido a mutação de um gene que induz o fenótipo Yellow em drosófila, resultando no rompimento da formação de melanização. Em Nematóceras, essa família é abundantemente expressa em flebotomíneos e tem

sido sugerida como importante marcador de exposição vetorial (Bahia *et al.*, 2007). Proteínas com domínio ML são sugeridas para o reconhecimento de lipídeo, particularmente no reconhecimento de produtos relacionados ao patógeno, mas poderia também ter função lisosomal e celular (Friedland *et al.*, 2003). As lipocalinas são proteínas amplamente distribuídas em animais e plantas. Essa proteína é altamente expressa em triatomíneos, como em *Rhodnius prolixus*, *Triatoma infestans* (Assumpção *et al.*, 2008), *Triatoma brasiliensis* (Santos *et al.*, 2007) e *Dipetalogaster maxima* (Assumpção *et al.*, 2010). Em triatomíneos, as lipocalinas são relatadas como carreadores de óxido nítrico, ligantes de agonistas de aminas biogênicas, nucleotídeos adenosine e tromboxano, bem como inibidores de agregação plaquetária induzida por colágeno e trombina e como alérgeno (Noeske-Jungblut, 1995). A função das lipocalinas secretada na saliva pode ser de remover os lipídeos no lugar da picada.

Os sialotranscriptomas de *Simulium* apresentam proteínas que codificam a família aegyptina. Essa proteínas são encontradas somente em sialotranscriptomas de mosquitos e tem sido primeiramente nomeadas como alérgeno de 30 kDa *Aedes* e como proteínas ricas em GE (Valenzuela *et al.*, 2003). Elas apresentam altos níveis de expressão em sialotranscriptomas de *Anopheles funestus* (Calvo *et al.*, 2007). Análises funcionais de um membro da família *A. aegypti*, conhecido como Aegyptin, bem como um membro de *An. stephensi* tem demonstrado atuar como antagonista de agregação plaquetária induzida por colágeno e como ferramenta útil para inibir a interação colágeno-plaqueta em *in vitro* e *in vivo* (Calvo *et al.*, 2007; 2010; Yoshida *et al.*, 2008). A filogenia de antígeno-5 encontrada no sialotranscriptoma de *T. guianense* revelou uma origem comum em mosquitos e simúlideos.

Interessantemente os sialotranscriptomas de mosquitos e simúlideos apresentam uma grande expressão de uma família de proteínas que se ligam a odorantes. A família de proteína D7 é específica em nematóceras hematófagos sendo reconhecida como um membro da superfamília OBP (Hekmat-Scafe *et al.*, 2000), entretanto, elas contém duas alfa-hélices adicionais (Mans *et al.*, 2007; Alvarenga *et al.*, 2010). A forma longa e curta da família D7 existe onde um ou dois domínios D7 existem na mesma proteína, produzindo proteínas com peso molecular de aproximadamente ~18 ou ~28 kDa.

No sialotranscriptoma de *T. guianense* encontramos três famílias de D7 (D7 longa, D7 16 kDa e D7 ultra-curtas). A superfamília D7 é altamente expressa na glândula salivar dessa espécie apresentando o maior número de transcritos de toda a biblioteca com 242 (18%) ESTs das 1324 ESTs secretoras. As proteínas D7 de *T. guianense* não apresentam similaridade a nível de

aminoácido com nenhuma outra proteína D7 de nematócera. O estudo funcional das proteínas D7 poderia ajudar a entender o papel de sua alta expressão nesse vetor.

Das 18 famílias de proteínas específicas para *Simulium* somente uma teve sua função esclarecida (família SVEP/Marydilan). A proteína SVEP é uma família de proteína específica para simulídeos e foi primeiramente descrita em homogenados salivares de *S. vittatum*, quando foi nomeada de SVEP (*Simulium* Vasodilator Erythema Protein), devido esta proteína produzir um efeito vasodilatador prolongado quando testada em pele de coelho (Cupp *et al.*, 1994). Sua forma recombinante (rSVEP) foi expressa e funcionalmente caracterizada como um potente vasodilatador, possivelmente ativando os canais de cálcio dependente de ATP (Cupp *et al.*, 1998). Essa proteína tem um importante papel durante a alimentação sanguínea e foi sugerida como um componente chave para a competência vatorial de simulídeos na transmissão de filárias do gênero *Onchocerca* (Cupp & Cupp, 1997).

Os sialotranscriptomas de duas espécies de simulídeos identificaram SVEP dentro de uma variada família multigene, com pelo menos 5 genes para cada espécie (Andersen *et al.*, 2009, Ribeiro *et al.*, 2010a). O sialotranscriptoma de *T. guianense* revelou proteínas homólogas com identidades variando de 50 para 70% e com sequências de tamanhos similares, mas com baixo nível de aminoácidos conservados. Além disso, a análise filogenética de SVEP de *T. guianense* indica que pelo menos dois genes (Sg-13 e Sg-344) possuem um ancestral comum com *S. nigrimanum* e um terceiro gene poderia ser reflexo do aumento de expressão dessa família de proteína. Interessantemente, *T. guianense* apresentou 190 ESTs codificando para a proteína SVEP (14,4% dos transcritos da classe secretora), sendo mais do que o dobro das sequências de *S. nigrimanum* (6,8% de suas ESTs da classe secretora) e mais do que três vezes em relação à *S. vittatum* (que codificam 4,5% de suas ESTs da classe secretora para SVEP). Esses aumentos apresentam diferença estatística altamente significativa pelo  $\chi^2=16.72$  ( $P=2.3E^{-8}$ ) e  $\chi^2=72.5$  ( $P=1.6E^{-17}$ ) para *S. nigrimanum* e *S. vittatum*, respectivamente.

Uma segunda família específica de *Simulium* e bastante expressa em simulídeos é a família de proteínas conhecida pelas suas repetições de histidina, prolina, glutamina e ácido glutâmico. A natureza dessas repetições nas proteínas tem sido sugerida na interação com proteínas de matrix, possivelmente colágeno, podendo funcionar de maneira análoga para Aegyptinas de mosquito (Misumi *et al.*, 1990). O simulídeo *S. vittatum* revelou essa família como a mais expressa em seu sialotranscriptoma com quatro regiões repetidas em suas sequências (rica em arginina, repetição HG, repetição HPH e repetição QPE (Andersen *et al.*, 2009), enquanto *T.*



*guianense* apresenta todas essas regiões além de uma região a mais de PKPED. Similarmente, sialotranscriptomas de mosquitos e *Culicoides* também contém proteínas com repetições de Pro-His e Gly-His, mas nenhuma outra similaridade das sequências.

As mucinas são proteínas de baixa complexidade ricas em resíduos de serina e treonina e comumente encontradas em sialotranscriptomas de mosquitos (Calvo *et al.*, 2009; 2010), ceratopogonídeos (Campbell *et al.*, 2005), cimecídeos (Franscischetti *et al.*, 2010) e simulídeos (Andersen *et al.*, 2009; Ribeiro *et al.*, 2010a). As mucinas não tem função biológica completamente conhecida mas são sugeridas para proteção das partes internas dos ductos salivares, além de funções antimicrobianas. As mucinas são proteínas com modificações pós-traducionais que possuem em sua forma madura resíduos de N-acetil galactosamina (Hang & Bertozzi, 2005). No sialotranscriptoma de *T. guianense* nós encontramos dois tipos de mucinas: mucinas de *Simulium* e as mucinas ácidas similares para as proteínas básicas de 7-13kDa. Na primeira podemos encontrar representantes de todas as espécies de simulídeos estudadas, enquanto a segunda é específica para *S. nigrimanum* e *T. guianense*. A função dessas proteínas em simulídeos não é conhecida. Assim como as mucinas nenhuma outra proteínas específica de *Simulium* possuem função conhecida. Algumas são comuns para os outros dois sialotranscriptomas estudados e outras são comuns apenas para *T. guianense* e *S. nigrimanum*.

Pela primeira vez em sialotranscriptoma de insetos uma proteína com similaridade para kunitoxina de cobra. Essa proteína é rica em Cys mas diferentemente não apresentam um domínio Kunitz. A função dessa proteína em simulídeos não é conhecida e deveria ser investigada molecularmente. Esse terceiro sialotranscriptoma foi importante para direcionar o estudo das proteínas candidatas para os estudos funcionais devido representar uma segunda experiência com espécie antropofílica e ter resultado numa maior homologia das proteínas de *T. guianense* com *S. nigrimanum*.

Com base nas homologia das sequências de proteínas encontradas no sialotranscriptoma de *T. guianense* poderíamos investigar se realmente o domínio Kunitz é a proteína responsável pela atividade de inibição de fXa, assim como estudar a presença de atividade apirásica e vasodilatadora em homogenados salivares dessa espécie. O sialotranscriptoma de *T. guianense* disponibiliza um banco domínio público de proteínas candidatas para à confecção de vacinas contra oncocercose, além de proporcionar a busca de novos fármacos de interesse biotecnológico e farmacológico.

## 7 CONCLUSÃO

Esse trabalho revela pela primeira vez uma coletânea de dados referentes às glândulas salivares de um simuliídeo Neotropical, antropofílico e vetor de filárias *O. volvulus*, que permitiram concluir:

1 – A anatomia do órgão salivar revelou duas glândulas secretoras. Com base na sua morfologia externa, nas características das células e histologia química sugerimos que a glândula principal secreta produtos associados a uma alimentação sanguínea, enquanto a glândula acessória secreta produtos necessários para uma alimentação rica em açúcares. A função de reservatório salivar não é aceita nesse trabalho.

2 - A glândula salivar secreta uma grande quantidade de proteínas em relação à de outros simuliídeos conhecidos, que possui a atividade de inibir a coagulação sanguínea, tendo como alvo a inibição de fXa mas não de trombina. Possivelmente, essa atividade inibitória seja desempenhada pela proteína da família Kunitz, encontrada no sialotranscriptoma desse simuliídeo.

3 – Os homogenados de glândula salivar de fêmeas de *T. guianense* até o segundo dia após a emergência possuem atividade de lisozima, principalmente em pH 7,0 e maior em glândulas de 48h. Além disso, genes para essa enzima foi confirmado no sialotranscriptoma

4 - O sialotranscriptoma revelou 32 famílias de proteínas sendo várias ubíquas, duas específicas de insetos e 18 específicas de *Simulium* representando 63,6% de seus transcritos secretores. O proteoma de *T. guianense* confirmou 28 das 32 famílias de proteínas encontradas no transcriptoma, onde um total de 101 novas sequências de proteínas de interesse foram disponibilizadas para o banco de dados de proteínas da não redundante database do NCBI, com 82 sequências codificando para proteínas secretoras.

5 - Algumas famílias de proteínas como hialuronidase, adenosina deaminase, serpina, domínios ML, lipocalina foram encontradas pela primeira vez em *Simulium*.

6 - O sialotranscriptoma de glândulas salivares de *T. guianense* também representa uma banco de dados de proteínas candidatas para à confecção de vacinas contra vetores da oncocercose.

## 8 PERSPECTIVAS

As perspectivas deste trabalho consistem em:

- ✚ Selecionar e expressar genes de interesse, com função desconhecida, em sistemas heterólogos visando sua caracterização molecular e funcional, assim como a identificação de ligantes para essas proteínas (seja alvos enzimáticos ou moléculas transportadas por proteínas carreadoras).
- ✚ Produzir o sialoma de machos de *T. guianense* para facilitar a classificação funcional das proteínas unicamente encontradas em simulídeos.

## 9 ANEXO I – Tabela Suplementar I e II

| Assembled contig                        | N EST | Name of larger sequence | Comments                                    | Larger orf (nt) | Frame | Length | Poly A at |
|---|-------|-------------------------|---|-----------------|-------|--------|-----------|
| <b>Secreted products</b>                |       |                         |   |                 |       |        |           |
| <b>Ubiquitous protein families</b>      |       |                         |   |                 |       |        |           |
| <b>Antigen-5 family</b>                 |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 453</a>           | 3     | SGL_P18_F02             | antigen 5-related salivary protein          | 612             | 1     | 876    | 1058      |
| <a href="#">Sg-contig 457</a>           | 2     | SGL_P16_A06             | salivary secreted antigen 5-related protein | 855             | 3     | 959    | 941       |
| <b>Yellow family</b>                    |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 522</a>           | 1     | SGL_P13_G08             | yellow                                      | 138             | 3     | 266    | 247       |
| <b>ML domain family</b>                 |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 705</a>           | 1     | SGM_P9_C05              | Niemann-Pick Type C-2                       | 144             | 2     | 203    |           |
| <b>Lipocalins</b>                       |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 671</a>           | 1     | SGM_P7_D03              | Apolipoprotein D/Lipocalin                  | 204             | 3     | 465    | 446       |
| <a href="#">Sg-contig 568</a>           | 1     | SGL_P18_A10             | Lipocalin-like domain                       | 588             | 2     | 907    | 888       |
| <b>Immunity related products</b>        |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 263</a>           | 5     | SGL_P11_B12             | Lysozyme                                    | 471             | 2     | 547    | 529       |
| <a href="#">Sg-contig 264</a>           | 5     | SGL_P17_B06             | Lysozyme                                    | 471             | 2     | 554    | 530       |
| <a href="#">Sg-contig 266</a>           | 4     | SGL_P15_A04             | Lysozyme                                    | 474             | 1     | 563    | 549       |
| <a href="#">Sg-contig 267</a>           | 4     | SGL_P14_C01             | Lysozyme                                    | 426             | 3     | 557    | 540       |
| <a href="#">Sg-contig 265</a>           | 2     | SGM_P6_G06              | Lysozyme                                    | 474             | 2     | 559    | 541       |
| <a href="#">Sg-contig 268</a>           | 2     | SGL_P15_D01             | Lysozyme                                    | 426             | 1     | 560    | 531       |
| <a href="#">Sg-contig 270</a>           | 1     | SGL_P20_A07             | Lysozyme                                    | 441             | 2     | 530    | 511       |
| <a href="#">Sg-contig 269</a>           | 1     | SGL_P11_D01             | Lysozyme                                    | 474             | 2     | 558    | 539       |
| <a href="#">Sg-contig 480</a>           | 1     | SGL_P11_E07             | C-type lysozyme                             | 495             | 1     | 852    | 833       |
| <a href="#">Sg-contig 369</a>           | 6     | SGM_P3_C09              | salivary expressed cecropin                 | 195             | 3     | 316    | 297       |
| <a href="#">Sg-contig 368</a>           | 4     | SGM_P7_E12              | salivary expressed cecropin                 | 183             | 2     | 315    | 318       |
| <a href="#">Sg-contig 674</a>           | 1     | SGM_P7_E09              | salivary expressed cecropin                 | 129             | 2     | 259    | 240       |
| <b>Protease inhibitor domains</b>       |       |                         |   |                 |       |        |           |
| <b>Serpin</b>                           |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 500</a>           | 1     | SGL_P12_E01             | Serpin                                      | 339             | 1     | 455    | 436       |
| <b>Kunitz domain protease inhibitor</b> |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 375</a>           | 6     | SGL_P15_B04             | Very short Kunitz inhibitor                 | 357             | 3     | 469    | 414       |
| <a href="#">Sg-contig 395</a>           | 3     | SGM_P8_C06              | Serine proteinase inhibitor (KU family)     | 336             | 1     | 387    | 349       |
| <a href="#">Sg-contig 376</a>           | 3     | SGM_P4_A09              | Protease inhibitor cvp2                     | 318             | 1     | 387    | 199       |
| <a href="#">Sg-contig 409</a>           | 2     | SGM_P9_F08              | Kunitoxin-Phi1                              | 309             | 3     | 424    | 405       |
| <a href="#">Sg-contig 549</a>           | 1     | SGL_P17_B07             | Serine proteinase inhibitor (KU family)     | 195             | 1     | 289    | 270       |
| <a href="#">Sg-contig 397</a>           | 1     | SGM_P10_G02             | Serine proteinase inhibitor (KU family)     | 348             | 1     | 399    | 380       |
| <a href="#">Sg-contig 396</a>           | 1     | SGL_P11_G05             | Serine proteinase inhibitor (KU family)     | 201             | 1     | 251    | 232       |
| <a href="#">Sg-contig 399</a>           | 1     | SGM_P7_H05              | single Kunitz protease inhibitor 2          | 201             | 2     | 253    | 234       |
| <a href="#">Sg-contig 398</a>           | 1     | SGS_P1_C01              | single Kunitz protease inhibitor 2          | 222             | 3     | 276    | 257       |
| <b>Enzymes</b>                          |       |                         |   |                 |       |        |           |
| <b>Trypsin</b>                          |       |                         |   |                 |       |        |           |
| <a href="#">Sg-contig 139</a>           | 9     | SGM_P7_G02              | Serine protease                             | 771             | 2     | 823    | 801       |
| <a href="#">Sg-contig 138</a>           | 7     | SGL_P17_F11             | Serine protease                             | 798             | 3     | 852    | 834       |
| <a href="#">Sg-contig 244</a>           | 6     | SGL_P16_D02             | Serine protease                             | 750             | 3     | 803    | 781       |
| <a href="#">Sg-contig 246</a>           | 5     | SGM_P6_A09              | Serine protease                             | 750             | 1     | 804    | 1073      |
| <a href="#">Sg-contig 418</a>           | 3     | SGL_P15_D03             | salivary serine protease                    | 243             | 3     | 315    | 297       |
| <a href="#">Sg-contig 242</a>           | 3     | SGL_P15_E12             | Serine protease                             | 756             | 1     | 810    | 789       |
| <a href="#">Sg-contig 241</a>           | 3     | SGM_P3_H06              | Serine protease                             | 756             | 1     | 808    | 791       |
| <a href="#">Sg-contig 243</a>           | 3     | SGS_P1_E11              | Serine protease                             | 756             | 1     | 806    | 1097      |
| <a href="#">Sg-contig 419</a>           | 2     | SGM_P7_G04              | salivary serine protease                    | 153             | 1     | 221    | 204       |
| <a href="#">Sg-contig 248</a>           | 2     | SGL_P14_E04             | Serine protease                             | 750             | 2     | 804    | 804       |

|                                    |    |                  |                                       |      |   |      |      |
|------------------------------------|----|------------------|---------------------------------------|------|---|------|------|
| <a href="#">Sg-contig_247</a>      | 2  | SGL_P11_E03      | Serine protease                       | 750  | 3 | 795  | 785  |
| <a href="#">Sg-contig_249</a>      | 2  | SGL_P17_G07      | Serine protease                       | 759  | 1 | 798  | 779  |
| <a href="#">Sg-contig_245</a>      | 2  | SGL_P14_B04      | Serine protease                       | 750  | 3 | 800  | 782  |
| <a href="#">Sg-contig_416</a>      | 2  | SGL_P14_A10      | Serine protease                       | 819  | 1 | 863  | 909  |
| <a href="#">Sg-contig_141</a>      | 2  | SGM_P4_F12       | Serine protease                       | 411  | 1 | 745  | 441  |
| <a href="#">Sg-contig_140</a>      | 2  | SGL_P18_E11      | Serine protease                       | 423  | 1 | 864  | 569  |
| <a href="#">Sg-contig_142</a>      | 2  | SGL_P15_G05      | Serine protease                       | 210  | 1 | 546  | 534  |
| <a href="#">Sg-contig_415</a>      | 2  | SGM_P7_A08       | Serine protease                       | 504  | 1 | 710  | 545  |
| <a href="#">Sg-contig_553</a>      | 1  | SGL_P17_C11      | salivary serine protease              | 492  | 3 | 525  | 506  |
| <a href="#">Sg-contig_251</a>      | 1  | SGL_P16_H02      | Serine protease                       | 750  | 3 | 802  | 783  |
| <a href="#">Sg-contig_254</a>      | 1  | SGL_P17_A06      | Serine protease                       | 489  | 3 | 529  | 510  |
| <a href="#">Sg-contig_252</a>      | 1  | SGL_P17_F04      | Serine protease                       | 750  | 3 | 832  | 813  |
| <a href="#">Sg-contig_250</a>      | 1  | SG_P1testing_A08 | Serine protease                       | 756  | 3 | 803  | 784  |
| <a href="#">Sg-contig_253</a>      | 1  | SGL_P16_F09      | Serine protease                       | 783  | 1 | 821  | 802  |
| <a href="#">Sg-contig_145</a>      | 1  | SGL_P13_E04      | Serine protease                       | 591  | 3 | 643  | 624  |
| <a href="#">Sg-contig_146</a>      | 1  | SGL_P13_H08      | Serine protease                       | 789  | 3 | 838  | 819  |
| <a href="#">Sg-contig_417</a>      | 1  | SGL_P14_E01      | Serine protease                       | 504  | 1 | 599  | 580  |
| <a href="#">Sg-contig_147</a>      | 1  | SGL_P20_F12      | Serine protease                       | 627  | 2 | 1132 | 1113 |
| <b>Hyaluronidase</b>               |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_414</a>      | 5  | SGL_P20_G05      | salivary hyaluronidase                | 1044 | 3 | 1169 | 1192 |
| <b>Apyrase</b>                     |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_354</a>      | 10 | SGL_P12_H09      | Apyrase                               | 918  | 2 | 1012 | 986  |
| <a href="#">Sg-contig_447</a>      | 2  | SGM_P4_H03       | apyrase/nucleotidase                  | 345  | 2 | 431  | 430  |
| <a href="#">Sg-contig_448</a>      | 1  | SGL_P15_A01      | apyrase/nucleotidase                  | 96   | 1 | 276  | 257  |
| <a href="#">Sg-contig_355</a>      | 1  | SGM_P5_G02       | Apyrase                               | 729  | 3 | 820  | 801  |
| <b>Amylase</b>                     |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_214</a>      | 18 | SGL_P13_D11      | Amylase                               | 1617 | 1 | 1735 | 792  |
| <a href="#">Sg-contig_213</a>      | 8  | SGL_P20_F03      | Amylase                               | 1077 | 3 | 1195 | 1131 |
| <a href="#">Sg-contig_296</a>      | 5  | SGM_P10_H01      | Amylase                               | 1257 | 1 | 1427 | 867  |
| <a href="#">Sg-contig_209</a>      | 5  | SGL_P19_G07      | Amylase                               | 561  | 1 | 677  | 694  |
| <a href="#">Sg-contig_211</a>      | 5  | SGM_P6_B08       | Amylase                               | 531  | 1 | 643  | 624  |
| <a href="#">Sg-contig_297</a>      | 4  | SGL_P17_G11      | Amylase                               | 549  | 2 | 731  | 709  |
| <a href="#">Sg-contig_298</a>      | 4  | SGL_P17_G12      | Amylase                               | 231  | 1 | 388  | 363  |
| <a href="#">Sg-contig_212</a>      | 3  | SGS_P2_F12       | Amylase                               | 207  | 3 | 320  | 327  |
| <a href="#">Sg-contig_295</a>      | 2  | SGL_P12_C04      | Amylase                               | 738  | 2 | 905  | 767  |
| <a href="#">Sg-contig_299</a>      | 2  | SGL_P19_H01      | Amylase                               | 444  | 1 | 622  | 604  |
| <a href="#">Sg-contig_210</a>      | 2  | SGL_P16_H06      | Amylase                               | 615  | 2 | 731  | 709  |
| <a href="#">Sg-contig_302</a>      | 1  | SGM_P3_E08       | Amylase                               | 381  | 2 | 534  | 515  |
| <a href="#">Sg-contig_300</a>      | 1  | SGL_P19_D05      | Amylase                               | 615  | 3 | 795  | 776  |
| <a href="#">Sg-contig_301</a>      | 1  | SGL_P13_A07      | Amylase                               | 222  | 3 | 375  | 356  |
| <b>Adenosine deaminase</b>         |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_525</a>      | 1  | SGL_P14_A04      | Adenine deaminase/adenosine deaminase | 660  | 3 | 799  | 780  |
| <b>Destabilase</b>                 |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_615</a>      | 1  | SGM_P10_D10      | salivary destabilase                  | 171  | 2 | 458  | 439  |
| <b>Insect-specific families</b>    |    |                  |                                       |      |   |      |      |
| <b>Aegyptin family</b>             |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_276</a>      | 6  | SGL_P20_F07      | Aegyptin                              | 915  | 3 | 1012 | 1020 |
| <a href="#">Sg-contig_271</a>      | 2  | SGL_P11_G06      | Aegyptin                              | 903  | 2 | 997  | 1009 |
| <a href="#">Sg-contig_272</a>      | 2  | SGL_P12_G06      | Aegyptin                              | 858  | 2 | 1013 | 994  |
| <a href="#">Sg-contig_274</a>      | 3  | SGL_P19_B10      | Aegyptin                              | 588  | 3 | 694  | 678  |
| <a href="#">Sg-contig_275</a>      | 2  | SGL_P13_B06      | Aegyptin                              | 579  | 3 | 677  | 684  |
| <a href="#">Sg-contig_277</a>      | 2  | SGM_P6_H09       | Aegyptin                              | 498  | 3 | 780  | 782  |
| <a href="#">Sg-contig_273</a>      | 2  | SGL_P20_F05      | Aegyptin                              | 612  | 1 | 718  | 693  |
| <a href="#">Sg-contig_280</a>      | 1  | SGL_P18_B10      | Aegyptin                              | 570  | 3 | 698  | 679  |
| <a href="#">Sg-contig_278</a>      | 1  | SGL_P14_D07      | Aegyptin                              | 828  | 2 | 920  | 901  |
| <a href="#">Sg-contig_281</a>      | 1  | SGL_P14_A06      | Aegyptin                              | 498  | 2 | 625  | 606  |
| <a href="#">Sg-contig_279</a>      | 1  | SGL_P19_F05      | Aegyptin                              | 909  | 3 | 1010 | 991  |
| <b>D7 family (OBP superfamily)</b> |    |                  |                                       |      |   |      |      |
| <a href="#">Sg-contig_190</a>      | 11 | SGL_P12_F02      | D7                                    | 360  | 2 | 540  | 547  |
| <a href="#">Sg-contig_75</a>       | 11 | SGM_P4_B03       | D7                                    | 345  | 2 | 616  | 595  |

|                      |    |                  |                                |     |   |     |      |
|----------------------|----|------------------|--------------------------------|-----|---|-----|------|
| <u>Sg-contig_62</u>  | 9  | SGM_P9_C02       | D7                             | 345 | 1 | 615 | 590  |
| <u>Sg-contig_66</u>  | 8  | SGM_P8_D11       | D7                             | 345 | 2 | 567 | 570  |
| <u>Sg-contig_77</u>  | 8  | SGL_P13_B10      | D7                             | 345 | 1 | 617 | 594  |
| <u>Sg-contig_187</u> | 7  | SGL_P11_H07      | D7                             | 360 | 2 | 533 | 511  |
| <u>Sg-contig_72</u>  | 7  | SGL_P18_D04      | D7                             | 345 | 1 | 615 | 622  |
| <u>Sg-contig_67</u>  | 5  | SGM_P4_G04       | D7                             | 345 | 2 | 616 | 595  |
| <u>Sg-contig_191</u> | 4  | SGL_P14_C05      | D7                             | 360 | 2 | 541 | 554  |
| <u>Sg-contig_188</u> | 4  | SGL_P14_H06      | D7                             | 360 | 1 | 537 | 518  |
| <u>Sg-contig_189</u> | 4  | SGL_P12_G05      | D7                             | 363 | 1 | 542 | 551  |
| <u>Sg-contig_78</u>  | 4  | SGL_P18_B12      | D7                             | 369 | 2 | 568 | 532  |
| <u>Sg-contig_73</u>  | 4  | SGM_P6_G02       | D7                             | 345 | 3 | 631 | 612  |
| <u>Sg-contig_363</u> | 4  | SGL_P15_C09      | salivary OBP/D7 family protein | 369 | 3 | 443 | 354  |
| <u>Sg-contig_65</u>  | 3  | SGL_P13_A03      | D7                             | 345 | 3 | 576 | 560  |
| <u>Sg-contig_68</u>  | 3  | SGL_P16_G11      | D7                             | 345 | 2 | 612 | 594  |
| <u>Sg-contig_64</u>  | 3  | SGL_P15_G06      | D7                             | 345 | 3 | 613 | 589  |
| <u>Sg-contig_71</u>  | 3  | SGM_P6_C06       | D7                             | 345 | 3 | 644 | 597  |
| <u>Sg-contig_69</u>  | 3  | SGL_P18_H02      | D7                             | 345 | 1 | 571 | 576  |
| <u>Sg-contig_63</u>  | 2  | SGL_P17_G08      | D7                             | 345 | 2 | 575 | 583  |
| <u>Sg-contig_76</u>  | 2  | SGM_P9_E03       | D7                             | 345 | 1 | 605 | 583  |
| <u>Sg-contig_70</u>  | 2  | SGM_P3_G10       | D7                             | 345 | 3 | 611 | 593  |
| <u>Sg-contig_220</u> | 18 | SGL_P20_C03      | D7                             | 840 | 2 | 944 | 970  |
| <u>Sg-contig_261</u> | 15 | SGL_P14_E07      | D7                             | 729 | 1 | 904 | 885  |
| <u>Sg-contig_221</u> | 14 | SGL_P14_G09      | D7                             | 852 | 3 | 960 | 960  |
| <u>Sg-contig_186</u> | 9  | SGM_P5_E05       | D7                             | 315 | 3 | 490 | 469  |
| <u>Sg-contig_260</u> | 8  | SGL_P14_E03      | D7                             | 690 | 3 | 837 | 806  |
| <u>Sg-contig_350</u> | 5  | SG_P1testing_D12 | D7                             | 441 | 1 | 556 | 537  |
| <u>Sg-contig_352</u> | 3  | SGL_P14_G02      | D7                             | 447 | 2 | 509 | 490  |
| <u>Sg-contig_382</u> | 3  | SGL_P17_D04      | short salivary D7 protein      | 198 | 3 | 453 | 435  |
| <u>Sg-contig_334</u> | 3  | SGL_P14_C04      | D7                             | 486 | 1 | 526 | 507  |
| <u>Sg-contig_74</u>  | 3  | SGL_P16_D11      | D7                             | 243 | 3 | 485 | 468  |
| <u>Sg-contig_262</u> | 2  | SGL_P13_B11      | D7                             | 729 | 3 | 900 | 1191 |
| <u>Sg-contig_219</u> | 2  | SGM_P8_E02       | D7                             | 852 | 2 | 950 | 932  |
| <u>Sg-contig_351</u> | 2  | SGL_P13_E11      | D7                             | 441 | 1 | 500 | 501  |
| <u>Sg-contig_383</u> | 2  | SGL_P18_D02      | short salivary D7 protein      | 387 | 1 | 628 | 632  |
| <u>Sg-contig_331</u> | 2  | SGL_P17_A04      | D7                             | 522 | 3 | 585 | 580  |
| <u>Sg-contig_335</u> | 2  | SGL_P17_E05      | D7                             | 501 | 3 | 550 | 533  |
| <u>Sg-contig_218</u> | 2  | SGL_P19_H08      | D7                             | 375 | 3 | 623 | 603  |
| <u>Sg-contig_332</u> | 2  | SGM_P10_E11      | D7                             | 525 | 2 | 566 | 547  |
| <u>Sg-contig_333</u> | 2  | SGL_P11_H04      | D7                             | 468 | 2 | 511 | 495  |
| <u>Sg-contig_362</u> | 2  | SGL_P11_E11      | salivary OBP/D7 family protein | 306 | 3 | 399 | 166  |
| <u>Sg-contig_353</u> | 1  | SGL_P17_H06      | D7                             | 441 | 2 | 523 | 504  |
| <u>Sg-contig_223</u> | 1  | SGL_P18_A01      | D7                             | 336 | 1 | 678 | 659  |
| <u>Sg-contig_81</u>  | 1  | SGS_P1_A01       | D7                             | 345 | 2 | 575 | 556  |
| <u>Sg-contig_82</u>  | 1  | SGM_P6_A02       | D7                             | 345 | 1 | 609 | 590  |
| <u>Sg-contig_194</u> | 1  | SGM_P4_D07       | D7                             | 129 | 3 | 232 | 213  |
| <u>Sg-contig_87</u>  | 1  | SGL_P15_A08      | D7                             | 288 | 1 | 448 | 429  |
| <u>Sg-contig_86</u>  | 1  | SGL_P15_G04      | D7                             | 261 | 1 | 467 | 448  |
| <u>Sg-contig_84</u>  | 1  | SGL_P11_C09      | D7                             | 390 | 3 | 630 | 611  |
| <u>Sg-contig_80</u>  | 1  | SGL_P13_F09      | D7                             | 345 | 3 | 563 | 544  |
| <u>Sg-contig_83</u>  | 1  | SGL_P14_F07      | D7                             | 345 | 2 | 565 | 546  |
| <u>Sg-contig_193</u> | 1  | SGM_P6_H04       | D7                             | 360 | 1 | 461 |      |
| <u>Sg-contig_79</u>  | 1  | SGL_P18_B09      | D7                             | 345 | 1 | 594 | 575  |
| <u>Sg-contig_367</u> | 1  | SGM_P8_H08       | salivary OBP/D7 family protein | 195 | 1 | 260 |      |
| <u>Sg-contig_337</u> | 1  | SGL_P15_E06      | D7                             | 504 | 3 | 567 | 548  |
| <u>Sg-contig_88</u>  | 1  | SGL_P12_D02      | D7                             | 231 | 3 | 601 | 582  |
| <u>Sg-contig_366</u> | 1  | SGM_P7_B08       | salivary OBP/D7 family protein | 312 | 2 | 385 | 366  |
| <u>Sg-contig_365</u> | 1  | SGL_P18_G10      | salivary OBP/D7 family protein | 354 | 1 | 427 | 408  |
| <u>Sg-contig_222</u> | 1  | SGM_P10_C07      | D7                             | 852 | 1 | 957 | 938  |
| <u>Sg-contig_385</u> | 1  | SGM_P4_C04       | short salivary D7 protein      | 384 | 3 | 647 | 628  |
| <u>Sg-contig_336</u> | 1  | SGL_P11_D06      | D7                             | 531 | 2 | 580 | 561  |
| <u>Sg-contig_364</u> | 1  | SGL_P11_C03      | salivary OBP/D7 family protein | 324 | 2 | 398 | 379  |
| <u>Sg-contig_192</u> | 1  | SGL_P12_D09      | D7                             | 246 | 3 | 430 | 411  |
| <u>Sg-contig_384</u> | 1  | SGL_P18_F06      | short salivary D7 protein      | 243 | 2 | 377 | 358  |
| <u>Sg-contig_85</u>  | 1  | SGL_P20_C05      | D7                             | 297 | 1 | 529 | 510  |

|                                   |    |                  |                                |     |   |     |      |
|-----------------------------------|----|------------------|--------------------------------|-----|---|-----|------|
| <u>Sg-contig 89</u>               | 1  | SGM_P8_H04       | D7                             | 264 | 2 | 548 | 529  |
| <u>Sg-contig 425</u>              | 1  | SGM_P7_C03       | salivary OBP/D7 family protein | 285 | 3 | 362 | 343  |
| <u>Sg-contig 426</u>              | 1  | SGM_P10_C11      | salivary OBP/D7 family protein | 363 | 2 | 438 | 419  |
| <u>Sg-contig 424</u>              | 1  | SGL_P13_A08      | salivary OBP/D7 family protein | 174 | 1 | 249 | 230  |
| <u>Sg-contig 427</u>              | 1  | SGM_P3_H08       | salivary OBP/D7 family protein | 111 | 3 | 184 | 165  |
| <u>Sg-contig 476</u>              | 1  | SGL_P11_C10      | salivary OBP/D7 family protein | 522 | 2 | 526 | 507  |
| <b>Simulium-specific families</b> |    |                  |                                |     |   |     |      |
| <b>SVEP vasodilator family</b>    |    |                  |                                |     |   |     |      |
| <u>Sg-contig 8</u>                | 20 | SGL_P15_D08      | SVEP family member             | 495 | 1 | 589 | 597  |
| <u>Sg-contig 13</u>               | 16 | SGM_P9_G02       | SVEP family member             | 492 | 1 | 772 | 1069 |
| <u>Sg-contig 1</u>                | 16 | SGM_P9_H05       | SVEP family member             | 495 | 2 | 596 | 594  |
| <u>Sg-contig 5</u>                | 15 | SGM_P9_D05       | SVEP family member             | 492 | 3 | 591 | 590  |
| <u>Sg-contig 11</u>               | 11 | SGL_P18_E09      | SVEP family member             | 489 | 2 | 754 | 731  |
| <u>Sg-contig 7</u>                | 7  | SGL_P15_F11      | SVEP family member             | 489 | 2 | 594 | 578  |
| <u>Sg-contig 12</u>               | 5  | SGM_P8_F09       | SVEP family member             | 492 | 3 | 592 | 598  |
| <u>Sg-contig 14</u>               | 5  | SGL_P13_G01      | SVEP family member             | 489 | 1 | 684 | 666  |
| <u>Sg-contig 9</u>                | 4  | SGL_P11_E12      | SVEP family member             | 492 | 3 | 585 | 594  |
| <u>Sg-contig 2</u>                | 3  | SGL_P16_D12      | SVEP family member             | 495 | 1 | 631 | 635  |
| <u>Sg-contig 6</u>                | 3  | SGL_P18_E02      | SVEP family member             | 489 | 2 | 575 | 556  |
| <u>Sg-contig 4</u>                | 2  | SGL_P20_B01      | SVEP family member             | 495 | 1 | 590 | 647  |
| <u>Sg-contig 16</u>               | 2  | SGL_P18_D06      | SVEP family member             | 471 | 2 | 580 | 563  |
| <u>Sg-contig 15</u>               | 2  | SGL_P17_D12      | SVEP family member             | 489 | 2 | 616 | 559  |
| <u>Sg-contig 10</u>               | 2  | SGL_P12_F12      | SVEP family member             | 489 | 2 | 584 | 565  |
| <u>Sg-contig 3</u>                | 2  | SGL_P19_D06      | SVEP family member             | 495 | 1 | 605 | 590  |
| <u>Sg-contig 102</u>              | 7  | SGL_P14_F05      | SVEP family member             | 465 | 2 | 602 | 582  |
| <u>Sg-contig 99</u>               | 5  | SGL_P19_C06      | SVEP family member             | 465 | 3 | 611 | 584  |
| <u>Sg-contig 95</u>               | 5  | SGL_P20_A05      | SVEP family member             | 468 | 1 | 567 | 570  |
| <u>Sg-contig 93</u>               | 4  | SG_P1testing_F02 | SVEP family member             | 438 | 1 | 573 | 557  |
| <u>Sg-contig 90</u>               | 4  | SG_P1testing_E07 | SVEP family member             | 300 | 1 | 365 | 347  |
| <u>Sg-contig 92</u>               | 3  | SGL_P11_E02      | SVEP family member             | 456 | 1 | 526 | 541  |
| <u>Sg-contig 100</u>              | 3  | SGM_P7_H02       | SVEP family member             | 438 | 3 | 575 | 541  |
| <u>Sg-contig 97</u>               | 2  | SG_P1testing_G06 | SVEP family member             | 477 | 3 | 614 | 595  |
| <u>Sg-contig 103</u>              | 2  | SGL_P14_F04      | SVEP family member             | 465 | 2 | 602 | 584  |
| <u>Sg-contig 94</u>               | 2  | SGL_P18_B01      | SVEP family member             | 453 | 2 | 590 | 579  |
| <u>Sg-contig 98</u>               | 2  | SGL_P17_C03      | SVEP family member             | 462 | 1 | 604 | 628  |
| <u>Sg-contig 101</u>              | 2  | SGL_P13_D05      | SVEP family member             | 336 | 1 | 608 | 589  |
| <u>Sg-contig 91</u>               | 2  | SGM_P3_F12       | SVEP family member             | 309 | 3 | 453 | 434  |
| <u>Sg-contig 96</u>               | 2  | SGL_P17_C04      | SVEP family member             | 297 | 1 | 528 | 528  |
| <u>Sg-contig 17</u>               | 2  | SGL_P12_C08      | SVEP family member             | 369 | 2 | 464 | 444  |
| <u>Sg-contig 18</u>               | 2  | SG_P1testing_A10 | SVEP family member             | 348 | 2 | 435 | 413  |
| <u>Sg-contig 20</u>               | 2  | SGM_P9_A09       | SVEP family member             | 249 | 2 | 345 | 351  |
| <u>Sg-contig 19</u>               | 2  | SG_P1testing_G03 | SVEP family member             | 354 | 1 | 437 | 174  |
| <u>Sg-contig 115</u>              | 1  | SGL_P16_F10      | SVEP family member             | 486 | 2 | 581 | 562  |
| <u>Sg-contig 24</u>               | 1  | SGL_P14_D04      | SVEP family member             | 516 | 2 | 609 | 590  |
| <u>Sg-contig 25</u>               | 1  | SGM_P4_D12       | SVEP family member             | 480 | 3 | 575 | 556  |
| <u>Sg-contig 26</u>               | 1  | SGM_P5_A11       | SVEP family member             | 486 | 2 | 578 | 559  |
| <u>Sg-contig 21</u>               | 1  | SGL_P11_A12      | SVEP family member             | 504 | 2 | 608 | 589  |
| <u>Sg-contig 22</u>               | 1  | SGL_P19_E12      | SVEP family member             | 489 | 1 | 599 | 580  |
| <u>Sg-contig 106</u>              | 1  | SGM_P5_H03       | SVEP family member             | 477 | 3 | 532 | 513  |
| <u>Sg-contig 107</u>              | 1  | SGM_P7_E06       | SVEP family member             | 480 | 3 | 555 | 536  |
| <u>Sg-contig 112</u>              | 1  | SGL_P18_G05      | SVEP family member             | 468 | 2 | 604 | 585  |
| <u>Sg-contig 23</u>               | 1  | SGL_P13_B12      | SVEP family member             | 330 | 1 | 583 | 564  |
| <u>Sg-contig 113</u>              | 1  | SGL_P12_F03      | SVEP family member             | 468 | 2 | 561 | 542  |
| <u>Sg-contig 109</u>              | 1  | SGM_P10_F06      | SVEP family member             | 324 | 3 | 461 | 442  |
| <u>Sg-contig 110</u>              | 1  | SGL_P16_A02      | SVEP family member             | 468 | 2 | 605 | 586  |
| <u>Sg-contig 105</u>              | 1  | SGL_P18_E07      | SVEP family member             | 300 | 3 | 426 | 407  |
| <u>Sg-contig 114</u>              | 1  | SGL_P16_G05      | SVEP family member             | 330 | 2 | 467 | 448  |
| <u>Sg-contig 111</u>              | 1  | SGM_P4_G03       | SVEP family member             | 465 | 1 | 531 | 512  |
| <u>Sg-contig 104</u>              | 1  | SGL_P14_B09      | SVEP family member             | 465 | 1 | 606 | 587  |
| <u>Sg-contig 108</u>              | 1  | SGM_P6_D12       | SVEP family member             | 465 | 2 | 535 | 516  |
| <u>Sg-contig 28</u>               | 1  | SGL_P20_A01      | SVEP family member             | 228 | 1 | 459 | 440  |
| <u>Sg-contig 27</u>               | 1  | SGL_P17_F03      | SVEP family member             | 240 | 1 | 628 | 609  |

|  |    |                  |   |      |   |      |      |
|--|----|------------------|---|------|---|------|------|
| Sg-contig_652                                  | 1  | SGM_P5_F02       | erythema protein SVEP-3                 | 204  | 2 | 348  | 329  |
| Sg-contig_609                                  | 1  | SGM_P10_B05      | erythema protein SVEP-5                 | 156  | 2 | 252  | 233  |
| <b>Other <i>Simulium</i> specific families</b> |    |                  |   |      |   |      |      |
| Sg-contig_195                                  | 23 | SGL_P12_A10      | hypothetical secreted protein           | 231  | 2 | 308  | 310  |
| Sg-contig_199                                  | 5  | SGM_P10_G08      | hypothetical secreted protein           | 231  | 1 | 305  | 287  |
| Sg-contig_196                                  | 5  | SGL_P12_G09      | hypothetical secreted protein           | 231  | 3 | 304  | 288  |
| Sg-contig_198                                  | 2  | SGL_P11_A01      | hypothetical secreted protein           | 219  | 3 | 308  | 289  |
| Sg-contig_216                                  | 21 | SGL_P12_D11      | hypothetical secreted protein           | 1125 | 1 | 1237 | 1129 |
| Sg-contig_215                                  | 16 | SGL_P17_B11      | hypothetical secreted protein           | 1116 | 1 | 1202 | 1201 |
| Sg-contig_136                                  | 9  | SGL_P14_H10      | hypothetical secreted protein           | 813  | 2 | 1124 | 869  |
| Sg-contig_344                                  | 6  | SGM_P7_D06       | hypothetical secreted peptide precursor | 468  | 3 | 544  | 519  |
| Sg-contig_303                                  | 6  | SGL_P16_D08      | hypothetical secreted peptide precursor | 513  | 2 | 600  | 607  |
| Sg-contig_343                                  | 5  | SGM_P8_E10       | hypothetical secreted peptide precursor | 465  | 3 | 536  | 543  |
| Sg-contig_305                                  | 5  | SGL_P19_D09      | hypothetical secreted peptide precursor | 513  | 1 | 600  | 585  |
| Sg-contig_227                                  | 5  | SGL_P16_F12      | hypothetical secreted peptide precursor | 285  | 3 | 425  | 421  |
| Sg-contig_134                                  | 5  | SGL_P13_C02      | hypothetical secreted protein           | 810  | 1 | 883  | 851  |
| Sg-contig_283                                  | 4  | SGL_P14_C12      | hypothetical secreted peptide precursor | 441  | 2 | 499  | 480  |
| Sg-contig_230                                  | 4  | SGL_P16_G07      | hypothetical secreted peptide precursor | 285  | 1 | 446  | 422  |
| Sg-contig_133                                  | 4  | SG_P1testing_C01 | hypothetical secreted protein           | 810  | 2 | 870  | 848  |
| Sg-contig_339                                  | 3  | SGM_P5_E11       | hypothetical secreted peptide precursor | 510  | 2 | 567  | 548  |
| Sg-contig_340                                  | 3  | SGL_P14_C07      | hypothetical secreted peptide precursor | 507  | 3 | 568  | 549  |
| Sg-contig_284                                  | 3  | SGL_P16_C01      | hypothetical secreted peptide precursor | 432  | 3 | 510  | 502  |
| Sg-contig_304                                  | 3  | SGL_P16_D03      | hypothetical secreted peptide precursor | 510  | 1 | 597  | 585  |
| Sg-contig_224                                  | 3  | SG_P1testing_D04 | hypothetical secreted peptide precursor | 285  | 1 | 412  | 398  |
| Sg-contig_225                                  | 3  | SGM_P10_E06      | hypothetical secreted peptide precursor | 285  | 2 | 418  | 399  |
| Sg-contig_228                                  | 3  | SGM_P5_F07       | hypothetical secreted peptide precursor | 285  | 1 | 421  | 425  |
| Sg-contig_135                                  | 3  | SGL_P15_H03      | hypothetical secreted protein           | 810  | 2 | 880  | 858  |
| Sg-contig_306                                  | 2  | SGM_P8_F01       | hypothetical secreted peptide precursor | 513  | 3 | 604  | 589  |
| Sg-contig_232                                  | 2  | SGL_P20_G08      | hypothetical secreted peptide precursor | 285  | 1 | 390  | 371  |
| Sg-contig_231                                  | 2  | SGM_P5_A04       | hypothetical secreted peptide precursor | 285  | 1 | 420  | 401  |
| Sg-contig_226                                  | 2  | SGL_P15_C06      | hypothetical secreted peptide precursor | 285  | 1 | 417  | 394  |
| Sg-contig_229                                  | 2  | SGL_P14_B08      | hypothetical secreted peptide precursor | 285  | 3 | 419  | 422  |
| Sg-contig_205                                  | 37 | SGL_P16_E04      | hypothetical secreted protein           | 402  | 1 | 491  | 479  |
| Sg-contig_258                                  | 22 | SGL_P20_H09      | hypothetical secreted peptide precursor | 333  | 3 | 482  | 491  |
| Sg-contig_282                                  | 9  | SG_P1testing_H06 | hypothetical secreted peptide precursor | 414  | 1 | 480  | 458  |
| Sg-contig_324                                  | 8  | SGL_P13_C09      | hypothetical secreted peptide precursor | 543  | 2 | 618  | 599  |
| Sg-contig_177                                  | 7  | SG_P1testing_D09 | hypothetical secreted protein           | 369  | 3 | 868  | 587  |
| Sg-contig_257                                  | 7  | SGL_P12_E08      | hypothetical secreted peptide precursor | 330  | 1 | 564  | 538  |
| Sg-contig_309                                  | 5  | SGM_P10_G11      | hypothetical secreted peptide precursor | 513  | 3 | 598  | 579  |
| Sg-contig_410                                  | 5  | SGL_P12_G03      | hypothetical secreted protein           | 189  | 3 | 295  | 168  |
| Sg-contig_319                                  | 5  | SGL_P20_G01      | hypothetical secreted peptide precursor | 498  | 3 | 740  | 731  |
| Sg-contig_338                                  | 5  | SGL_P16_B06      | hypothetical secreted peptide precursor | 354  | 1 | 434  | 429  |
| Sg-contig_321                                  | 4  | SGL_P11_D09      | hypothetical secreted peptide precursor | 690  | 2 | 1110 | 1075 |
| Sg-contig_311                                  | 4  | SGL_P20_A04      | hypothetical secreted peptide precursor | 516  | 2 | 610  | 574  |
| Sg-contig_165                                  | 4  | SGM_P10_H09      | hypothetical secreted protein           | 429  | 2 | 607  | 578  |
| Sg-contig_170                                  | 4  | SGM_P3_E06       | hypothetical secreted protein           | 399  | 2 | 582  | 508  |
| Sg-contig_422                                  | 4  | SGM_P4_E01       | hypothetical secreted protein           | 294  | 3 | 373  | 353  |
| Sg-contig_420                                  | 4  | SGL_P16_F07      | hypothetical secreted peptide precursor | 312  | 3 | 499  | 480  |
| Sg-contig_174                                  | 4  | SGL_P20_F08      | hypothetical secreted protein           | 312  | 3 | 446  | 428  |
| Sg-contig_356                                  | 3  | SGL_P11_D02      | hypothetical secreted peptide precursor | 426  | 3 | 560  | 543  |
| Sg-contig_171                                  | 3  | SGM_P8_H09       | hypothetical secreted protein           | 369  | 3 | 527  | 503  |
| Sg-contig_320                                  | 3  | SGL_P13_E12      | hypothetical secreted peptide precursor | 834  | 2 | 1060 | 917  |
| Sg-contig_357                                  | 3  | SGL_P14_B10      | hypothetical secreted peptide precursor | 426  | 2 | 555  | 547  |
| Sg-contig_176                                  | 3  | SGL_P17_E09      | hypothetical secreted protein           | 339  | 1 | 549  | 529  |
| Sg-contig_446                                  | 3  | SGL_P15_B11      | hypothetical secreted protein           | 423  | 2 | 567  | 553  |
| Sg-contig_168                                  | 2  | SGL_P13_F06      | hypothetical secreted protein           | 411  | 3 | 599  | 568  |
| Sg-contig_169                                  | 2  | SGL_P11_B07      | hypothetical secreted protein           | 393  | 3 | 567  | 555  |
| Sg-contig_173                                  | 2  | SGM_P4_F01       | hypothetical secreted protein           | 381  | 3 | 573  | 568  |
| Sg-contig_166                                  | 2  | SGM_P9_F05       | hypothetical secreted protein           | 411  | 2 | 561  | 541  |
| Sg-contig_167                                  | 2  | SGL_P18_B05      | hypothetical secreted protein           | 417  | 2 | 601  | 558  |
| Sg-contig_172                                  | 2  | SGL_P11_F11      | hypothetical secreted protein           | 393  | 2 | 526  | 505  |
| Sg-contig_310                                  | 2  | SGL_P14_B05      | hypothetical secreted peptide precursor | 516  | 1 | 606  | 616  |
| Sg-contig_175                                  | 2  | SGL_P17_F02      | hypothetical secreted protein           | 393  | 1 | 542  | 510  |
| Sg-contig_371                                  | 2  | SG_P1testing_E05 | hypothetical secreted peptide precursor | 402  | 3 | 451  | 440  |



|                      |   |                  |   |     |   |      |      |
|----------------------|---|------------------|---|-----|---|------|------|
| <u>Sg-contig 197</u> | 2 | SGL_P11_B11      | hypothetical secreted protein           | 213 | 3 | 265  | 246  |
| <u>Sg-contig 438</u> | 2 | SGL_P12_B04      | hypothetical secreted peptide precursor | 216 | 2 | 300  |      |
| <u>Sg-contig 370</u> | 2 | SG_P1testing_A03 | hypothetical secreted peptide precursor | 234 | 2 | 271  | 254  |
| <u>Sg-contig 412</u> | 2 | SGL_P19_A02      | hypothetical secreted protein           | 564 | 1 | 781  | 675  |
| <u>Sg-contig 411</u> | 2 | SGM_P3_F06       | hypothetical secreted protein           | 534 | 1 | 1294 | 869  |
| <u>Sg-contig 403</u> | 2 | SGL_P12_C12      | hypothetical secreted protein           | 330 | 1 | 418  | 403  |
| <u>Sg-contig 206</u> | 2 | SGL_P16_B09      | hypothetical secreted protein           | 318 | 2 | 460  | 455  |
| <u>Sg-contig 318</u> | 2 | SGL_P11_B05      | hypothetical secreted peptide precursor | 345 | 2 | 554  | 535  |
| <u>Sg-contig 137</u> | 2 | SGL_P17_C10      | hypothetical secreted protein           | 459 | 3 | 879  | 888  |
| <u>Sg-contig 413</u> | 1 | SGL_P14_B07      | hypothetical secreted protein           | 171 | 3 | 340  | 321  |
| <u>Sg-contig 361</u> | 1 | SGM_P4_H04       | hypothetical secreted protein           | 141 | 2 | 274  | 255  |
| <u>Sg-contig 323</u> | 1 | SGM_P3_D01       | hypothetical secreted peptide precursor | 162 | 2 | 367  | 348  |
| <u>Sg-contig 423</u> | 1 | SGM_P4_C12       | hypothetical secreted protein           | 291 | 1 | 362  |      |
| <u>Sg-contig 421</u> | 1 | SGM_P5_H01       | hypothetical secreted peptide precursor | 183 | 1 | 283  | 264  |
| <u>Sg-contig 237</u> | 1 | SGM_P7_G08       | hypothetical secreted peptide precursor | 285 | 2 | 415  | 396  |
| <u>Sg-contig 345</u> | 1 | SGL_P16_G06      | hypothetical secreted peptide precursor | 468 | 1 | 543  | 524  |
| <u>Sg-contig 307</u> | 1 | SG_P1testing_E04 | hypothetical secreted peptide precursor | 411 | 3 | 492  | 473  |
| <u>Sg-contig 407</u> | 1 | SGM_P3_A02       | hypothetical secreted protein           | 420 | 3 | 491  | 472  |
| <u>Sg-contig 405</u> | 1 | SGL_P14_H01      | hypothetical secreted protein           | 366 | 3 | 441  | 422  |
| <u>Sg-contig 317</u> | 1 | SGM_P3_D12       | hypothetical secreted peptide precursor | 315 | 1 | 391  |      |
| <u>Sg-contig 285</u> | 1 | SGL_P13_E07      | hypothetical secreted peptide precursor | 435 | 2 | 494  | 475  |
| <u>Sg-contig 183</u> | 1 | SGM_P7_C09       | hypothetical secreted protein           | 369 | 3 | 546  | 527  |
| <u>Sg-contig 200</u> | 1 | SGM_P10_A11      | hypothetical secreted protein           | 246 | 2 | 298  | 279  |
| <u>Sg-contig 203</u> | 1 | SGM_P4_D08       | hypothetical secreted protein           | 231 | 2 | 278  | 259  |
| <u>Sg-contig 322</u> | 1 | SGL_P20_G04      | hypothetical secreted peptide precursor | 834 | 2 | 1044 | 1025 |
| <u>Sg-contig 314</u> | 1 | SGL_P19_F02      | hypothetical secreted peptide precursor | 519 | 2 | 600  | 581  |
| <u>Sg-contig 313</u> | 1 | SGL_P13_B09      | hypothetical secreted peptide precursor | 513 | 2 | 593  | 574  |
| <u>Sg-contig 181</u> | 1 | SGL_P17_A03      | hypothetical secreted protein           | 393 | 2 | 524  | 505  |
| <u>Sg-contig 204</u> | 1 | SGM_P8_A03       | hypothetical secreted protein           | 231 | 1 | 306  | 287  |
| <u>Sg-contig 315</u> | 1 | SGL_P17_B10      | hypothetical secreted peptide precursor | 519 | 1 | 587  | 568  |
| <u>Sg-contig 326</u> | 1 | SGL_P11_E10      | hypothetical secreted peptide precursor | 504 | 1 | 578  | 559  |
| <u>Sg-contig 325</u> | 1 | SGL_P18_H10      | hypothetical secreted peptide precursor | 411 | 1 | 582  | 563  |
| <u>Sg-contig 179</u> | 1 | SGL_P13_F02      | hypothetical secreted protein           | 423 | 1 | 598  | 579  |
| <u>Sg-contig 207</u> | 1 | SGL_P17_D03      | hypothetical secreted protein           | 423 | 1 | 484  | 465  |
| <u>Sg-contig 328</u> | 1 | SGM_P7_B05       | hypothetical secreted peptide precursor | 180 | 3 | 247  |      |
| <u>Sg-contig 360</u> | 1 | SGL_P15_E04      | hypothetical secreted protein           | 288 | 2 | 372  | 353  |
| <u>Sg-contig 180</u> | 1 | SGM_P5_D04       | hypothetical secreted protein           | 315 | 3 | 397  | 378  |
| <u>Sg-contig 178</u> | 1 | SGL_P11_A10      | hypothetical secreted protein           | 393 | 2 | 594  | 575  |
| <u>Sg-contig 286</u> | 1 | SGS_P2_E09       | hypothetical secreted peptide precursor | 441 | 3 | 510  | 491  |
| <u>Sg-contig 287</u> | 1 | SGL_P19_B03      | hypothetical secreted peptide precursor | 294 | 3 | 499  | 480  |
| <u>Sg-contig 235</u> | 1 | SGL_P18_H09      | hypothetical secreted peptide precursor | 285 | 3 | 417  | 398  |
| <u>Sg-contig 238</u> | 1 | SGM_P5_C06       | hypothetical secreted peptide precursor | 285 | 2 | 427  | 408  |
| <u>Sg-contig 239</u> | 1 | SGL_P14_C06      | hypothetical secreted peptide precursor | 285 | 1 | 419  | 400  |
| <u>Sg-contig 233</u> | 1 | SGL_P17_E08      | hypothetical secreted peptide precursor | 303 | 1 | 413  | 394  |
| <u>Sg-contig 234</u> | 1 | SGS_P1_D04       | hypothetical secreted peptide precursor | 285 | 1 | 410  | 391  |
| <u>Sg-contig 236</u> | 1 | SGL_P11_A08      | hypothetical secreted peptide precursor | 285 | 1 | 416  | 397  |
| <u>Sg-contig 289</u> | 1 | SGM_P3_F07       | hypothetical secreted peptide precursor | 255 | 1 | 317  | 298  |
| <u>Sg-contig 406</u> | 1 | SGL_P12_E02      | hypothetical secreted protein           | 447 | 2 | 527  | 508  |
| <u>Sg-contig 316</u> | 1 | SGM_P3_F04       | hypothetical secreted peptide precursor | 474 | 2 | 548  | 529  |
| <u>Sg-contig 240</u> | 1 | SGL_P19_B04      | hypothetical secreted peptide precursor | 285 | 3 | 422  | 403  |
| <u>Sg-contig 342</u> | 1 | SGL_P15_H04      | hypothetical secreted peptide precursor | 507 | 2 | 567  | 548  |
| <u>Sg-contig 308</u> | 1 | SGL_P11_G08      | hypothetical secreted peptide precursor | 558 | 2 | 645  | 626  |
| <u>Sg-contig 143</u> | 1 | SG_P1testing_G02 | hypothetical secreted protein           | 810 | 1 | 873  | 854  |
| <u>Sg-contig 341</u> | 1 | SGL_P16_D09      | hypothetical secreted peptide precursor | 417 | 2 | 478  | 459  |
| <u>Sg-contig 144</u> | 1 | SGM_P8_C07       | hypothetical secreted protein           | 405 | 2 | 470  | 451  |
| <u>Sg-contig 201</u> | 1 | SGS_P2_C02       | hypothetical secreted protein           | 219 | 2 | 300  | 281  |
| <u>Sg-contig 327</u> | 1 | SGM_P4_F09       | hypothetical secreted peptide precursor | 507 | 2 | 581  | 562  |
| <u>Sg-contig 329</u> | 1 | SGM_P8_G06       | hypothetical secreted peptide precursor | 174 | 3 | 258  |      |
| <u>Sg-contig 373</u> | 1 | SGM_P8_E07       | hypothetical secreted peptide precursor | 285 | 1 | 318  | 299  |
| <u>Sg-contig 346</u> | 1 | SGM_P10_F02      | hypothetical secreted peptide precursor | 276 | 2 | 351  | 332  |
| <u>Sg-contig 202</u> | 1 | SGM_P3_H10       | hypothetical secreted protein           | 219 | 2 | 301  | 282  |
| <u>Sg-contig 312</u> | 1 | SGL_P16_H07      | hypothetical secreted peptide precursor | 447 | 2 | 582  | 563  |
| <u>Sg-contig 374</u> | 1 | SGL_P15_D09      | hypothetical secreted peptide precursor | 279 | 2 | 313  | 294  |
| <u>Sg-contig 404</u> | 1 | SGL_P11_F06      | hypothetical secreted protein           | 213 | 2 | 284  | 265  |

|                           |    |                  |   |      |   |      |      |
|---------------------------|----|------------------|---|------|---|------|------|
| <u>Sg-contig_288</u>      | 1  | SGM_P4_E12       | hypothetical secreted peptide precursor | 291  | 2 | 356  | 337  |
| <u>Sg-contig_182</u>      | 1  | SG_P1testing_E08 | hypothetical secreted protein           | 351  | 1 | 534  | 515  |
| <u>Sg-contig_439</u>      | 1  | SGL_P11_F07      | hypothetical secreted peptide precursor | 213  | 1 | 281  | 262  |
| <u>Sg-contig_290</u>      | 1  | SGM_P6_C02       | hypothetical secreted peptide precursor | 258  | 1 | 321  | 302  |
| <u>Sg-contig_217</u>      | 1  | SGM_P4_H09       | hypothetical secreted protein           | 822  | 3 | 1223 | 1204 |
| <u>Sg-contig_359</u>      | 1  | SGL_P17_C08      | hypothetical secreted protein           | 294  | 3 | 518  | 499  |
| <u>Sg-contig_208</u>      | 1  | SGL_P12_D03      | hypothetical secreted protein           | 243  | 2 | 438  | 419  |
| <u>Sg-contig_330</u>      | 1  | SGL_P16_A09      | hypothetical secreted peptide precursor | 339  | 3 | 628  | 609  |
| <u>Sg-contig_185</u>      | 1  | SGL_P17_D01      | hypothetical secreted protein           | 360  | 1 | 568  | 549  |
| <u>Sg-contig_467</u>      | 1  | SG_P1testing_D10 | hypothetical secreted protein           | 165  | 1 | 219  |      |
| <u>Sg-contig_593</u>      | 1  | SGL_P19_H05      | hypothetical secreted protein           | 552  | 2 | 643  | 624  |
| <u>Sg-contig_526</u>      | 1  | SGL_P14_A09      | hypothetical secreted protein           | 273  | 2 | 335  | 316  |
| <u>Sg-contig_435</u>      | 1  | SG_P1testing_G09 | hypothetical secreted peptide precursor | 294  | 3 | 350  |      |
| <u>Sg-contig_437</u>      | 1  | SGL_P18_A08      | hypothetical secreted peptide precursor | 279  | 1 | 427  | 408  |
| <u>Sg-contig_436</u>      | 1  | SGM_P7_H03       | hypothetical secreted peptide precursor | 261  | 3 | 414  | 395  |
| <u>Sg-contig_594</u>      | 1  | SGL_P19_H10      | hypothetical secreted protein           | 141  | 2 | 191  | 172  |
| <u>Sg-contig_751</u>      | 1  | SGS_P2_G11       | hypothetical secreted protein           | 285  | 3 | 521  | 502  |
| <u>Sg-contig_487</u>      | 1  | SGL_P12_A07      | hypothetical secreted peptide precursor | 366  | 3 | 473  | 454  |
| <u>Sg-contig_596</u>      | 1  | SGL_P20_B06      | hypothetical secreted peptide precursor | 357  | 1 | 500  | 481  |
| <u>Sg-contig_602</u>      | 1  | SGL_P20_F01      | hypothetical secreted protein           | 603  | 3 | 692  | 673  |
| <u>Sg-contig_653</u>      | 1  | SGM_P5_F04       | hypothetical secreted protein           | 162  | 2 | 283  |      |
| <b>PolyQ mucin family</b> |    |                  |   |      |   |      |      |
| <u>Sg-contig_126</u>      | 14 | SGL_P13_E08      | PolyQ family member                     | 1089 | 2 | 1288 | 535  |
| <u>Sg-contig_129</u>      | 3  | SGL_P16_D06      | PolyQ family member                     | 975  | 1 | 1133 | 1192 |
| <u>Sg-contig_121</u>      | 2  | SGL_P18_E08      | PolyQ family member                     | 1113 | 1 | 1302 | 828  |
| <u>Sg-contig_125</u>      | 2  | SGL_P16_H01      | PolyQ family member                     | 1071 | 1 | 1252 | 661  |
| <u>Sg-contig_128</u>      | 2  | SGM_P9_E10       | PolyQ family member                     | 1035 | 3 | 1218 | 663  |
| <u>Sg-contig_120</u>      | 6  | SGL_P11_F03      | PolyQ family member                     | 978  | 2 | 1162 | 1155 |
| <u>Sg-contig_117</u>      | 4  | SGL_P12_E06      | PolyQ family member                     | 507  | 2 | 689  | 676  |
| <u>Sg-contig_119</u>      | 4  | SGL_P13_F01      | PolyQ family member                     | 387  | 1 | 568  | 566  |
| <u>Sg-contig_127</u>      | 3  | SGL_P18_B02      | PolyQ family member                     | 963  | 2 | 1145 | 1134 |
| <u>Sg-contig_122</u>      | 2  | SGL_P20_D04      | PolyQ family member                     | 528  | 2 | 709  | 715  |
| <u>Sg-contig_118</u>      | 2  | SGM_P5_F09       | PolyQ family member                     | 618  | 2 | 802  | 784  |
| <u>Sg-contig_123</u>      | 2  | SG_P1testing_E03 | PolyQ family member                     | 753  | 2 | 936  | 914  |
| <u>Sg-contig_124</u>      | 2  | SGL_P20_H05      | PolyQ family member                     | 612  | 2 | 794  | 991  |
| <u>Sg-contig_132</u>      | 1  | SGL_P11_B03      | PolyQ family member                     | 114  | 2 | 265  | 246  |
| <u>Sg-contig_130</u>      | 1  | SGL_P20_B04      | PolyQ family member                     | 363  | 2 | 544  | 525  |
| <u>Sg-contig_131</u>      | 1  | SGL_P12_G07      | PolyQ family member                     | 969  | 2 | 1241 |      |
| <b>GH repeat family</b>   |    |                  |   |      |   |      |      |
| <u>Sg-contig_41</u>       | 17 | SGS_P1_C05       | GH repeat family member                 | 456  | 2 | 615  | 612  |
| <u>Sg-contig_33</u>       | 13 | SGL_P16_H05      | GH repeat family member                 | 540  | 3 | 647  | 620  |
| <u>Sg-contig_38</u>       | 7  | SGL_P15_E03      | GH repeat family member                 | 510  | 3 | 614  | 617  |
| <u>Sg-contig_39</u>       | 5  | SGM_P5_F01       | GH repeat family member                 | 459  | 3 | 619  | 590  |
| <u>Sg-contig_40</u>       | 4  | SGL_P19_F06      | GH repeat family member                 | 480  | 3 | 576  | 564  |
| <u>Sg-contig_42</u>       | 2  | SGM_P4_H10       | GH repeat family member                 | 441  | 3 | 592  | 572  |
| <u>Sg-contig_30</u>       | 15 | SGL_P20_B09      | GH repeat family member                 | 522  | 2 | 618  | 598  |
| <u>Sg-contig_37</u>       | 15 | SGL_P18_C10      | GH repeat family member                 | 471  | 3 | 621  | 606  |
| <u>Sg-contig_31</u>       | 12 | SGL_P17_G05      | GH repeat family member                 | 492  | 1 | 618  | 619  |
| <u>Sg-contig_29</u>       | 5  | SGL_P20_A09      | GH repeat family member                 | 519  | 1 | 612  | 593  |
| <u>Sg-contig_36</u>       | 4  | SGM_P8_H07       | GH repeat family member                 | 564  | 3 | 658  | 662  |
| <u>Sg-contig_32</u>       | 4  | SGL_P20_C07      | GH repeat family member                 | 546  | 1 | 645  | 627  |
| <u>Sg-contig_35</u>       | 3  | SGL_P11_C05      | GH repeat family member                 | 495  | 1 | 687  | 690  |
| <u>Sg-contig_34</u>       | 2  | SGL_P16_B01      | GH repeat family member                 | 495  | 1 | 618  | 619  |
| <u>Sg-contig_46</u>       | 1  | SGL_P19_C11      | GH repeat family member                 | 540  | 1 | 638  | 619  |
| <u>Sg-contig_59</u>       | 1  | SGM_P3_G04       | GH repeat family member                 | 240  | 1 | 304  | 285  |
| <u>Sg-contig_53</u>       | 1  | SGM_P3_G07       | GH repeat family member                 | 546  | 3 | 645  | 626  |
| <u>Sg-contig_48</u>       | 1  | SGL_P19_B05      | GH repeat family member                 | 510  | 3 | 631  | 612  |
| <u>Sg-contig_47</u>       | 1  | SGL_P13_D04      | GH repeat family member                 | 510  | 2 | 634  | 615  |
| <u>Sg-contig_44</u>       | 1  | SGL_P17_H03      | GH repeat family member                 | 552  | 1 | 653  | 634  |
| <u>Sg-contig_55</u>       | 1  | SGL_P13_A09      | GH repeat family member                 | 492  | 1 | 624  | 605  |
| <u>Sg-contig_56</u>       | 1  | SGM_P8_G09       | GH repeat family member                 | 492  | 1 | 620  | 601  |
| <u>Sg-contig_57</u>       | 1  | SGM_P10_C04      | GH repeat family member                 | 492  | 1 | 619  | 600  |
| <u>Sg-contig_54</u>       | 1  | SGS_P2_D09       | GH repeat family member                 | 468  | 2 | 564  | 545  |

|                             |    |                  |  |     |   |     |     |
|-----------------------------|----|------------------|--|-----|---|-----|-----|
| <u>Sg-contig_49</u>         | 1  | SGM_P8_E11       | GH repeat family nenber  | 501 | 2 | 601 | 582 |
| <u>Sg-contig_50</u>         | 1  | SGM_P9_B03       | GH repeat family nenber  | 471 | 1 | 526 | 507 |
| <u>Sg-contig_51</u>         | 1  | SGL_P12_H03      | GH repeat family nenber  | 489 | 2 | 619 | 600 |
| <u>Sg-contig_45</u>         | 1  | SGL_P15_C02      | GH repeat family nenber  | 525 | 3 | 676 | 657 |
| <u>Sg-contig_58</u>         | 1  | SGM_P6_B04       | GH repeat family nenber  | 288 | 3 | 369 | 350 |
| <u>Sg-contig_60</u>         | 1  | SGM_P4_H08       | GH repeat family nenber  | 258 | 2 | 323 | 304 |
| <u>Sg-contig_43</u>         | 1  | SGL_P18_C12      | GH repeat family nenber  | 525 | 2 | 615 | 596 |
| <u>Sg-contig_52</u>         | 1  | SGM_P5_C03       | GH repeat family nenber  | 390 | 3 | 439 | 420 |
| <u>Sg-contig_61</u>         | 1  | SGL_P13_H07      | GH repeat family nenber  | 279 | 1 | 703 | 684 |
| <b>Collagen-like family</b> |    |                  |  |     |   |     |     |
| <u>Sg-contig_152</u>        | 12 | SGL_P20_G02      | Collagen-like family   | 702 | 2 | 776 | 865 |
| <u>Sg-contig_153</u>        | 7  | SGL_P18_G01      | Collagen-like family   | 693 | 1 | 767 | 745 |
| <u>Sg-contig_148</u>        | 2  | SGL_P12_H06      | Collagen-like family   | 891 | 1 | 964 | 944 |
| <u>Sg-contig_151</u>        | 2  | SGL_P12_D07      | Collagen-like family   | 702 | 3 | 776 | 758 |
| <u>Sg-contig_149</u>        | 11 | SGL_P13_C04      | Collagen-like family   | 546 | 2 | 633 | 635 |
| <u>Sg-contig_150</u>        | 3  | SGM_P4_G06       | Collagen-like family   | 618 | 2 | 692 | 701 |
| <u>Sg-contig_159</u>        | 1  | SGL_P15_A02      | Collagen-like family   | 708 | 1 | 778 | 759 |
| <u>Sg-contig_160</u>        | 1  | SGL_P11_C12      | Collagen-like family   | 702 | 1 | 772 | 753 |
| <u>Sg-contig_162</u>        | 1  | SGM_P3_F02       | Collagen-like family   | 630 | 3 | 702 | 683 |
| <u>Sg-contig_156</u>        | 1  | SGL_P13_F10      | Collagen-like family   | 909 | 3 | 991 | 972 |
| <u>Sg-contig_157</u>        | 1  | SGL_P13_H05      | Collagen-like family   | 888 | 3 | 967 | 948 |
| <u>Sg-contig_161</u>        | 1  | SGL_P17_B01      | Collagen-like family   | 705 | 1 | 778 | 759 |
| <u>Sg-contig_154</u>        | 1  | SG_P1testing_E02 | Collagen-like family   | 822 | 3 | 897 | 878 |
| <u>Sg-contig_155</u>        | 1  | SGL_P18_D09      | Collagen-like family   | 786 | 3 | 858 | 839 |
| <u>Sg-contig_158</u>        | 1  | SGL_P17_B04      | Collagen-like family   | 330 | 2 | 713 | 694 |
| <u>Sg-contig_163</u>        | 1  | SGM_P3_E04       | Collagen-like family   | 558 | 1 | 631 | 612 |
| <u>Sg-contig_164</u>        | 1  | SGL_P12_C11      | Collagen-like family   | 507 | 2 | 651 | 632 |
| <b>Housekeeping</b>         |    |                  |  |     |   |     |     |
| <b>Protein synthesis</b>    |    |                  |  |     |   |     |     |
| <u>Sg-contig_259</u>        | 28 | SGL_P16_F05      | rRNA   | 177 | 2 | 748 | 734 |
| <u>Sg-contig_390</u>        | 3  | SGM_P6_E03       | Ribosomal protein L41  | 219 | 3 | 420 | 370 |
| <u>Sg-contig_449</u>        | 2  | SGL_P17_E02      | 60S ribosomal protein L23  | 429 | 3 | 547 | 528 |
| <u>Sg-contig_441</u>        | 3  | SGL_P12_G10      | ribosomal protein L13  | 765 | 3 | 866 | 860 |
| <u>Sg-contig_452</u>        | 3  | SGM_P6_B09       | 60S ribosomal protein L35  | 360 | 3 | 452 | 164 |
| <u>Sg-contig_442</u>        | 3  | SGL_P14_G03      | Elongation factor 1alpha48D  | 324 | 3 | 936 | 923 |
| <u>Sg-contig_377</u>        | 3  | SGS_P2_H04       | 60S acidic ribosomal protein P1                                    | 342 | 1 | 543 | 528 |
| <u>Sg-contig_378</u>        | 2  | SGM_P7_B01       | 60S acidic ribosomal protein P1                                    | 342 | 2 | 589 | 570 |
| <u>Sg-contig_459</u>        | 2  | SGS_P2_E10       | 60S ribosomal protein L7   | 75  | 2 | 175 | 156 |
| <u>Sg-contig_454</u>        | 2  | SGS_P1_A06       | 40S ribosomal protein S3   | 606 | 1 | 733 | 716 |
| <u>Sg-contig_433</u>        | 2  | SGM_P6_D08       | 40S ribosomal protein S14  | 183 | 1 | 218 | 200 |
| <u>Sg-contig_391</u>        | 1  | SGL_P11_C06      | Ribosomal protein L41  | 201 | 3 | 409 | 390 |
| <u>Sg-contig_392</u>        | 1  | SGL_P20_A03      | Ribosomal protein L41  | 219 | 2 | 427 | 408 |
| <u>Sg-contig_393</u>        | 1  | SGL_P14_F11      | Ribosomal protein L41  | 201 | 2 | 426 | 407 |
| <u>Sg-contig_381</u>        | 1  | SGL_P20_C04      | 60S acidic ribosomal protein P1                                    | 342 | 2 | 588 | 569 |
| <u>Sg-contig_394</u>        | 1  | SGM_P3_F10       | Ribosomal protein L41  | 147 | 1 | 445 | 426 |
| <u>Sg-contig_379</u>        | 1  | SGL_P15_G02      | 60S acidic ribosomal protein P1                                    | 342 | 1 | 607 | 588 |
| <u>Sg-contig_380</u>        | 1  | SGL_P19_D08      | 60S acidic ribosomal protein P1                                    | 342 | 2 | 578 | 559 |
| <u>Sg-contig_450</u>        | 1  | SGL_P15_B05      | 60S ribosomal protein L23  | 231 | 3 | 287 | 268 |
| <u>Sg-contig_556</u>        | 1  | SGL_P17_D08      | 60S ribosomal protein L11  | 285 | 1 | 325 | 306 |
| <u>Sg-contig_544</u>        | 1  | SGL_P16_F01      | 40S ribosomal protein S20  | 117 | 1 | 191 | 172 |
| <u>Sg-contig_742</u>        | 1  | SGS_P2_C01       | 60S ribosomal protein L44  | 138 | 2 | 209 | 190 |
| <u>Sg-contig_557</u>        | 1  | SGL_P17_E01      | 60S acidic ribosomal protein P2                                    | 210 | 2 | 279 | 260 |
| <u>Sg-contig_563</u>        | 1  | SGL_P17_H07      | 40S ribosomal protein S16  | 273 | 2 | 364 | 345 |
| <u>Sg-contig_741</u>        | 1  | SGS_P2_B10       | 60S acidic ribosomal protein P0                                    | 534 | 2 | 679 | 660 |
| <u>Sg-contig_434</u>        | 1  | SGM_P9_D12       | 40S ribosomal protein S14  | 168 | 2 | 204 | 185 |
| <u>Sg-contig_612</u>        | 1  | SGM_P10_C12      | Uncharacterized conserved protein related to ribosomal protein S8E | 330 | 1 | 459 | 440 |
| <u>Sg-contig_633</u>        | 1  | SGM_P3_G05       | 60s ribosomal protein L23  | 99  | 1 | 172 | 153 |
| <u>Sg-contig_586</u>        | 1  | SGL_P19_E09      | ribosomal protein L3   | 492 | 1 | 580 | 561 |
| <u>Sg-contig_634</u>        | 1  | SGM_P3_G12       | 60S ribosomal protein L5   | 156 | 1 | 175 | 156 |
| <u>Sg-contig_663</u>        | 1  | SGM_P6_E08       | 40S ribosomal protein S8   | 606 | 3 | 768 | 749 |
| <u>Sg-contig_532</u>        | 1  | SGL_P15_D11      | Polypeptide release factor 3                                       | 417 | 3 | 689 | 670 |
| <u>Sg-contig_474</u>        | 1  | SGL_P11_B01      | 60S acidic ribosomal protein P2                                    | 228 | 2 | 385 | 366 |

|                          |   |                  |   |     |   |      |      |
|--------------------------|---|------------------|---|-----|---|------|------|
| <u>Sg-contig_511</u>     | 1 | SGL_P13_C08      | 40S ribosomal protein S15/S22   | 261 | 2 | 325  | 306  |
| <u>Sg-contig_581</u>     | 1 | SGL_P19_B11      | 40S ribosomal protein S2  | 441 | 3 | 514  | 495  |
| <u>Sg-contig_683</u>     | 1 | SGM_P8_C01       | 40S ribosomal protein S15/S22   | 171 | 1 | 446  | 427  |
| <u>Sg-contig_572</u>     | 1 | SGL_P18_F03      | Ribosomal protein S7e   | 405 | 1 | 500  | 481  |
| <u>Sg-contig_745</u>     | 1 | SGS_P2_D10       | Ribosomal L38e protein family   | 270 | 3 | 462  | 443  |
| <u>Sg-contig_587</u>     | 1 | SGL_P19_F01      | Aspartyl-tRNA synthetase  | 939 | 2 | 1034 | 1015 |
| <u>Sg-contig_752</u>     | 1 | SGS_P2_H01       | 60S ribosomal protein L30   | 105 | 1 | 161  | 142  |
| <u>Sg-contig_658</u>     | 1 | SGM_P6_B03       | 60S ribosomal protein L27a  | 483 | 3 | 588  | 569  |
| <u>Sg-contig_443</u>     | 1 | SGL_P13_G04      | S25 ribosomal protein   | 252 | 2 | 431  |      |
| <u>Sg-contig_485</u>     | 1 | SGL_P11_H01      | 40S ribosomal protein S4  | 99  | 3 | 205  |      |
| <u>Sg-contig_700</u>     | 1 | SGM_P9_B01       | 40S ribosomal protein S2  | 435 | 3 | 509  | 490  |
| <u>Sg-contig_445</u>     | 1 | SGL_P18_G07      | 40S ribosomal protein S25   | 357 | 2 | 571  | 552  |
| <u>Sg-contig_598</u>     | 1 | SGL_P20_D06      | 40S ribosomal protein S29   | 294 | 3 | 431  | 412  |
| <u>Sg-contig_559</u>     | 1 | SGL_P17_F08      | Poly-adenylate binding protein, unique domain   | 345 | 1 | 400  | 381  |
| <u>Sg-contig_622</u>     | 1 | SGM_P10_G10      | glycyl-tRNA synthetase  | 180 | 2 | 234  | 215  |
| <u>Sg-contig_555</u>     | 1 | SGL_P17_D07      | ribosomal protein L35   | 573 | 1 | 656  | 637  |
| <u>Sg-contig_607</u>     | 1 | SGM_P10_A08      | 60S ribosomal protein L21   | 417 | 3 | 558  | 539  |
| <u>Sg-contig_604</u>     | 1 | SGL_P20_G06      | rRNA  | 162 | 1 | 517  | 498  |
| <u>Sg-contig_729</u>     | 1 | SGS_P1_F09       | Ribosomal protein L36e  | 198 | 1 | 250  | 231  |
| <u>Sg-contig_667</u>     | 1 | SGM_P6_G08       | Ubiquitin-40S ribosomal protein S27a  | 471 | 3 | 696  | 677  |
| <u>Sg-contig_548</u>     | 1 | SGL_P17_A08      | Eukaryotic translation initiation factor 5A   | 504 | 3 | 1044 | 1025 |
| <u>Sg-contig_664</u>     | 1 | SGM_P6_F04       | 60s ribosomal protein L15   | 282 | 3 | 363  |      |
| <u>Sg-contig_608</u>     | 1 | SGM_P10_A09      | 40S ribosomal protein SA  | 216 | 3 | 323  | 304  |
| <u>Sg-contig_479</u>     | 1 | SGL_P11_D12      | Eukaryotic translation initiation factor 5A ubiquitin-ribosomal protein S30e fusion protein | 504 | 2 | 1032 | 1013 |
| <u>Sg-contig_478</u>     | 1 | SGL_P11_D11      | 60s ribosomal protein L15   | 240 | 2 | 361  | 342  |
| <u>Sg-contig_567</u>     | 1 | SGL_P18_A09      | 60S ribosomal protein L14   | 423 | 1 | 514  | 495  |
| <u>Sg-contig_590</u>     | 1 | SGL_P19_G04      | 60S ribosomal protein L17   | 345 | 2 | 451  | 432  |
| <u>Sg-contig_530</u>     | 1 | SGL_P15_C03      | 60s ribosomal protein L10   | 219 | 2 | 336  | 317  |
| <u>Sg-contig_571</u>     | 1 | SGL_P18_D05      | 40S ribosomal S30 protein-like protein  | 684 | 3 | 865  | 846  |
| <u>Sg-contig_601</u>     | 1 | SGL_P20_E03      | 50S ribosomal protein L1  | 426 | 3 | 549  | 530  |
| <u>Sg-contig_747</u>     | 1 | SGS_P2_E07       | 40S ribosomal protein S17   | 369 | 1 | 459  | 440  |
| <u>Sg-contig_628</u>     | 1 | SGM_P3_C01       | S25 ribosomal protein   | 264 | 3 | 382  |      |
| <u>Sg-contig_444</u>     | 1 | SGL_P18_D11      | 60S ribosomal protein L29-like protein  | 267 | 2 | 446  | 427  |
| <u>Sg-contig_540</u>     | 1 | SGL_P16_B11      | ribosomal protein L28e-like protein   | 240 | 2 | 385  | 366  |
| <u>Sg-contig_521</u>     | 1 | SGL_P13_G07      | Elongation factor 2b  | 387 | 1 | 461  | 442  |
| <u>Sg-contig_465</u>     | 1 | SG_P1testing_D01 | 60S ribosomal protein L29-like protein  | 480 | 1 | 703  | 684  |
| <u>Sg-contig_535</u>     | 1 | SGL_P15_H09      | 40S ribosomal protein S15   | 240 | 3 | 423  | 404  |
| <u>Sg-contig_599</u>     | 1 | SGL_P20_D07      |   | 327 | 3 | 470  | 451  |
| <b>Energy metabolism</b> |   |                  |   |     |   |      |      |
| <u>Sg-contig_432</u>     | 3 | SGL_P14_F09      | NADH dehydrogenase subunit 1  | 327 | 3 | 836  | 822  |
| <u>Sg-contig_451</u>     | 3 | SGM_P7_E02       | cytochrome oxidase subunit 1  | 291 | 2 | 739  | 742  |
| <u>Sg-contig_428</u>     | 2 | SGM_P6_C03       | cytochrome oxidase subunit III  | 279 | 1 | 811  | 825  |
| <u>Sg-contig_430</u>     | 1 | SGM_P5_D03       | Cytochrome c oxidase subunit III  | 279 | 3 | 536  | 517  |
| <u>Sg-contig_429</u>     | 1 | SGM_P6_G03       | cytochrome c oxidase subunit III  | 342 | 1 | 565  | 546  |
| <u>Sg-contig_481</u>     | 1 | SGL_P11_F02      | Cytochrome c oxidase subunit 2  | 174 | 2 | 439  | 420  |
| <u>Sg-contig_645</u>     | 1 | SGM_P4_H05       | NADH dehydrogenase subunits 2, 5,   | 135 | 2 | 317  | 298  |
| <u>Sg-contig_576</u>     | 1 | SGL_P18_G02      | NADH dehydrogenase subunit 2  | 186 | 1 | 591  | 572  |
| <u>Sg-contig_517</u>     | 1 | SGL_P13_E10      | cytochrome c oxidase subunit Va/COX6  | 231 | 3 | 367  | 348  |
| <u>Sg-contig_546</u>     | 1 | SGL_P16_H10      | Acyl carrier protein/NADH-ubiquinone oxidoreductase, NDUFB1/SDAP subunit                    | 177 | 3 | 303  | 284  |
| <u>Sg-contig_616</u>     | 1 | SGM_P10_E08      | NADH dehydrogenase subunits 2, 5,   | 837 | 1 | 1296 |      |
| <u>Sg-contig_676</u>     | 1 | SGM_P7_F07       | NADH-cytochrome B5 reductase  | 195 | 3 | 264  |      |
| <u>Sg-contig_495</u>     | 1 | SGL_P12_C05      | cytochrome c oxidase subunit I  | 327 | 2 | 1352 |      |
| <u>Sg-contig_648</u>     | 1 | SGM_P5_B11       | Ubiquinol cytochrome c reductase, subunit QCR8  | 285 | 3 | 466  | 447  |
| <u>Sg-contig_677</u>     | 1 | SGM_P7_F10       | cytochrome b  | 534 | 2 | 1095 | 1076 |
| <u>Sg-contig_488</u>     | 1 | SGL_P12_A12      | cytochrome b  | 267 | 1 | 794  | 775  |
| <u>Sg-contig_626</u>     | 1 | SGM_P3_A11       | NADH-ubiquinone oxidoreductase B15 subunit (NDUFB4)   | 375 | 1 | 457  | 438  |
| <u>Sg-contig_673</u>     | 1 | SGM_P7_D07       | cytochrome c oxidase subunit VIb  | 201 | 1 | 324  | 305  |
| <u>Sg-contig_707</u>     | 1 | SGM_P9_C11       | NADH:ubiquinone oxidoreductase, NDUFS5/15kDa  | 378 | 2 | 530  | 511  |
| <u>Sg-contig_649</u>     | 1 | SGM_P5_C08       | NADH-ubiquinone oxidoreductase, subunit   | 177 | 3 | 236  | 217  |

|                                       |   |                  |  |   |     |   |      |      |
|---------------------------------------|---|------------------|--|---|-----|---|------|------|
|                                       |   |                  |  | NDUFB10/PDSW  |     |   |      |      |
| <a href="#">Sg-contig_493</a>         | 1 | SGL_P12_B10      |  | Ubiquinol-cytochrome C reductase, UQCRX/QCR9 like   | 252 | 3 | 404  | 385  |
| <a href="#">Sg-contig_538</a>         | 1 | SGL_P16_A05      |  | F0F1-type ATP synthase, beta subunit  | 312 | 3 | 1237 |      |
| <a href="#">Sg-contig_661</a>         | 1 | SGM_P6_D05       |  | NADH-ubiquinone oxidoreductase chain 6 CDGSH iron-sulfur domain-containing                                | 387 | 2 | 471  | 452  |
| <a href="#">Sg-contig_473</a>         | 1 | SG_P1testing_H03 |  | protein 3, mitochondrial  | 399 | 2 | 551  | 532  |
| <a href="#">Sg-contig_639</a>         | 1 | SGM_P4_C02       |  | NADH:ubiquinone oxidoreductase, NDUFA8/PGIV/19 kDa subunit  | 435 | 1 | 629  | 610  |
| <a href="#">Sg-contig_491</a>         | 1 | SGL_P12_B03      |  | ATPase subunit 6  | 141 | 1 | 168  | 149  |
| <a href="#">Sg-contig_469</a>         | 1 | SG_P1testing_F03 |  | similar to malate dehydrogenase   | 129 | 2 | 410  | 391  |
| <a href="#">Sg-contig_463</a>         | 1 | SG_P1testing_B07 |  | cytochrome c oxidase subunit Vb/COX4  | 366 | 1 | 779  | 760  |
| <b>Signal transduction</b>            |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_401</a>         | 4 | SGM_P3_F05       |  | DNA-binding nuclear phosphoprotein p8   | 327 | 2 | 721  | 692  |
| <a href="#">Sg-contig_400</a>         | 2 | SGM_P4_A02       |  | DNA-binding nuclear phosphoprotein p8   | 270 | 3 | 710  | 692  |
| <a href="#">Sg-contig_402</a>         | 1 | SGM_P4_H01       |  | DNA-binding nuclear phosphoprotein p8   | 270 | 2 | 675  | 656  |
| <a href="#">Sg-contig_566</a>         | 1 | SGL_P18_A07      |  | Renin receptor-like protein   | 216 | 2 | 579  | 560  |
| <a href="#">Sg-contig_501</a>         | 1 | SGL_P12_E09      |  | Selenophosphate synthetase  | 369 | 3 | 620  | 601  |
| <a href="#">Sg-contig_560</a>         | 1 | SGL_P17_G02      |  | Ras-specific guanine nucleotide-releasing factor RalGPS1  | 423 | 1 | 641  | 622  |
| <a href="#">Sg-contig_486</a>         | 1 | SGL_P11_H02      |  | NADH-ubiquinone reductase complex 1 MLRQ subunit  | 276 | 2 | 493  | 474  |
| <a href="#">Sg-contig_496</a>         | 1 | SGL_P12_C06      |  | NADH dehydrogenase  | 276 | 1 | 509  | 490  |
| <a href="#">Sg-contig_719</a>         | 1 | SGS_P1_A09       |  | Serine-threonine phosphatase 2B, catalytic subunit  | 381 | 3 | 1125 | 1106 |
| <a href="#">Sg-contig_738</a>         | 1 | SGS_P2_B04       |  | Serine-threonine phosphatase 2B, catalytic subunit  | 381 | 3 | 1112 | 1093 |
| <a href="#">Sg-contig_505</a>         | 1 | SGL_P12_G04      |  | 40S ribosomal protein sa (P40) (34/67 kDa laminin receptor) (colon carcinoma laminin-binding protein)     | 870 | 1 | 1029 | 1010 |
| <a href="#">Sg-contig_561</a>         | 1 | SGL_P17_G03      |  | CDC42 Rho GTPase-activating protein   | 330 | 1 | 1284 |      |
| <a href="#">Sg-contig_475</a>         | 1 | SGL_P11_B09      |  | Putative gustatory receptor 43a   | 354 | 3 | 456  | 437  |
| <a href="#">Sg-contig_617</a>         | 1 | SGM_P10_F01      |  | G protein beta subunit-like protein   | 414 | 2 | 487  | 468  |
| <b>Transporters and storage</b>       |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_547</a>         | 1 | SGL_P17_A07      |  | Vacuolar H <sup>+</sup> -ATPase V0 sector, subunit a  | 561 | 3 | 1135 | 1116 |
| <a href="#">Sg-contig_731</a>         | 1 | SGS_P1_G07       |  | Mitochondrial ADP/ATP carrier protein   | 222 | 3 | 644  | 625  |
| <a href="#">Sg-contig_462</a>         | 1 | SG_P1testing_B03 |  | Mitochondrial carrier protein   | 276 | 3 | 496  | 477  |
| <a href="#">Sg-contig_611</a>         | 1 | SGM_P10_C10      |  | Vacuolar H <sup>+</sup> -ATPase V1 sector, subunit F  | 345 | 1 | 685  | 666  |
| <a href="#">Sg-contig_662</a>         | 1 | SGM_P6_D06       |  | vitellogenin 1  | 372 | 1 | 543  | 524  |
| <b>Proteasome machinery</b>           |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_640</a>         | 1 | SGM_P4_C06       |  | proteasome subunit alpha type 2   | 99  | 3 | 157  | 138  |
| <b>Protein modification machinery</b> |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_621</a>         | 1 | SGM_P10_G06      |  | Oligosaccharyltransferase, beta subunit   | 366 | 1 | 456  | 437  |
| <a href="#">Sg-contig_472</a>         | 1 | SG_P1testing_H01 |  | Cysteine proteinase Cathepsin L   | 216 | 3 | 372  | 353  |
| <a href="#">Sg-contig_503</a>         | 1 | SGL_P12_F09      |  | Peptidyl-prolyl cis-trans isomerase   | 330 | 2 | 500  | 481  |
| <a href="#">Sg-contig_695</a>         | 1 | SGM_P8_H05       |  | asparagine-linked glycosylation 14 homolog  | 705 | 2 | 825  | 806  |
| <a href="#">Sg-contig_637</a>         | 1 | SGM_P4_B01       |  | Glycoprotein 6-alpha-L-fucosyltransferase   | 435 | 2 | 510  | 491  |
| <a href="#">Sg-contig_574</a>         | 1 | SGL_P18_F08      |  | Defender against cell death protein/oligosaccharyltransferase, epsilon subunit                            | 402 | 2 | 642  | 623  |
| <b>Protein export machinery</b>       |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_539</a>         | 1 | SGL_P16_B07      |  | Vesicle coat complex COPI, zeta subunit   | 195 | 1 | 484  | 465  |
| <a href="#">Sg-contig_618</a>         | 1 | SGM_P10_F03      |  | Membrane trafficking protein, emp24/gp25L/p24 family  | 189 | 1 | 526  | 507  |
| <a href="#">Sg-contig_732</a>         | 1 | SGS_P1_G09       |  | translocon-associated protein, beta subunit precursor (trap-beta) (signal sequence receptor beta subunit) | 216 | 2 | 319  | 300  |
| <a href="#">Sg-contig_727</a>         | 1 | SGS_P1_E01       |  | Translocon-associated protein beta (TRAPB)  | 633 | 2 | 776  | 757  |
| <b>Nuclear regulation</b>             |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_580</a>         | 1 | SGL_P19_B07      |  | cyclin g  | 162 | 2 | 486  | 467  |
| <a href="#">Sg-contig_537</a>         | 1 | SGL_P16_A03      |  | Chromatin-associated protein Dek  | 345 | 1 | 432  | 413  |
| <b>Nucleotide metabolism</b>          |   |                  |  |   |     |   |      |      |
| <a href="#">Sg-contig_638</a>         | 1 | SGM_P4_B09       |  | Dihydropyrimidine dehydrogenase   | 732 | 2 | 828  | 809  |



|                                  |    |                  |  |      |   |      |      |
|----------------------------------|----|------------------|--|------|---|------|------|
| Sg-contig_672                    | 1  | SGM_P7_D05       | Protein involved in ubiquinone biosynthesis    | 579  | 1 | 647  | 628  |
| <b>Transcription machinery</b>   |    |                  |  |      |   |      |      |
| Sg-contig_552                    | 1  | SGL_P17_C09      | DNA-directed RNA polymerase subunit beta       | 165  | 1 | 737  | 718  |
| Sg-contig_669                    | 1  | SGM_P7_B07       | RNA-binding protein FUS                        | 342  | 2 | 457  | 438  |
| <b>Carbohydrate metabolism</b>   |    |                  |  |      |   |      |      |
| Sg-contig_484                    | 1  | SGL_P11_G12      | Glyceraldehyde 3-phosphate dehydrogenase       | 183  | 2 | 586  | 567  |
| Sg-contig_527                    | 1  | SGL_P14_B11      | 3-phosphoglycerate kinase                      | 198  | 2 | 241  | 222  |
| Sg-contig_497                    | 1  | SGL_P12_C10      | UDP-glucuronosyl and UDP-glucosyl transferase  | 504  | 1 | 767  | 748  |
| <b>Amino acid metabolism</b>     |    |                  |  |      |   |      |      |
| Sg-contig_686                    | 1  | SGM_P8_D10       | Dipeptidyl aminopeptidase                      | 240  | 1 | 291  |      |
| <b>Detoxification metabolism</b> |    |                  |  |      |   |      |      |
| Sg-contig_569                    | 1  | SGL_P18_C07      | Reactive oxygen species modulator 1            | 267  | 3 | 430  | 411  |
| Sg-contig_606                    | 1  | SGM_P10_A03      | superoxide dismutase 3                         | 768  | 2 | 1108 | 1089 |
| Sg-contig_520                    | 1  | SGL_P13_G02      | microsomal glutathione s-transferase           | 345  | 3 | 402  |      |
| <b>Cytoskeletal</b>              |    |                  |  |      |   |      |      |
| Sg-contig_466                    | 1  | SG_P1testing_D03 | Tubulin alpha-3 chain                          | 243  | 2 | 265  | 246  |
| Sg-contig_468                    | 1  | SG_P1testing_D11 | Flocculation protein FLO11                     | 432  | 2 | 640  | 621  |
| <b>Unknown conserved</b>         |    |                  |  |      |   |      |      |
| Sg-contig_292                    | 18 | SGL_P17_H04      | Conserved hypothetical protein                 | 1578 | 1 | 1753 | 1062 |
| Sg-contig_256                    | 30 | SGL_P13_B02      | Conserved hypothetical protein                 | 1014 | 1 | 1047 | 1029 |
| Sg-contig_347                    | 10 | SGM_P6_C05       | Conserved hypothetical protein                 | 837  | 2 | 914  | 880  |
| Sg-contig_431                    | 4  | SGM_P6_H03       | hypothetical conserved protein                 | 291  | 1 | 405  | 222  |
| Sg-contig_372                    | 3  | SGM_P3_A12       | hypothetical secreted peptide precursor        | 405  | 3 | 448  | 205  |
| Sg-contig_386                    | 3  | SGL_P11_B06      | hypothetical conserved protein                 | 384  | 1 | 436  | 417  |
| Sg-contig_387                    | 2  | SGM_P9_B10       | hypothetical conserved protein                 | 129  | 2 | 240  | 153  |
| Sg-contig_255                    | 2  | SGM_P7_H09       | Conserved hypothetical protein                 | 573  | 2 | 600  | 591  |
| Sg-contig_291                    | 2  | SGL_P12_A01      | Conserved hypothetical protein                 | 705  | 3 | 871  | 833  |
| Sg-contig_389                    | 1  | SGM_P5_B01       | hypothetical conserved protein                 | 252  | 2 | 294  | 275  |
| Sg-contig_388                    | 1  | SGL_P15_B12      | hypothetical conserved protein                 | 243  | 1 | 284  | 265  |
| Sg-contig_348                    | 1  | SGM_P5_E06       | Conserved hypothetical protein                 | 213  | 1 | 268  | 249  |
| Sg-contig_293                    | 1  | SGM_P10_F05      | Conserved hypothetical protein                 | 198  | 1 | 336  | 317  |
| Sg-contig_349                    | 1  | SGL_P18_B03      | Conserved hypothetical protein                 | 240  | 1 | 280  | 261  |
| Sg-contig_294                    | 1  | SGL_P15_A03      | Conserved hypothetical protein                 | 222  | 2 | 386  | 367  |
| Sg-contig_644                    | 1  | SGM_P4_H02       | hypothetical conserved protein                 | 195  | 2 | 297  | 278  |
| Sg-contig_654                    | 1  | SGM_P5_G03       | hypothetical conserved protein                 | 369  | 1 | 722  | 703  |
| Sg-contig_461                    | 1  | SG_P1testing_A12 | hypothetical protein EAI_11049                 | 117  | 1 | 371  | 352  |
| Sg-contig_585                    | 1  | SGL_P19_D10      | conserved hypothetical protein                 | 249  | 1 | 432  |      |
| Sg-contig_516                    | 1  | SGL_P13_E06      | WW domain-containing protein                   | 558  | 1 | 1339 |      |
| Sg-contig_523                    | 1  | SGL_P13_G12      | hypothetical conserved protein                 | 438  | 1 | 938  | 919  |
| Sg-contig_682                    | 1  | SGM_P8_B04       | conserved hypothetical protein                 | 270  | 2 | 406  | 387  |
| Sg-contig_689                    | 1  | SGM_P8_F05       | hypothetical conserved protein                 | 282  | 1 | 383  | 364  |
| Sg-contig_610                    | 1  | SGM_P10_B12      | hypothetical conserved protein                 | 120  | 1 | 153  |      |
| Sg-contig_490                    | 1  | SGL_P12_B02      | hypothetical salivary protein                  | 300  | 1 | 492  | 473  |
| Sg-contig_508                    | 1  | SGL_P12_H04      | Uncharacterized conserved protein              | 327  | 1 | 526  | 507  |
| Sg-contig_678                    | 1  | SGM_P7_G01       | hypothetical conserved protein with Ig domains | 204  | 1 | 261  | 242  |
| Sg-contig_709                    | 1  | SGM_P9_D04       | hypothetical conserved protein                 | 696  | 3 | 927  | 908  |
| <b>Transposable element</b>      |    |                  |  |      |   |      |      |
| Sg-contig_541                    | 1  | SGL_P16_C03      | similar to gag-pol protein                     | 255  | 2 | 446  | 427  |
| <b>Bacterial contaminant</b>     |    |                  |  |      |   |      |      |
| Sg-contig_588                    | 1  | SGL_P19_F03      | bacterial contaminant                          | 123  | 1 | 124  |      |
| <b>Unknown products</b>          |    |                  |  |      |   |      |      |
| Sg-contig_358                    | 2  | SGM_P6_E07       | Unkown   | 87   | 2 | 144  | 125  |
| Sg-contig_635                    | 1  | SGM_P3_H03       | Unkown   | 66   | 2 | 122  | 103  |
| Sg-contig_730                    | 1  | SGS_P1_G05       | Unkown   | 51   | 2 | 106  | 87   |
| Sg-contig_694                    | 1  | SGM_P8_H03       | Unkown   | 78   | 3 | 171  | 152  |
| Sg-contig_506                    | 1  | SGL_P12_G12      | Unkown   | 84   | 3 | 194  |      |
| Sg-contig_518                    | 1  | SGL_P13_F05      | Unkown   | 108  | 1 | 168  | 149  |
| Sg-contig_558                    | 1  | SGL_P17_E04      | Unkown   | 84   | 3 | 89   | 70   |

|                      |   |                  |                 |     |   |     |     |
|----------------------|---|------------------|-----------------|-----|---|-----|-----|
| <u>Sg-contig_583</u> | 1 | SGL_P19_C10      | Unkown          | 156 | 2 | 159 | 140 |
| <u>Sg-contig_614</u> | 1 | SGM_P10_D04      | Unkown          | 57  | 1 | 78  | 59  |
| <u>Sg-contig_623</u> | 1 | SGM_P10_H04      | Unkown          | 57  | 3 | 91  | 72  |
| <u>Sg-contig_646</u> | 1 | SGM_P5_A09       | Unkown          | 111 | 2 | 115 | 96  |
| <u>Sg-contig_657</u> | 1 | SGM_P5_H05       | Unkown          | 99  | 1 | 170 | 151 |
| <u>Sg-contig_659</u> | 1 | SGM_P6_B11       | Unkown          | 51  | 1 | 96  |     |
| <u>Sg-contig_670</u> | 1 | SGM_P7_C10       | Unkown          | 63  | 1 | 91  | 72  |
| <u>Sg-contig_680</u> | 1 | SGM_P8_A05       | Unkown          | 90  | 1 | 92  | 73  |
| <u>Sg-contig_688</u> | 1 | SGM_P8_F02       | Unkown          | 69  | 1 | 90  | 71  |
| <u>Sg-contig_692</u> | 1 | SGM_P8_G10       | Unkown          | 75  | 1 | 121 | 102 |
| <u>Sg-contig_706</u> | 1 | SGM_P9_C10       | Unkown          | 60  | 2 | 107 | 88  |
| <u>Sg-contig_716</u> | 1 | SGS_P1_A03       | Unkown          | 87  | 1 | 173 | 154 |
| <u>Sg-contig_733</u> | 1 | SGS_P1_H03       | Unkown          | 90  | 3 | 97  | 78  |
| <u>Sg-contig_734</u> | 1 | SGS_P1_H09       | Unkown          | 51  | 1 | 113 | 94  |
| <u>Sg-contig_739</u> | 1 | SGS_P2_B05       | Unkown          | 45  | 2 | 64  | 45  |
| <u>Sg-contig_740</u> | 1 | SGS_P2_B09       | Unkown          | 90  | 2 | 95  | 76  |
| <u>Sg-contig_743</u> | 1 | SGS_P2_C10       | Unkown          | 63  | 2 | 66  | 47  |
| <u>Sg-contig_748</u> | 1 | SGS_P2_F01       | Unkown          | 78  | 1 | 97  | 78  |
| <u>Sg-contig_753</u> | 1 | SGS_P2_H05       | Unkown          | 72  | 2 | 141 |     |
| <u>Sg-contig_630</u> | 1 | SGM_P3_D08       | Unkown          | 63  | 1 | 108 | 89  |
| <u>Sg-contig_533</u> | 1 | SGL_P15_F10      | Unkown          | 54  | 2 | 76  | 57  |
| <u>Sg-contig_613</u> | 1 | SGM_P10_D02      | Unkown          | 60  | 1 | 62  | 43  |
| <u>Sg-contig_631</u> | 1 | SGM_P3_D09       | Unkown          | 54  | 2 | 76  | 57  |
| <u>Sg-contig_656</u> | 1 | SGM_P5_G09       | Unkown          | 63  | 1 | 65  | 46  |
| <u>Sg-contig_699</u> | 1 | SGM_P9_A10       | Unkown          | 51  | 2 | 71  | 52  |
| <u>Sg-contig_708</u> | 1 | SGM_P9_D01       | Unkown          | 72  | 2 | 77  | 58  |
| <u>Sg-contig_713</u> | 1 | SGM_P9_F01       | Unkown          | 60  | 2 | 106 | 87  |
| <u>Sg-contig_714</u> | 1 | SGM_P9_G07       | Unkown          | 57  | 1 | 60  | 41  |
| <u>Sg-contig_717</u> | 1 | SGS_P1_A07       | Unkown          | 60  | 3 | 67  | 48  |
| <u>Sg-contig_735</u> | 1 | SGS_P2_A01       | Unkown          | 39  | 1 | 68  | 49  |
| <u>Sg-contig_749</u> | 1 | SGS_P2_F09       | Unkown          | 51  | 3 | 66  | 47  |
| <u>Sg-contig_408</u> | 3 | SG_P1testing_B02 | unknown product | 327 | 3 | 434 | 415 |
| <u>Sg-contig_440</u> | 3 | SGL_P12_A09      | unknown product | 168 | 1 | 378 | 369 |
| <u>Sg-contig_456</u> | 3 | SGL_P20_B02      | unknown product | 282 | 1 | 675 | 659 |
| <u>Sg-contig_458</u> | 2 | SGL_P16_C05      | unknown product | 174 | 1 | 578 | 559 |
| <u>Sg-contig_712</u> | 1 | SGM_P9_E05       | unknown product | 234 | 3 | 291 | 272 |
| <u>Sg-contig_573</u> | 1 | SGL_P18_F04      | unknown product | 96  | 3 | 169 | 150 |
| <u>Sg-contig_647</u> | 1 | SGM_P5_B10       | unknown product | 105 | 3 | 232 | 213 |
| <u>Sg-contig_750</u> | 1 | SGS_P2_G10       | unknown product | 198 | 3 | 424 | 405 |
| <u>Sg-contig_534</u> | 1 | SGL_P15_G11      | unknown product | 201 | 1 | 955 | 936 |
| <u>Sg-contig_619</u> | 1 | SGM_P10_F04      | unknown product | 45  | 1 | 91  | 72  |
| <u>Sg-contig_592</u> | 1 | SGL_P19_G09      | unknown product | 138 | 1 | 141 | 122 |
| <u>Sg-contig_684</u> | 1 | SGM_P8_D04       | unknown product | 96  | 1 | 149 | 130 |
| <u>Sg-contig_494</u> | 1 | SGL_P12_B12      | unknown product | 165 | 2 | 280 |     |
| <u>Sg-contig_685</u> | 1 | SGM_P8_D07       | unknown product | 162 | 2 | 225 | 206 |
| <u>Sg-contig_595</u> | 1 | SGL_P20_B05      | unknown product | 369 | 1 | 848 | 829 |
| <u>Sg-contig_737</u> | 1 | SGS_P2_A12       | unknown product | 96  | 3 | 125 | 106 |
| <u>Sg-contig_579</u> | 1 | SGL_P19_A07      | unknown product | 147 | 3 | 505 | 486 |
| <u>Sg-contig_514</u> | 1 | SGL_P13_E01      | unknown product | 231 | 2 | 643 | 624 |
| <u>Sg-contig_725</u> | 1 | SGS_P1_D08       | unknown product | 108 | 2 | 232 | 213 |
| <u>Sg-contig_624</u> | 1 | SGM_P3_A03       | unknown product | 72  | 2 | 137 | 118 |
| <u>Sg-contig_724</u> | 1 | SGS_P1_C08       | unknown product | 75  | 1 | 170 |     |
| <u>Sg-contig_642</u> | 1 | SGM_P4_F03       | unknown product | 63  | 3 | 102 | 83  |
| <u>Sg-contig_643</u> | 1 | SGM_P4_F08       | unknown product | 75  | 3 | 131 | 112 |
| <u>Sg-contig_455</u> | 1 | SGL_P19_H06      | unknown product | 105 | 2 | 241 | 222 |
| <u>Sg-contig_698</u> | 1 | SGM_P9_A08       | unknown product | 96  | 1 | 97  | 78  |
| <u>Sg-contig_722</u> | 1 | SGS_P1_B08       | unknown product | 102 | 1 | 103 | 84  |
| <u>Sg-contig_744</u> | 1 | SGS_P2_D08       | unknown product | 195 | 1 | 278 | 259 |
| <u>Sg-contig_632</u> | 1 | SGM_P3_F11       | unknown product | 99  | 3 | 173 | 154 |
| <u>Sg-contig_605</u> | 1 | SGL_P20_H06      | unknown product | 207 | 3 | 317 | 298 |
| <u>Sg-contig_668</u> | 1 | SGM_P7_A01       | unknown product | 99  | 3 | 105 | 86  |
| <u>Sg-contig_620</u> | 1 | SGM_P10_F11      | unknown product | 93  | 1 | 197 | 178 |
| <u>Sg-contig_470</u> | 1 | SG_P1testing_F08 | unknown product | 219 | 3 | 225 | 206 |
| <u>Sg-contig_691</u> | 1 | SGM_P8_G08       | unknown product | 138 | 2 | 258 | 239 |

|                      |   |                  |                      |     |   |      |      |
|----------------------|---|------------------|----------------------|-----|---|------|------|
| <u>Sg-contig_513</u> | 1 | SGL_P13_D08      | unknown product      | 141 | 2 | 393  | 374  |
| <u>Sg-contig_679</u> | 1 | SGM_P7_G11       | unknown product      | 165 | 2 | 168  | 149  |
| <u>Sg-contig_597</u> | 1 | SGL_P20_B11      | unknown product      | 180 | 3 | 322  |      |
| <u>Sg-contig_603</u> | 1 | SGL_P20_G03      | unknown product      | 189 | 3 | 528  | 509  |
| <u>Sg-contig_655</u> | 1 | SGM_P5_G04       | unknown product      | 228 | 2 | 412  | 393  |
| <u>Sg-contig_666</u> | 1 | SGM_P6_F11       | unknown product      | 171 | 2 | 521  | 502  |
| <u>Sg-contig_550</u> | 1 | SGL_P17_B08      | unknown product      | 144 | 2 | 212  | 193  |
| <u>Sg-contig_721</u> | 1 | SGS_P1_B07       | unknown product      | 135 | 2 | 138  | 119  |
| <u>Sg-contig_627</u> | 1 | SGM_P3_B03       | unknown product      | 105 | 1 | 367  |      |
| <u>Sg-contig_589</u> | 1 | SGL_P19_G03      | unknown product      | 144 | 1 | 211  | 192  |
| <u>Sg-contig_718</u> | 1 | SGS_P1_A08       | unknown product      | 117 | 1 | 234  | 215  |
| <u>Sg-contig_650</u> | 1 | SGM_P5_C11       | unknown product      | 141 | 2 | 157  | 138  |
| <u>Sg-contig_697</u> | 1 | SGM_P9_A07       | unknown product      | 78  | 3 | 139  | 120  |
| <u>Sg-contig_710</u> | 1 | SGM_P9_E01       | unknown product      | 159 | 2 | 353  |      |
| <u>Sg-contig_702</u> | 1 | SGM_P9_B11       | unknown product      | 147 | 2 | 273  | 254  |
| <u>Sg-contig_543</u> | 1 | SGL_P16_E08      | unknown product      | 318 | 1 | 353  | 334  |
| <u>Sg-contig_681</u> | 1 | SGM_P8_B01       | unknown product      | 162 | 3 | 324  | 305  |
| <u>Sg-contig_551</u> | 1 | SGL_P17_B12      | unknown product      | 219 | 1 | 757  | 738  |
| <u>Sg-contig_720</u> | 1 | SGS_P1_B02       | unknown product      | 93  | 3 | 122  | 103  |
| <u>Sg-contig_675</u> | 1 | SGM_P7_F02       | unknown product      | 84  | 3 | 89   | 70   |
| <u>Sg-contig_531</u> | 1 | SGL_P15_D10      | unknown product      | 96  | 2 | 225  | 206  |
| <u>Sg-contig_564</u> | 1 | SGL_P17_H10      | unknown product      | 201 | 2 | 375  | 356  |
| <u>Sg-contig_696</u> | 1 | SGM_P8_H06       | unknown product      | 105 | 1 | 243  | 224  |
| <u>Sg-contig_184</u> | 1 | SGM_P8_B09       | Unknown              | 129 | 3 | 334  | 315  |
| <u>Sg-contig_728</u> | 1 | SGS_P1_F03       | unknown product      | 111 | 3 | 129  | 110  |
| <u>Sg-contig_600</u> | 1 | SGL_P20_D09      | unknown product      | 111 | 2 | 515  | 496  |
| <u>Sg-contig_570</u> | 1 | SGL_P18_D01      | unknown product      | 126 | 2 | 308  | 289  |
| <u>Sg-contig_641</u> | 1 | SGM_P4_D04       | unknown product      | 183 | 1 | 498  | 479  |
| <u>Sg-contig_464</u> | 1 | SG_P1testing_B11 | unknown product      | 144 | 3 | 193  | 174  |
| <u>Sg-contig_703</u> | 1 | SGM_P9_B12       | unknown product      | 87  | 3 | 119  | 100  |
| <u>Sg-contig_515</u> | 1 | SGL_P13_E03      | unknown product      | 330 | 3 | 404  | 385  |
| <u>Sg-contig_512</u> | 1 | SGL_P13_C10      | unknown product      | 126 | 3 | 498  | 479  |
| <u>Sg-contig_499</u> | 1 | SGL_P12_D06      | unknown product      | 186 | 3 | 308  | 289  |
| <u>Sg-contig_483</u> | 1 | SGL_P11_G09      | unknown product      | 186 | 2 | 235  |      |
| <u>Sg-contig_726</u> | 1 | SGS_P1_D09       | unknown product      | 141 | 1 | 498  | 479  |
| <u>Sg-contig_575</u> | 1 | SGL_P18_F12      | unknown product      | 144 | 3 | 405  | 386  |
| <u>Sg-contig_625</u> | 1 | SGM_P3_A05       | unknown product      | 333 | 1 | 1281 |      |
| <u>Sg-contig_651</u> | 1 | SGM_P5_E08       | unknown product      | 195 | 3 | 292  |      |
| <u>Sg-contig_482</u> | 1 | SGL_P11_F05      | hypothetical protein | 93  | 2 | 343  | 324  |
| <u>Sg-contig_507</u> | 1 | SGL_P12_H01      | unknown product      | 288 | 1 | 682  | 663  |
| <u>Sg-contig_565</u> | 1 | SGL_P18_A02      | unknown product      | 138 | 3 | 337  | 318  |
| <u>Sg-contig_509</u> | 1 | SGL_P13_A02      | unknown product      | 192 | 3 | 443  |      |
| <u>Sg-contig_591</u> | 1 | SGL_P19_G05      | unknown product      | 294 | 1 | 622  | 603  |
| <u>Sg-contig_629</u> | 1 | SGM_P3_C05       | unknown product      | 159 | 3 | 428  | 409  |
| <u>Sg-contig_562</u> | 1 | SGL_P17_G10      | unknown product      | 138 | 3 | 342  | 323  |
| <u>Sg-contig_492</u> | 1 | SGL_P12_B08      | unknown product      | 201 | 2 | 681  | 662  |
| <u>Sg-contig_711</u> | 1 | SGM_P9_E02       | unknown product      | 159 | 2 | 259  | 240  |
| <u>Sg-contig_489</u> | 1 | SGL_P12_B01      | unknown product      | 96  | 2 | 298  |      |
| <u>Sg-contig_665</u> | 1 | SGM_P6_F06       | unknown product      | 153 | 1 | 442  | 423  |
| <u>Sg-contig_636</u> | 1 | SGM_P4_A06       | unknown product      | 189 | 1 | 959  | 940  |
| <u>Sg-contig_504</u> | 1 | SGL_P12_G01      | unknown product      | 237 | 1 | 1021 | 1002 |
| <u>Sg-contig_498</u> | 1 | SGL_P12_D04      | unknown product      | 135 | 2 | 904  | 885  |
| <u>Sg-contig_502</u> | 1 | SGL_P12_E10      | unknown product      | 207 | 1 | 382  | 363  |
| <u>Sg-contig_542</u> | 1 | SGL_P16_D07      | unknown product      | 129 | 3 | 245  | 226  |
| <u>Sg-contig_554</u> | 1 | SGL_P17_D06      | unknown product      | 159 | 2 | 603  | 584  |
| <u>Sg-contig_701</u> | 1 | SGM_P9_B02       | unknown product      | 240 | 1 | 334  |      |
| <u>Sg-contig_723</u> | 1 | SGS_P1_B09       | unknown product      | 237 | 1 | 390  | 371  |
| <u>Sg-contig_529</u> | 1 | SGL_P14_F01      | unknown product      | 87  | 3 | 430  | 411  |
| <u>Sg-contig_528</u> | 1 | SGL_P14_D06      | unknown product      | 186 | 3 | 222  | 203  |
| <u>Sg-contig_545</u> | 1 | SGL_P16_F03      | unknown product      | 108 | 1 | 385  | 366  |
| <u>Sg-contig_460</u> | 1 | SG_P1testing_A05 | unknown product      | 93  | 1 | 131  |      |
| <u>Sg-contig_578</u> | 1 | SGL_P18_H08      | unknown product      | 135 | 2 | 614  | 595  |
| <u>Sg-contig_471</u> | 1 | SG_P1testing_G08 | unknown product      | 120 | 1 | 190  | 171  |
| <u>Sg-contig_116</u> | 1 | SGS_P1_C12       | Unknown              | 165 | 1 | 178  | 159  |



|                      |   |             |                 |     |   |      |      |
|----------------------|---|-------------|-----------------|-----|---|------|------|
| <u>Sg-contig_584</u> | 1 | SGL_P19_D07 | unknown product | 117 | 3 | 301  | 282  |
| <u>Sg-contig_660</u> | 1 | SGM_P6_D03  | unknown product | 123 | 3 | 185  |      |
| <u>Sg-contig_693</u> | 1 | SGM_P8_H02  | unknown product | 141 | 2 | 293  | 274  |
| <u>Sg-contig_477</u> | 1 | SGL_P11_C11 | unknown product | 129 | 3 | 651  | 632  |
| <u>Sg-contig_746</u> | 1 | SGS_P2_E04  | unknown product | 144 | 2 | 147  | 128  |
| <u>Sg-contig_519</u> | 1 | SGL_P13_F07 | unknown product | 297 | 3 | 1289 | 1270 |
| <u>Sg-contig_582</u> | 1 | SGL_P19_C02 | unknown product | 183 | 3 | 718  | 699  |
| <u>Sg-contig_715</u> | 1 | SGM_P9_G11  | unknown product | 258 | 3 | 318  | 299  |
| <u>Sg-contig_687</u> | 1 | SGM_P8_E03  | unknown product | 156 | 2 | 161  | 142  |
| <u>Sg-contig_690</u> | 1 | SGM_P8_G03  | unknown product | 159 | 3 | 362  | 343  |
| <u>Sg-contig_704</u> | 1 | SGM_P9_C04  | unknown product | 306 | 3 | 635  | 616  |
| <u>Sg-contig_510</u> | 1 | SGL_P13_A06 | unknown product | 276 | 2 | 337  | 318  |
| <u>Sg-contig_536</u> | 1 | SGL_P15_H10 | unknown product | 321 | 2 | 450  | 431  |
| <u>Sg-contig_524</u> | 1 | SGL_P13_H03 | unknown product | 306 | 1 | 441  | 422  |
| <u>Sg-contig_577</u> | 1 | SGL_P18_G11 | unknown product | 171 | 1 | 490  | 471  |
| <u>Sg-contig_736</u> | 1 | SGS_P2_A11  | unknown product | 69  | 1 | 114  |      |

| Seq name   | Number of ESTs | Seq size | Seq size | First codon | * | Mature MW | pI    | SigP Result | Cleavage Position | N Tryptic peptides |
|--|----------------|----------|----------|-------------|---|-----------|-------|-------------|-------------------|--------------------|
| <b>Secreted proteins</b>   |                |          |          |             |   |           |       |             |                   |                    |
| <b>Ubiquitous protein families</b>   |                |          |          |             |   |           |       |             |                   |                    |
| <b>Amylase/Maltase</b>   |                |          |          |             |   |           |       |             |                   |                    |
| Sg-214   | 18             | 535      | 1608     | D           | * |           |       | CYT         |                   | 116                |
| Sg-296   | 5              | 417      | 1254     | G           | * |           |       | CYT         |                   | 32                 |
| <b>Serine proteases</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-138   | 7              | 260      | 783      | M           | * | 27,069    | 8,83  | SIG         | 26-27             | 28                 |
| Sg-244   | 6              | 242      | 729      | M           | * | 25,45     | 9,03  | SIG         | 18-19             | 48                 |
| Sg-416   | 2              | 263      | 792      | M           | * | 27,023    | 6,57  | SIG         | 16-17             | 6                  |
| <b>Hyaluronidase</b>   |                |          |          |             |   |           |       |             |                   |                    |
| Sg-414   | 5              | 342      | 1029     | M           | * | 37,805    | 9,2   | SIG         | 21-22             | 14                 |
| <b>Apyrase</b>   |                |          |          |             |   |           |       |             |                   |                    |
| Sg-354   | 10             | 304      | 915      | G           | * |           |       | CYT         |                   | 58                 |
| <b>Kunitz salivary protein</b>   |                |          |          |             |   |           |       |             |                   |                    |
| Sg-395   | 3              | 106      | 321      | M           |   | 10,087    | 9,45  | SIG         | 19-20             | 2                  |
| <b>Lysozyme</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-263   | 5              | 141      | 426      | M           | * | 13,748    | 9,22  | SIG         | 21-22             | 7                  |
| <b>Cecropin</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-369   | 6              | 56       | 171      | M           | * | 3,337     | 10,66 | SIG         | 23-24             | 2                  |
| Sg-368   | 4              | 56       | 171      | M           | * | 3,337     | 10,66 | SIG         | 23-24             | 2                  |
| <b>Antigen-5 family</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-453   | 3              | 204      | 615      | F           | * |           |       | CYT         |                   |                    |
| Sg-457   | 2              | 278      | 837      | M           | * | 30,006    | 9,36  | SIG         | 20-21             | 30                 |
| <b>Insect-specific proteins</b>  |                |          |          |             |   |           |       |             |                   |                    |
| <b>Aegyptin</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-276   | 6              | 290      | 873      | M           | * | 28,623    | 4,17  | SIG         | 20-21             | 37                 |
| <b>Diptera secreted protein from conserved insect family</b>                                 |                |          |          |             |   |           |       |             |                   |                    |
| Sg-256   | 30             | 334      | 1005     | Q           | * |           |       | CYT         |                   | 28                 |
| Sg-216   | 21             | 363      | 1092     | M           | * | 37,849    | 9,62  | SIG         | 19-20             | 133                |
| Sg-292   | 18             | 521      | 1566     | M           | * | 57,228    | 6,15  | SIG         | 17-18             | 49                 |
| Sg-215   | 16             | 363      | 1092     | M           | * | 37,981    | 9,61  | SIG         | 19-20             | 134                |
| Sg-347   | 10             | 233      | 702      | M           | * |           |       | CYT         |                   | 36                 |
| <b>Laminin-like secreted salivary protein found in <i>Simulium</i> and <i>Culicoides</i></b> |                |          |          |             |   |           |       |             |                   |                    |
| Sg-431   | 4              | 95       | 288      | S           | * |           |       | BL          |                   | 2                  |
| <b>Long D7 family</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-220   | 18             | 264      | 795      | M           | * | 28,352    | 4,67  | SIG         | 19-20             | 24                 |
| Sg-261   | 15             | 241      | 726      | M           | * | 24,898    | 4,84  | SIG         | 19-20             | 42                 |
| Sg-218   | 2              | 170      | 513      | Y           | * |           |       | CYT         |                   | 14                 |
| <b>D7 16 kDa family</b>  |                |          |          |             |   |           |       |             |                   |                    |
| Sg-350   | 5              | 141      | 426      | M           | * | 14,213    | 9,11  | SIG         | 16-17             | 9                  |
| Sg-331   | 2              | 166      | 501      | L           | * | 16,658    | 7,75  | SIG         | 17-18             | 14                 |
| <b>Ultra-short D7 proteins (10-12 kDa mature weight)</b>                                     |                |          |          |             |   |           |       |             |                   |                    |
| Sg-75  | 11             | 112      | 339      | M           | * | 10,817    | 4,43  | SIG         | 18-19             |                    |
| Sg-190   | 11             | 107      | 324      | M           | * | 10,706    | 4,54  | SIG         | 16-17             |                    |
| Sg-363   | 4              | 114      | 345      | M           | * | 11,246    | 4,79  | SIG         | 17-18             |                    |
| Sg-383   | 2              | 120      | 363      | M           | * | 12,088    | 5,88  | SIG         | 18-19             | 3                  |

| <i>Simulium</i> -specific proteins                                  |    |     |      |   |   |        |      |     |       |     |
|---|----|-----|------|---|---|--------|------|-----|-------|-----|
| SVEP/Marydilan  |    |     |      |   |   |        |      |     |       |     |
| Sg-8  | 20 | 152 | 459  | M | * | 15,069 | 6,57 | SIG | 20-21 | 149 |
| Sg-1  | 16 | 152 | 459  | M | * | 15,024 | 8,47 | SIG | 20-21 | 164 |
| Sg-102  | 7  | 149 | 450  | M | * | 14,941 | 6,17 | SIG | 19-20 | 43  |
| Sg-344  | 6  | 145 | 438  | M | * | 13,855 | 8,65 | SIG | 22-23 | 14  |
| Sg-95   | 5  | 149 | 450  | M | * | 14,742 | 7,94 | SIG | 19-20 | 26  |
| Sg-99   | 5  | 149 | 450  | M | * | 14,737 | 7,94 | SIG | 19-20 | 25  |
| Sg-90   | 4  | 98  | 297  | K | * |        |      | CYT |       | 25  |
| Sg-93   | 4  | 140 | 423  | L | * |        |      | CYT |       | 33  |
| <a href="#">Sg-92</a>   | 3  | 149 | 450  | M | * | 14,917 | 7,93 | SIG | 19-20 | 53  |
| Sg-100  | 3  | 140 | 423  | L | * |        |      | CYT |       | 32  |
| Sg-103  | 2  | 149 | 450  | M | * | 14,893 | 7,01 | SIG | 19-20 | 22  |
| Sg-94   | 2  | 149 | 450  | M | * | 14,846 | 8,54 | SIG | 19-20 | 34  |
| Sg-101  | 2  | 149 | 450  | M | * | 14,697 | 6,34 | SIG | 19-20 | 20  |
| H-rich, acidic proteins of <i>Simulium</i>                          |    |     |      |   |   |        |      |     |       |     |
| Sg-41   | 17 | 148 | 447  | M | * | 14,524 | 4,98 | SIG | 19-20 |     |
| Sg-30   | 15 | 152 | 459  | M | * | 15,119 | 5,05 | SIG | 19-20 |     |
| Sg-37   | 15 | 153 | 462  | M | * | 15,091 | 4,93 | SIG | 19-20 |     |
| Sg-33   | 13 | 159 | 480  | M | * | 15,979 | 5,06 | SIG | 19-20 |     |
| Sg-31   | 12 | 153 | 462  | M | * | 15,218 | 5,02 | SIG | 19-20 |     |
| Sg-38   | 7  | 149 | 450  | M | * | 14,685 | 5,21 | SIG | 19-20 |     |
| Sg-40   | 4  | 151 | 456  | M | * | 15,098 | 5,07 | SIG | 19-20 |     |
| Sg-36   | 4  | 166 | 501  | M | * | 16,827 | 4,94 | SIG | 19-20 |     |
| Sg-32   | 4  | 160 | 483  | M | * | 16,007 | 5,3  | SIG | 19-20 |     |
| Sg-34   | 2  | 154 | 465  | M | * | 15,35  | 4,95 | SIG | 19-20 |     |
| Sg-42   | 2  | 143 | 432  | M | * | 14,029 | 4,97 | SIG | 19-20 |     |
| <i>Simulium</i> mucin   |    |     |      |   |   |        |      |     |       |     |
| Sg-126  | 14 | 349 | 1050 | M | * | 33     | 4,45 | SIG | 21-22 | 12  |
| Sg-120  | 6  | 324 | 975  | V | * |        |      | CYT |       | 11  |
| Sg-117  | 4  | 169 | 510  | T | * |        |      | CYT |       | 10  |
| Sg-119  | 4  | 127 | 384  | T | * |        |      | CYT |       | 9   |
| Sg-127  | 3  | 320 | 963  | V | * |        |      | CYT |       | 11  |
| Sg-129  | 3  | 313 | 942  | M | * | 29,583 | 4,73 | SIG | 21-22 | 12  |
| Sg-121  | 2  | 359 | 1080 | M | * | 34,072 | 4,5  | SIG | 21-22 | 12  |
| Sg-125  | 2  | 340 | 1023 | M | * | 32,141 | 4,64 | SIG | 21-22 | 7   |
| Sg-128  | 2  | 330 | 993  | M | * | 31,223 | 4,59 | SIG | 21-22 | 10  |
| Acidic mucins proteins similar to Basic 7-13 <i>Simulium</i> family |    |     |      |   |   |        |      |     |       |     |
| Sg-177  | 7  | 114 | 345  | M | * | 9,054  | 4,48 | SIG | 23-24 |     |
| Sg-170  | 4  | 124 | 375  | M | * | 9,836  | 4,48 | SIG | 23-24 |     |
| Sg-165  | 4  | 134 | 405  | M | * | 10,585 | 4,48 | SIG | 23-24 |     |
| Sg-174  | 4  | 88  | 267  | F | * |        |      | CYT |       |     |
| Sg-171  | 3  | 114 | 345  | M | * | 9,016  | 4,22 | SIG | 23-24 |     |
| Sg-176  | 3  | 110 | 333  | M | * | 8,726  | 4,22 | SIG | 23-24 |     |
| Sg-168  | 2  | 128 | 387  | M | * | 10,056 | 4,22 | SIG | 23-24 |     |
| Sg-169  | 2  | 122 | 369  | M | * | 9,674  | 4,48 | SIG | 23-24 |     |
| Sg-173  | 2  | 118 | 357  | M | * | 9,377  | 4,48 | SIG | 23-24 |     |
| Sg-166  | 2  | 128 | 387  | M | * | 10,141 | 4,48 | SIG | 23-24 |     |
| Sg-167  | 2  | 130 | 393  | M | * | 10,269 | 4,48 | SIG | 23-24 |     |
| Sg-172  | 2  | 122 | 369  | M | * | 9,639  | 4,1  | SIG | 23-24 |     |
| Sg-175  | 2  | 122 | 369  | M | * | 9,593  | 4,22 | SIG | 23-24 |     |

|   |    |     |     |   |   |        |       |     |       |    |
|---|----|-----|-----|---|---|--------|-------|-----|-------|----|
| <b><i>Simulium</i> collagen-like family</b>                               |    |     |     |   |   |        |       |     |       |    |
| Sg-152  | 12 | 218 | 657 | M | * | 21,89  | 9,79  | SIG | 19-20 | 87 |
| Sg-149  | 11 | 182 | 549 | S | * |        |       | CYT |       | 84 |
| <b><i>Sv</i> 7.8 kDa family</b>   |    |     |     |   |   |        |       |     |       |    |
| Sg-372  | 3  | 126 | 381 | M | * | 11,153 | 10,34 | SIG | 22-23 | 9  |
| Sg-356  | 3  | 138 | 417 | M | * | 13,474 | 9,08  | SIG | 22-23 | 11 |
| Sg-205  | 37 | 107 | 324 | M | * | 8,348  | 10,72 | SIG | 28-29 | 3  |
| Sg-227  | 5  | 88  | 267 | M | * | 7,145  | 10,47 | SIG | 23-24 | 4  |
| <b>Basic 7-13 <i>Simulium</i> family</b>                                  |    |     |     |   |   |        |       |     |       |    |
| Sg-420  | 4  | 102 | 309 | M | * | 8,648  | 11,35 | SIG | 23-24 | 3  |
| Sg-403  | 2  | 106 | 321 | V |   |        |       | CYT |       | 8  |
| <b><i>Simulium</i> 4.8 kDa family</b>                                     |    |     |     |   |   |        |       |     |       |    |
| Sg-195  | 23 | 65  | 198 | M | * | 5,006  | 4,43  | SIG | 17-18 |    |
| Sg-199  | 5  | 65  | 198 | M | * | 4,992  | 4,43  | SIG | 17-18 |    |
| Sg-196  | 5  | 65  | 198 | M | * | 5,004  | 4,37  | SIG | 17-18 |    |
| <b><i>Simulium</i> basic 7.4 kDa family</b>                               |    |     |     |   |   |        |       |     |       |    |
| Sg-422  | 4  | 87  | 264 | M | * | 6,897  | 11,75 | SIG | 25-26 | 4  |
| <b><i>Simulium</i> basic 13 kDa family</b>                                |    |     |     |   |   |        |       |     |       |    |
| Sg-446  | 3  | 139 | 420 | M | * | 13,741 | 10,66 | SIG | 22-23 | 2  |
| <b>5-Cys <i>Simulium</i> family</b>                                       |    |     |     |   |   |        |       |     |       |    |
| Sg-282  | 9  | 135 | 408 | L | * | 14,235 | 5,72  | SIG | 16-17 | 3  |
| <b>Deorphanized <i>Simulium nigrimanum</i> 8-10 Cys W family</b>          |    |     |     |   |   |        |       |     |       |    |
| Sg-324  | 8  | 163 | 492 | M | * | 16,386 | 4,63  | SIG | 22-23 | 12 |
| Sg-340  | 3  | 163 | 492 | M | * | 16,944 | 8,06  | SIG | 20-21 | 22 |
| <b>Acid 28 kDa family</b>   |    |     |     |   |   |        |       |     |       |    |
| Sg-319  | 5  | 166 | 501 | S | * |        |       | CYT |       | 23 |
| Sg-321  | 4  | 221 | 666 | M | * | 22,61  | 7,06  | SIG | 18-19 | 41 |
| Sg-320  | 3  | 268 | 807 | M | * | 27,423 | 4,98  | SIG | 18-19 | 45 |
| <b><i>Simulium</i> basic 28 kDa family</b>                                |    |     |     |   |   |        |       |     |       |    |
| Sg-136  | 9  | 263 | 792 | M | * | 28,117 | 8,8   | SIG | 19-20 | 34 |
| <b>Deorphanized 19 kDa family first seen in <i>S. nigrimanum</i> only</b> |    |     |     |   |   |        |       |     |       |    |
| Sg-303  | 6  | 164 | 495 | M | * | 16,887 | 9,25  | SIG | 21-22 | 17 |
| Sg-309  | 5  | 165 | 498 | M | * | 16,815 | 9,09  | SIG | 22-23 | 9  |
| <b>Deorphanized 8kDa basic protein</b>                                    |    |     |     |   |   |        |       |     |       |    |
| Sg-258  | 22 | 96  | 291 | M | * | 8,187  | 10,57 | SIG | 19-20 | 5  |
| <b>Deorphanized Sn basic 4.4kDa family</b>                                |    |     |     |   |   |        |       |     |       |    |
| Sg-438  | 2  | 61  | 186 | M | * | 4,8    | 12,01 | SIG | 19-20 |    |
| <b>Proteins so far unique to <i>S. guianense</i></b>                      |    |     |     |   |   |        |       |     |       |    |
| <b>Novel peptide similar to kunitoxin</b>                                 |    |     |     |   |   |        |       |     |       |    |
| Sg-375  | 6  | 114 | 345 | M | * | 9,569  | 9,91  | SIG | 23-24 | 4  |
| Sg-409  | 2  | 96  | 291 | M | * | 8,057  | 9,87  | SIG | 23-24 |    |
| <b>Housekeeping proteins</b>  |    |     |     |   |   |        |       |     |       |    |
| <b>Ribosomal proteins</b>   |    |     |     |   |   |        |       |     |       |    |
| Sg-377  | 3  | 110 | 333 | M | * |        |       | BL  |       | 3  |
| Sg-390  | 3  | 25  | 78  | M | * |        |       | CYT |       |    |
| Sg-441  | 3  | 234 | 705 | F | * |        |       | CYT |       | 5  |
| Sg-452  | 3  | 117 | 354 | S | * |        |       | CYT |       | 4  |
| Sg-449  | 2  | 140 | 423 | M | * |        |       | CYT |       | 5  |

| Energy metabolism               |   |     |     |   |   |  |     |  |   |
|---------------------------------|---|-----|-----|---|---|--|-----|--|---|
| Sg-432                          | 3 | 267 | 804 | G | * |  | CYT |  |   |
| Sg-428                          | 2 | 260 | 783 | T | * |  | CYT |  |   |
| DNA-binding nuclear protein p8  |   |     |     |   |   |  |     |  |   |
| Sg-401                          | 4 | 66  | 201 | M | * |  | CYT |  |   |
| Conserved hypothetical proteins |   |     |     |   |   |  |     |  |   |
| Sg-412                          | 2 | 185 | 558 | V | * |  | CYT |  | 3 |
| Sg-411                          | 2 | 176 | 531 | V | * |  | CYT |  |   |

| Seq name  | Clust 70% at Sim on 50% of length Cluster | # seq | Clust 75% at Sim on 50% f length Cluster | # seq | Clust 80% at Sim on 50% of length Cluster | # seq | Clust at 85% Sim On 50% of length Cluster | # seqs | Clust At 90% Sim on 50% of length Cluster | # seq | Clust At 95% Sim On 50% of length-Cluster | # seq |
|---|---|-------|--|-------|---|-------|---|--------|---|-------|---|-------|
| Secreted proteins                                     |   |       |  |       |   |       |   |        |   |       |   |       |
| Ubiquitous protein families                           |   |       |  |       |   |       |   |        |   |       |   |       |
| Amylase/Maltase                                       |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-214  | 41  | 1     | 39                                       | 1     | 38  | 1     | 38  | 1      | 37  | 1     | 34  | 1     |
| Sg-296  | 69  | 1     | 67                                       | 1     | 66  | 1     | 66  | 1      | 67  | 1     | 70  | 1     |
| Serine proteases                                      |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-138  | 22  | 2     | 23                                       | 2     | 23  | 2     | 23  | 2      | 23  | 2     | 40  | 1     |
| Sg-244  | 5   | 9     | 5  | 9     | 5   | 9     | 5   | 9      | 5   | 9     | 4   | 9     |
| Sg-416  | 53  | 1     | 51                                       | 1     | 50  | 1     | 50  | 1      | 50  | 1     | 49  | 1     |
| Hyaluronidase   |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-414  | 66  | 1     | 64                                       | 1     | 63  | 1     | 63  | 1      | 64  | 1     | 67  | 1     |
| Apyrase   |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-354  | 62  | 1     | 60                                       | 1     | 59  | 1     | 59  | 1      | 60  | 1     | 60  | 1     |
| Kunitz salivary protein                               |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-395  | 80  | 1     | 83                                       | 1     | 83  | 1     | 83  | 1      | 84  | 1     | 92  | 1     |
| Lysozyme  |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-263  | 9   | 6     | 8  | 6     | 8   | 6     | 8   | 6      | 8   | 6     | 7   | 6     |
| Cecropin  |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-369  | 24  | 2     | 25                                       | 2     | 25  | 2     | 25  | 2      | 25  | 2     | 22  | 2     |
| Sg-368  | 24  | 2     | 25                                       | 2     | 25  | 2     | 25  | 2      | 25  | 2     | 22  | 2     |
| Antigen-5 family                                      |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-453  | 81  | 1     | 84                                       | 1     | 84  | 1     | 84  | 1      | 85  | 1     | 93  | 1     |
| Sg-457  | 84  | 1     | 87                                       | 1     | 87  | 1     | 87  | 1      | 88  | 1     | 101                                       | 1     |
| Insect-specific proteins                              |   |       |  |       |   |       |   |        |   |       |   |       |
| Aegyptin  |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-276  | 14  | 3     | 14                                       | 3     | 14  | 3     | 14  | 3      | 14  | 3     | 21  | 2     |
| Diptera secreted protein from conserved insect family |   |       |  |       |   |       |   |        |   |       |   |       |
| Sg-256  | 58  | 1     | 56                                       | 1     | 55  | 1     | 55  | 1      | 55  | 1     | 55  | 1     |
| Sg-216  | 21  | 2     | 22                                       | 2     | 22  | 2     | 22  | 2      | 22  | 2     | 19  | 2     |

|  |    |    |    |    |    |    |    |    |    |    |     |    |
|--|----|----|----|----|----|----|----|----|----|----|-----|----|
| Sg-292   | 42 | 1  | 40 | 1  | 39 | 1  | 39 | 1  | 38 | 1  | 35  | 1  |
| Sg-215   | 21 | 2  | 22 | 2  | 22 | 2  | 22 | 2  | 22 | 2  | 19  | 2  |
| Sg-347   | 61 | 1  | 59 | 1  | 58 | 1  | 58 | 1  | 59 | 1  | 59  | 1  |
| <b>Laminin-like secreted salivary protein found in <i>Simulium</i> and <i>Culicoides</i></b> |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-431   | 75 | 1  | 74 | 1  | 73 | 1  | 73 | 1  | 74 | 1  | 80  | 1  |
| <b>Long D7 family</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-220   | 16 | 3  | 16 | 3  | 16 | 3  | 16 | 3  | 16 | 3  | 14  | 3  |
| Sg-261   | 17 | 3  | 17 | 3  | 17 | 3  | 17 | 3  | 17 | 3  | 27  | 2  |
| Sg-218   | 91 | 1  | 94 | 1  | 94 | 1  | 94 | 1  | 95 | 1  | 110 | 1  |
| <b>D7 16 kDa family</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-350   | 18 | 3  | 19 | 3  | 19 | 3  | 19 | 3  | 19 | 3  | 16  | 3  |
| Sg-331   | 13 | 5  | 12 | 5  | 12 | 5  | 12 | 5  | 11 | 5  | 11  | 4  |
| <b>Ultra-short D7 proteins (10-12 kDa mature weight)</b>                                     |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-75  | 2  | 16 | 2  | 16 | 2  | 16 | 2  | 16 | 2  | 16 | 2   | 16 |
| Sg-190   | 8  | 6  | 7  | 6  | 7  | 6  | 7  | 6  | 7  | 6  | 6   | 6  |
| Sg-363   | 26 | 2  | 27 | 2  | 27 | 2  | 27 | 2  | 43 | 1  | 41  | 1  |
| Sg-383   | 88 | 1  | 91 | 1  | 91 | 1  | 91 | 1  | 92 | 1  | 105 | 1  |
| <b><i>Simulium</i>-specific proteins</b>   |    |    |    |    |    |    |    |    |    |    |     |    |
| <b>SVEP/Marydilan</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-8   | 1  | 16 | 1  | 16 | 1  | 16 | 1  | 16 | 1  | 16 | 1   | 16 |
| Sg-1   | 1  | 16 | 1  | 16 | 1  | 16 | 1  | 16 | 1  | 16 | 1   | 16 |
| Sg-102   | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 28  | 2  |
| Sg-344   | 25 | 2  | 26 | 2  | 26 | 2  | 26 | 2  | 26 | 2  | 23  | 2  |
| Sg-95  | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 17  | 3  |
| Sg-99  | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 66  | 1  |
| Sg-90  | 33 | 2  | 32 | 2  | 32 | 2  | 32 | 2  | 31 | 2  | 29  | 2  |
| Sg-93  | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 83  | 1  |
| <a href="#">Sg-92</a>  | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 17  | 3  |
| Sg-100   | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 95  | 1  |
| Sg-103   | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 98  | 1  |
| Sg-94  | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 99  | 1  |
| Sg-101   | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 4  | 11 | 100 | 1  |
| <b>H-rich, acidic proteins of <i>Simulium</i></b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-41  | 43 | 1  | 41 | 1  | 40 | 1  | 40 | 1  | 39 | 1  | 36  | 1  |
| Sg-30  | 30 | 2  | 31 | 2  | 31 | 2  | 31 | 2  | 30 | 2  | 26  | 2  |
| Sg-37  | 59 | 1  | 57 | 1  | 56 | 1  | 56 | 1  | 56 | 1  | 56  | 1  |
| Sg-33  | 45 | 1  | 43 | 1  | 42 | 1  | 42 | 1  | 41 | 1  | 38  | 1  |
| Sg-31  | 60 | 1  | 58 | 1  | 57 | 1  | 57 | 1  | 57 | 1  | 57  | 1  |
| Sg-38  | 23 | 2  | 24 | 2  | 24 | 2  | 24 | 2  | 24 | 2  | 20  | 2  |
| Sg-40  | 47 | 1  | 45 | 1  | 44 | 1  | 44 | 1  | 44 | 1  | 42  | 1  |
| Sg-36  | 71 | 1  | 69 | 1  | 68 | 1  | 68 | 1  | 69 | 1  | 74  | 1  |
| Sg-32  | 72 | 1  | 71 | 1  | 70 | 1  | 70 | 1  | 71 | 1  | 77  | 1  |
| Sg-34  | 38 | 2  | 36 | 2  | 36 | 2  | 36 | 2  | 35 | 2  | 33  | 2  |
| Sg-42  | 54 | 1  | 52 | 1  | 51 | 1  | 51 | 1  | 51 | 1  | 50  | 1  |
| <b><i>Simulium</i> mucin</b>   |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-126   | 44 | 1  | 42 | 1  | 41 | 1  | 41 | 1  | 40 | 1  | 37  | 1  |
| Sg-120   | 65 | 1  | 63 | 1  | 62 | 1  | 62 | 1  | 63 | 1  | 65  | 1  |
| Sg-117   | 34 | 2  | 76 | 1  | 76 | 1  | 76 | 1  | 77 | 1  | 84  | 1  |
| Sg-119   | 34 | 2  | 77 | 1  | 77 | 1  | 77 | 1  | 78 | 1  | 85  | 1  |
| Sg-127   | 83 | 1  | 86 | 1  | 86 | 1  | 86 | 1  | 87 | 1  | 96  | 1  |
| Sg-129   | 50 | 1  | 48 | 1  | 47 | 1  | 47 | 1  | 47 | 1  | 45  | 1  |

|  |    |    |    |    |    |    |    |    |    |    |     |    |
|--|----|----|----|----|----|----|----|----|----|----|-----|----|
| Sg-121   | 55 | 1  | 53 | 1  | 52 | 1  | 52 | 1  | 52 | 1  | 51  | 1  |
| Sg-125   | 56 | 1  | 54 | 1  | 53 | 1  | 53 | 1  | 53 | 1  | 52  | 1  |
| Sg-128   | 57 | 1  | 55 | 1  | 54 | 1  | 54 | 1  | 54 | 1  | 53  | 1  |
| <b>Acidic mucins proteins similar to Basic 7-13 <i>Simulium</i> family</b> |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-177   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-170   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-165   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 76  | 1  |
| Sg-174   | 3  | 13 | 3  | 13 | 74 | 1  | 74 | 1  | 75 | 1  | 81  | 1  |
| Sg-171   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-176   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 87  | 1  |
| Sg-168   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-169   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-173   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-166   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-167   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-172   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| Sg-175   | 3  | 13 | 3  | 13 | 3  | 12 | 3  | 12 | 3  | 12 | 3   | 10 |
| <b><i>Simulium</i> collagen-like family</b>                                |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-152   | 11 | 5  | 10 | 5  | 10 | 5  | 10 | 5  | 12 | 4  | 12  | 3  |
| Sg-149   | 11 | 5  | 10 | 5  | 10 | 5  | 10 | 5  | 58 | 1  | 58  | 1  |
| <b>Sv 7.8 kDa family</b>   |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-372   | 36 | 2  | 34 | 2  | 34 | 2  | 34 | 2  | 33 | 2  | 31  | 2  |
| Sg-356   | 35 | 2  | 33 | 2  | 33 | 2  | 33 | 2  | 32 | 2  | 30  | 2  |
| Sg-205   | 28 | 2  | 29 | 2  | 29 | 2  | 29 | 2  | 28 | 2  | 54  | 1  |
| Sg-227   | 6  | 9  | 6  | 9  | 6  | 9  | 6  | 9  | 6  | 9  | 5   | 9  |
| <b>Basic 7-13 <i>Simulium</i> family</b>                                   |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-420   | 74 | 1  | 73 | 1  | 72 | 1  | 72 | 1  | 73 | 1  | 79  | 1  |
| Sg-403   | 90 | 1  | 93 | 1  | 93 | 1  | 93 | 1  | 94 | 1  | 109 | 1  |
| <b><i>Simulium</i> 4.8 kDa family</b>                                      |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-195   | 10 | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 8   | 5  |
| Sg-199   | 10 | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 8   | 5  |
| Sg-196   | 10 | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 9  | 5  | 8   | 5  |
| <b><i>Simulium</i> basic 7.4 kDa family</b>                                |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-422   | 73 | 1  | 72 | 1  | 71 | 1  | 71 | 1  | 72 | 1  | 78  | 1  |
| <b><i>Simulium</i> basic 13 kDa family</b>                                 |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-446   | 79 | 1  | 82 | 1  | 82 | 1  | 82 | 1  | 83 | 1  | 91  | 1  |
| <b>5-Cys <i>Simulium</i> family</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-282   | 15 | 3  | 15 | 3  | 15 | 3  | 15 | 3  | 15 | 3  | 13  | 3  |
| <b>Deorphanized <i>Simulium nigrimanum</i> 8-10 Cys W family</b>           |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-324   | 63 | 1  | 61 | 1  | 60 | 1  | 60 | 1  | 61 | 1  | 62  | 1  |
| Sg-340   | 27 | 2  | 28 | 2  | 28 | 2  | 28 | 2  | 27 | 2  | 24  | 2  |
| <b>Acid 28 kDa family</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-319   | 68 | 1  | 66 | 1  | 65 | 1  | 65 | 1  | 66 | 1  | 69  | 1  |
| Sg-321   | 31 | 2  | 70 | 1  | 69 | 1  | 69 | 1  | 70 | 1  | 75  | 1  |
| Sg-320   | 31 | 2  | 78 | 1  | 78 | 1  | 78 | 1  | 79 | 1  | 86  | 1  |
| <b><i>Simulium</i> basic 28 kDa family</b>                                 |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-136   | 12 | 5  | 11 | 5  | 11 | 5  | 11 | 5  | 10 | 5  | 9   | 5  |
| <b>Deorphanized 19 kDa family first seen in <i>S. nigrimanum</i> only</b>  |    |    |    |    |    |    |    |    |    |    |     |    |
| Sg-303   | 7  | 7  | 13 | 4  | 13 | 4  | 13 | 4  | 13 | 4  | 10  | 4  |

|  |    |   |    |   |    |   |    |   |    |   |     |   |
|--|----|---|----|---|----|---|----|---|----|---|-----|---|
| Sg-309   | 7  | 7 | 18 | 3 | 18 | 3 | 18 | 3 | 18 | 3 | 15  | 3 |
| <b>Deorphanized 8kDa basic protein</b>               |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-258   | 29 | 2 | 30 | 2 | 30 | 2 | 30 | 2 | 29 | 2 | 25  | 2 |
| <b>Deorphanized Sn basic 4.4kDa family</b>           |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-438   | 86 | 1 | 89 | 1 | 89 | 1 | 89 | 1 | 90 | 1 | 103 | 1 |
| <b>Proteins so far unique to <i>S. guianense</i></b> |    |   |    |   |    |   |    |   |    |   |     |   |
| <b>Novel peptide similar to kunitoxin</b>            |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-375   | 64 | 1 | 62 | 1 | 61 | 1 | 61 | 1 | 62 | 1 | 64  | 1 |
| Sg-409   | 85 | 1 | 88 | 1 | 88 | 1 | 88 | 1 | 89 | 1 | 102 | 1 |
| <b>Housekeeping proteins</b>                         |    |   |    |   |    |   |    |   |    |   |     |   |
| <b>Ribosomal proteins</b>                            |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-377   | 37 | 2 | 35 | 2 | 35 | 2 | 35 | 2 | 34 | 2 | 32  | 2 |
| Sg-390   | 49 | 1 | 47 | 1 | 46 | 1 | 46 | 1 | 46 | 1 | 44  | 1 |
| Sg-441   | 76 | 1 | 79 | 1 | 79 | 1 | 79 | 1 | 80 | 1 | 88  | 1 |
| Sg-452   | 77 | 1 | 80 | 1 | 80 | 1 | 80 | 1 | 81 | 1 | 89  | 1 |
| Sg-449   | 51 | 1 | 49 | 1 | 48 | 1 | 48 | 1 | 48 | 1 | 46  | 1 |
| <b>Energy metabolism</b>                             |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-432   | 78 | 1 | 81 | 1 | 81 | 1 | 81 | 1 | 82 | 1 | 90  | 1 |
| Sg-428   | 89 | 1 | 92 | 1 | 92 | 1 | 92 | 1 | 93 | 1 | 106 | 1 |
| <b>DNA-binding nuclear protein p8</b>                |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-401   | 48 | 1 | 46 | 1 | 45 | 1 | 45 | 1 | 45 | 1 | 43  | 1 |
| <b>Conserved hypothetical proteins</b>               |    |   |    |   |    |   |    |   |    |   |     |   |
| Sg-412   | 39 | 2 | 37 | 2 | 37 | 2 | 37 | 2 | 36 | 2 | 107 | 1 |
| Sg-411   | 39 | 2 | 37 | 2 | 37 | 2 | 37 | 2 | 36 | 2 | 108 | 1 |



## 10 REFERÊNCIAS BIBLIOGRÁFICAS

- Abebe M, Cupp MS, Champagne D, Cupp EW. Simulidin: a black fly (*Simulium vittatum*) salivary gland protein with anti-thrombin activity. *J Insect Physiol* 1995; 41:1001-6.
- Abebe M, Cupp MS, Ramberg FB, Cupp EW. Anticoagulant activity in salivary gland extracts of black flies (Diptera: Simuliidae). *J Med Entomol* 1994; 31(6):908-11.
- Abrams PJ, Emerson CR. Rivaroxaban: a novel, oral, direct factor Xa inhibitor. *Pharmacotherapy* 2009; 29(2):167-81.
- Adler P. *Biology of Disease vectors*. 2<sup>nd</sup> ed. Oxford (UK): Elsevier Academy Press; 2005.
- Adler PH, Crosskey RW. *World Blackflies (Diptera: Simuliidae): A fully revised edition of taxonomic and geographical inventory*. [s.l.]: [s.n]. 2008; 105 p. Disponível em : <http://entweb.clemson.edu/biomia/pdfs/blackflyinventory.pdf>. Acesso em: 20 de maio 2011.
- Al-Ahdal MN, Al-Hussain K, Thorogood RJ, Reilly HC, Wilson JD. Protein constituents of mosquito saliva: studies on *Culex molestus*. *J Trop Med Hyg* 1990; 93:98-105.
- Alarcon-Chaidez FJ, Sun J, Wikel SK. Transcriptome analysis of the salivary glands of *Dermacentor andersoni stiles* (Acari: Ixodidae). *Insect Biochem Mol Biol* 2007; 37(1):48-71.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997; 25(17):3389-402.
- Alvarenga PH, Francischetti IM, Calvo E, Sa-Nunes A, Ribeiro JM, Andersen JF. The function and three-dimensional structure of a thromboxane a(2)/cysteinyl leukotriene-binding protein from the saliva of a mosquito vector of the malaria parasite. *PLoS Biology* 2010; 8(11):e1000547.

Alves-Silva J, Ribeiro JM, Van Den Abbeele J, Attardo G, Hao Z, Haines LR, Soares MB, Berriman M, Aksoy S, Lehane MJ. An insight into the sialome of *Glossina morsitans morsitans*. BMC Genomics 2010; 11:213.

Ameri M, Wang X, Wilkerson MJ, Kanost MR, Broce AB. An immunoglobulin binding protein (antigen 5) of the stable fly (Diptera: Muscidae) salivary gland stimulates bovine immune responses. J Med Entomol 2008; 45(1):94-101.

Andersen JF, Hinnebusch BJ, Lucas DA, Conrads TP, Veenstra TD, Pham VM, Ribeiro JM. An insight into the sialome of the oriental rat flea, *Xenopsylla cheopis* (Rots). BMC Genomics 2007; 8:102.

Andersen JF, Pham VM, Meng Z, Champagne DE, Ribeiro JM. Insight into the sialome of the Black Fly, *Simulium vittatum*. J Proteome Res 2009; 8(3):1474-88.

Arca B, Lombardo F, Francischetti IM, Pham VM, Mestres-Simon M, Andersen JF, Ribeiro JM. An insight into the sialome of the adult female mosquito *Aedes albopictus*. Insect Biochem Mol Biol 2007; 37(2):107-27.

Arca B, Lombardo F, Valenzuela JG, Francischetti IM, Marinotti O, Coluzzi M, Ribeiro JM. An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. J Exp Biol 2005; 208(Pt 20):3971-86.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25(1):25-9.

Assumpcao TC, Charneau S, Santiago PB, Francischetti IM, Meng Z, Araujo CN, Pham VM, Queiroz RM, Castro CN, Ricart CA, Santana J, Ribeiro JM. Insight into the Salivary Transcriptome and Proteome of *Dipetalogaster maxima*. J Proteome Res 2010; Nov 8.

Assumpcao TC, Francischetti IM, Andersen JF, Schwarz A, Santana JM, Ribeiro JM. An insight into the sialome of the blood-sucking bug *Triatoma infestans*, a vector of Chagas' disease. Insect Biochem Mol Biol 2008; 38(2):213-32.

Bahia D, Gontijo NF, Leon IR, Perales J, Pereira MH, Oliveira G, Correa-Oliveira R, Reis AB. Antibodies from dogs with canine visceral leishmaniasis recognise two proteins from the saliva of *Lutzomyia longipalpis*. Parasitology Res 2007; 100(3):449-54.

Bajaj MS, Birktoft JJ, Steer SA, Bajaj SP. Structure and biology of tissue factor pathway inhibitor. Thromb Haemost 2001; 86(4):959-72.

Balashov I. Ultrastructural characteristics of the salivary glands of the taiga tick, *Ixodes persulcatus* (Ixodidae). I. The granulosecreting alveoles of the fasting female. Parazitologiya 1979; 13:572-81.

Baptist BA. The morphology and physiology of the salivary glands of Hemiptera-Heteroptera. Q J Microsc Sci 1941; 82:91-139.

Barrow PM, McIver SB, Wright KA. Salivary glands of female *Culex pipiens*: morphological changes associated with maturation and blood-feeding. Can Entomol 1975; 1153-60.

Barth R. Estudo Anatômico e histológico sobre a subfamília Triatominae (Heteroptera, Reduviidae). Mem Inst Oswaldo Cruz 1954; (52):517-85.

Basañez MG, Takaoka H, Suzuki H, Noda S, Tada I. The vectorial of several blackfly species (Diptera: Simuliidae) in relation to human onchocerciasis in the Sierra Parima and upper Orinoco region of Venezuela. *Ann Trop Med Parasitol* 1988; 82:597-611.

Bateman A, Birney E, Durbin R, Eddy SR, Howe KL, Sonnhammer EL. The Pfam protein families database. *Nucleic Acids Res* 2000; 28(1):263-6.

Bearzoti P, Lane E, Menezes J. Relato de um caso de oncocercose adquirida no Brasil. *Rev Paul Med* 1967; (70):102.

Beckett EB. Peripheral cells in the salivary glands of female *Aedes aegypti* and *Aedes togoi* mosquitoes. *J Insect Physiol* 1977; (107):805-7.

Beckett EB. Development and ageing of the salivary glands of adult male *Aedes aegypti* (L.) and *Aedes togoi* (Theobald) mosquitoes (Diptera:Culicidae). *Int J Insect Morphol Embryol* 1988; (17):327-33.

Bennett GF. The salivary gland as an aid in the identification of some simuliids. *Can J Zool* 1963; 41:947-52.

Brennan JDG, Kent M, Dhar R, Fujioka H, Kumar N. *Anopheles gambiae* salivary gland proteins as putative targets for blocking transmission of malaria parasites. *Proc Natl Acad Sci USA* 2000; 97: 13859-64.

Calvo E, Andersen J, Francischetti IM, de LCM, deBianchi AG, James AA, Ribeiro JM, Marinotti O. The transcriptome of adult female *Anopheles darlingi* salivary glands. *Insect Mol Biol* 2004; 13(1):73-88.

Calvo E, Dao A, Pham VM, Ribeiro JM. An insight into the sialome of *Anopheles funestus* reveals an emerging pattern in anopheline salivary protein families. *Insect Biochem Mol Biol* 2007; 37(2):164-75.

Calvo E, Mans BJ, Andersen JF, Ribeiro JM. Function and evolution of a mosquito salivary protein family. *J Biol Chem* 2006; 281(4):1935-42.

Calvo E, Pham VM, Lombardo F, Arca B, Ribeiro JM. The sialotranscriptome of adult male *Anopheles gambiae* mosquitoes. *Insect Biochem Mol Biol* 2006; 36(7):570-5.

Calvo E, Pham VM, Marinotti O, Andersen JF, Ribeiro JM. The salivary gland transcriptome of the neotropical malaria vector *Anopheles darlingi* reveals accelerated evolution of genes relevant to hematophagy. *BMC Genomics* 2009; 10:57.

Calvo E, Pham VM, Ribeiro JM. An insight into the sialotranscriptome of the non-blood feeding *Toxorhynchites amboinensis* mosquito. *Insect Biochem Mol Biol* 2008; 38(5):499-507.

Calvo E, Sanchez-Vargas I, Favreau AJ, Barbian KD, Pham VM, Olson KE, Ribeiro JM. An insight into the sialotranscriptome of the West Nile mosquito vector, *Culex tarsalis*. *BMC Genomics* 2010; 11:51.

Calvo E, Sanchez-Vargas I, Kotsyfakis M, Favreau AJ, Barbian KD, Pham VM, Olson KE, Ribeiro JM. The salivary gland transcriptome of the eastern tree hole mosquito, *Ochlerotatus triseriatus*. *J Med Entomol* 2010; 47(3):376-86.

Campbell CL, Vandyke KA, Letchworth GJ, Drolet BS, Hanekamp T, Wilson WC. Midgut and salivary gland transcriptomes of the arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *Insect Mol Biol* 2005; 14(2):121-36.

Campos IT, Amino R, Sampaio CA, Auerswald EA, Friedrich T, Lemaire HG, Schenkman S, Tanaka AS. Infestin, a thrombin inhibitor presents in *Triatoma infestans* midgut, a Chagas' disease vector: gene cloning, expression and characterization of the inhibitor. *Insect Biochem Mol Biol* 2002; 32(9):991-7.

Cappello M, Bergum PW, Vlasuk GP, Furmidge BA, Pritchard DI, Aksoy S. Isolation and characterization of the tsetse thrombin inhibitor: a potent antithrombotic peptide from the saliva of *Glossina morsitans morsitans*. *Am J Trop Med Hyg* 1996; 54:475-80.

Cappello M, Li S, Chen X, Li C, Harrison L, Narashimhan S, Beard CB, Aksoy S. Tsetse thrombin inhibitor: Bloodmeal-induced expression of an anticoagulant in salivary glands and gut tissue of *Glossina morsitans morsitans*. *Proc Natl Acad Sci USA* 1998; 95:14290-5.

Cerna P, Mikes L, Volf P. Salivary gland hyaluronidase in various species of phlebotomine sand flies (Diptera: Psychodidae). *Insect Biochem Mol Biol* 2002; 32(12):1691-7.

Chagas AC, Medeiros JF, Astolfi-Filho S, Py-Daniel V. Anticoagulant activity in salivary gland homogenates of *Thyrsopelma guianense* (Diptera: Simuliidae), the primary vector of onchocerciasis in the Brazilian Amazon. *Mem Inst Oswaldo Cruz* 2010; 105(2):174-8.

Champagne D, Nussenzveig RH, Ribeiro JMC. Purification, characterization, and cloning of nitric oxide-carrying heme proteins (nitrophorins) from salivary glands of the blood sucking insect *Rhodnius prolixus*. *J Biol Chem* 1995; 270:8691-5.

Champagne DE. Antihemostatic strategies of blood-feeding arthropods. *Curr Drug Targets* 2004; 4(4):375-96.

Champagne DE. Antihemostatic molecules from saliva of blood-feeding arthropods. *Pathophysiol Haemost Thromb* 2005; 34(4-5):221-7.

Champagne DE, Smartt CT, Ribeiro JM, James AA. The salivary gland-specific apyrase of the mosquito *Aedes aegypti* is a member of the 5'-nucleotidase family. *Proc Natl Acad Sci USA* 1995; 92(3):694-8.

Chapman RF. *The Insects: Structure and Function*. 4<sup>th</sup> ed. Cambridge (NY): Cambridge University Press; 1998.

Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly *Lutzomyia longipalpis*. *Proc Natl Acad Sci USA* 1999; 96(26):15155-60.

Ciprandi A, Oliveira SK, Masuda A, Horn F, Termignoni C. *Boophilus microplus*: Its saliva contains microphilin, a small thrombin inhibitor. *Exp Parasitol* 2006; (114):40-6.

Colman RW, Hirsh J, Marder VJ, Clowes AW. Overview of coagulation, fibrinolysis, and their regulation. In: Colman RW, Hirsh J, Marder VJ, Clowes AW, George JN editors. *Hemostasis and thrombosis. Basic principles and clinical practice*, 4nd ed. Lippincott; Williams & Wilkins, Philadelphia, 2001; P. 17-20.

Crosskey RW. *The natural history of black flies*. Chichester: John Wiley & Sons; 1990.

Cupp EW, Cupp MS. Black fly (Diptera:Simuliidae) salivary secretions: importance in vector competence and disease. *J Med Entomol* 1997; 34(2):87-94.

Cupp EW, Cupp MS, Ribeiro JM, Kunz SE. Blood-feeding strategy of *Haematobia irritans* (Diptera: Muscidae). *J Med Entomol* 1998; 35(4):591-5.

Cupp MS, Cupp EW, Ochoa AJ, Moulton JK. Salivary apyrase in New World blackflies (Diptera: Simuliidae) and its relationship to onchocerciasis vector status. *Med Vet Entomol* 1995; 9(3):325-30.

Cupp MS, Cupp EW, Ramberg FB. Salivary gland apyrase in black flies (*Simulium vittatum*). *J Insect Physiol* 1993; 39:817-21.

Cupp MS, Ribeiro JM, Champagne DE, Cupp EW. Analyses of cDNA and recombinant protein for a potent vasoactive protein in saliva of a blood-feeding black fly, *Simulium vittatum*. *J Exp Biol* 1998; 201(Pt 10):1553-61.

Cupp MS, Ribeiro JM, Cupp EW. Vasodilative activity in black fly salivary glands. *Am J Trop Med Hyg* 1994; 50(2):241-6.

Daffre S, Kylsten P, Samakovlis C, Hultmark D. The lysozyme locus in *Drosophila melanogaster*: an expanded gene family adapted for expression in the digestive tract. *Mol Gen Genet* 1994; (242):152-62.

Davie EW, Ratnoff OD. Waterfall sequence for intrinsic blood clotting. *Science* 1964; 145: 1310-1312.

Dan A, Pereira MH, Pesquero JL, Diotaiuti L, Beirao PS. Action of the saliva of *Triatoma infestans* (Heteroptera: Reduviidae) on sodium channels. *J Med Entomol* 1999; 36(6):875-9.

Diaz LA, Sampaio SA, Rivitti EA, Martins CR, Cunha PR, Lombardi C, Almeida FA, Castro RM, Macca ML, Lavrado C, et al. Endemic pemphigus foliaceus (Fogo Selvagem): II. Current and historic epidemiologic studies. *J Invest Dermatol* 1989; 92(1):4-12.



Dickinson RG, O'Hagan JE, Sholtz M, Binnington KC, Hegarty MP. Prostaglandin in the saliva of the cattle tick *Boophilus microplus*. Aust J Exp Biol Med Sci 1976; 54:475-86.

Eriksson BI, Smith H, Yasothan U, Kirkpatrick P. Dabigatran etexilate. Nat Rev Drug Discov 2008; 7(7):557-8.

Faudry E, Lozzi SP, Santana JM, D'Souza-Ault M, Kieffer S, Felix CR, Ricart CA, Sousa MV, Vernet T, Teixeira AR. *Triatoma infestans* apyrases belong to the 5'-nucleotidase family. J Biol Chem 2004; 279(19):19607-13.

Faudry E, Santana JM, Ebel C, Vernet T, Teixeira AR. Salivary apyrases of *Triatoma infestans* are assembled into homo-oligomers. Biochem J 2006; 396(3):509-15.

Fenton JW. Thrombin. Ann N Y Acad Sci 1986; 485: 5-15.

Fenton JW, Landis BH, Walz DA, Finlayson JS. Human Thrombins. In: Chemistry and Biology of Thrombin. Lundblad RL, Fenton JW, Mann KG editors. Ann Arbor Science, Ann Arbor, MI; 1977; P. 43-70.

Forattini OP. The Cimicidae and their importance in public health (Hemiptera-Heteroptera:Cimicidae). Rev Saude Publica 1990; 24:1-37.

Francischetti IM. Platelet aggregation inhibitors from hematophagous animals. Toxicon 2010; 56(7):1130-44.

Francischetti IM, Mather TN, Ribeiro JM. Cloning of a salivary gland metalloprotease and characterization of gelatinase and fibrin(ogen)lytic activities in the saliva of the Lyme disease tick vector *Ixodes scapularis*. Biochem Biophys Res Commun 2003; 305(4):869-75.

Francischetti IM, Meng Z, Mans BJ, Gudderra N, Hall M, Veenstra TD, Pham VM, Kotsyfakis M, Ribeiro JM. An insight into the salivary transcriptome and proteome of the soft tick and vector of epizootic bovine abortion, *Ornithodoros coriaceus*. J Proteomics 2008; 71(5):493-512.

Francischetti IM, My Pham V, Mans BJ, Andersen JF, Mather TN, Lane RS, Ribeiro JM. The transcriptome of the salivary glands of the female western black-legged tick *Ixodes pacificus* (Acari: Ixodidae). Insect Biochem Mol Biol 2005; 35(10):1142-61.

Francischetti IM, Valenzuela JG, Pham VM, Garfield MK, Ribeiro JM. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. J Exp Biol 2002; 205(Pt 16):2429-51.

Francischetti IM, Valenzuela JG, Ribeiro JM. Anophelin: kinetics and mechanism of thrombin inhibition. Biochemistry 1999; 38(50):16678-85.

Franco RF. Fisiologia da Coagulação, Anticoagulação e Fibrinólise. Medicina 2001; 34: 229 – 237.

Friedland N, Liou HL, Lobel P, Stock AM. Structure of a cholesterol-binding protein deficient in Niemann-Pick type C2 disease. Proc Natl Acad Sci USA 2003; 100(5):2512-7.

Friedrich R, Kroger B, Biajolan S, Lemaire HG, Hoffken HW, Reuschenbach P, Otte M, Dodt J. A Kazal-type inhibitor with thrombin specificity from *Rhodnius prolixus*. J Biol Chem 1993; 268:16216-22.

Fuentes-Prior P, Noeske-Jungblut C, Donner P, Schleuning WD, Huber R, Bode W. Structure of the thrombin complex with triabin, a lipocalin-like exosite-binding inhibitor derived from a triatomine bug. Proc Natl Acad Sci USA 1997; 94(22):11845-50.

Gaspar AR, Joubert AM, Crause JC, Neitz AW. Isolation and characterization of an anticoagulant from the salivary glands of the tick, *Ornithodoros savignyi* (Acari: Argasidae). Exp Applied Acarol 1996; 20(10):583-98.

Gibbs GM, Roelants K, O'Bryan MK. The CAP superfamily: cysteine-rich secretory proteins, antigen 5, and pathogenesis-related 1 proteins--roles in reproduction, cancer, and immune defense. Endocr Rev 2008; 29(7):865-97.

Gil-Azevedo LH, Santos-Mallet JR, Maia-Herzog M. Caracteres diagnósticos de *Simulium (Chirostilbia) pertinax* Kollar (Diptera: Simuliidae). Neotrop Entomol 2004; 433-7.

Gillespie RD, Dolan MC, Piesman J, Titus RG. Identification of an IL-2 binding protein in the saliva of the Lyme disease vector tick, *Ixodes scapularis*. J Immunol 2001; 166(7):4319-26.

Gosbee J, Allen JR, West AS. The salivary glands of adult blackflies. Can J Zool 1969; (47):1341-4.

Grimaldi D, Engel M. Evolution of the insects. Cambridge (NY): Cambridge University Press; 2005.

Hamada N, Fouque F. Black flies (Diptera: Simuliidae) of French Guiana: cytotaxonomy and a preliminary list of species. Mem Inst Oswaldo Cruz 2001; 96(7):955-9.

Hang HC, Bertozzi CR. The chemistry and biology of mucin-type O-linked glycosylation. Bioorg Med Chem 2005; 13(17):5021-34.

Hekmat-Scafe DS, Dorit RL, Carlson JR. Molecular evolution of odorant-binding protein genes OS-E and OS-F in *Drosophila*. Genetics 2000; 155(1):117-27.

- Higgs GA, Vane JR, Hart RJ, Porter C, Wilson RG. Prostaglandins in the saliva of the cattle tick, *Boophilus microplus*(Canestrini) (Acarina, Ixodidae). Bull Ent Res 1976; 66:665-70.
- Homan EJ, Zuluaga FN, Yuill TM, Lorbacher H. Studies on the transmission of Venezuelan equine encephalitis virus by Colombian simuliidae (Diptera). Am J Trop Med Hyg 1985; 34(4):799-804.
- Horn F, dos Santos PC, Termignoni C. Boophilus microplus anticoagulant protein: An antithrombin inhibitor isolated from the cattle tick saliva. Arch Biochem Biophys 2000; (384):68-73.
- House CR, Ginsborg BL. Salivary gland. Oxford: Pergamon Press; 1985.
- Huang X, Madan A. CAP3: A DNA sequence assembly program. Genome Res 1999; 9(9):868-77.
- Hulo N, Bairoch A, Bulliard V, Cerutti L, De Castro E, Langendijk-Genevaux PS, Pagni M, Sigrist CJ. The PROSITE database. Nucleic Acids Res 2006; 34:D227-30.
- Jacobs JW, Cupp EW, Sardana M, Friedman PA. Isolation and characterization of a coagulation factor Xa inhibitor from black fly salivary glands. Thromb Haemost 1990; 64(2):235-8.
- James AA, Rossignol PA. Mosquito salivary glands: Parasitological and molecular aspects. Parasitol Today 1991; 7(10):267-71.
- Janzen HG, Wright KA. The salivary glands of *Aedes aegypti* (L): an electron microscope study. Can J Zool 1971; (49):1343-5.

- Jariyapan N, Takaoka H, Choochote W, Poovorawan Y, Otsuka Y, Jitpakdi A, Junkum A. Morphology and electrophoretic protein profiles of female salivary glands in four Oriental black fly species (Diptera: Simuliidae). *J Vector Ecol* 2006; 31(2):406-11.
- Jenny NS, Mann KG. Coagulation cascade: an overview. In: Loscalzo J, Schafer AI editors. *Thrombosis and hemorrhage*, 2nd ed. Williams & Wilkins, Baltimore; 1998; P. 3-27.
- Juhn J, Naeem-Ullah U, Maciel Guedes BA, Majid A, Coleman J, Paolucci Pimenta PF, Akram W, James AA, Marinotti O. Spatial mapping of gene expression in the salivary glands of the dengue vector mosquito, *Aedes aegypti*. *Parasit Vectors* 2011; 4:1.
- Julenius K, Molgaard A, Gupta R, Brunak S. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 2005; 15(2):153-64.
- Kato H, Jochim RC, Lawyer PG, Valenzuela JG. Identification and characterization of a salivary adenosine deaminase from the sand fly *Phlebotomus duboscqi*, the vector of *Leishmania major* in sub-Saharan Africa. *J Exp Biol* 2007; 210(Pt 5):733-40.
- Koh CY, Kini RM. Molecular diversity of anticoagulants from haematophagous animals. *Thromb Haemost* 2009; 102(3):437-53.
- Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 2004; 5(2):150-63.
- Kylsten P, Kinbrell DA, Daffre S, Samakovlis C, Hultmark D. The lysozyme locus in *Drosophila melanogaster*: different genes are expressed in midgut and salivary glands. *Mol Gen Genet* 1992; (232):335-43.

Lacey LA. Simulídeos antropofílicos no Parque Nacional da Amazônia (Tapajós), Brasil, com referências aos efeitos no Homem. Bol Oficina Sanit Panam 1981; (90):326-38.

Lacey LA, Charlwood JD. On the biting activities of some anthropophilic Simuliidae (Diptera). Bull Entomol Res 1980; (90):495-509.

Lacombe D. Anatomia e histologia das glândulas salivares nos triatomíneos. Mem Inst Oswaldo Cruz 1999; (94):557-64.

Lane RP, Crosskey RW. Medical insects and arachnids. Mosquitoes (Culicidae) and blackflies (Simuliidae). London: London Press; 1993.

Law J, Ribeiro JMC, Wells M. Biochemical insights derived from diversity in insects. Ann Rev Biochem 1992; 61:87-112.

Leboulle G, Crippa M, Decrem Y, Mejri N, Brossard M, Bollen A, Godfroid E. Characterization of a novel salivary immunosuppressive protein from *Ixodes ricinus* ticks. J Biol Chem 2002; 15:15.

Lerner EA, Ribeiro JMC, Nelson RJ, Lerner MR. Isolation of maxadilan, a potent vasodilatory peptide from the salivary glands of the sand fly *Lutzomyia longipalpis*. J Biol Chem 1991; 266:11234-6.

Lerner EA, Shoemaker CB. Maxadilan: Cloning and functional expression of the gene encoding this potent vasodilator peptide. J Biol Chem 1992; 267:1062-6.

Macedo-Ribeiro S, Almeida C, Calisto BM, Friedrich T, Mentele R, Sturzebecher J, Fuentes-Prior P, Pereira PJ. Isolation, cloning and structural characterisation of boophilin, a

multifunctional Kunitz-type proteinase inhibitor from the cattle tick. PLoS ONE 2008; 3(2):e1624.

MacFarlane RG. An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. Nature 1964; 202: 498- 499.

Maia-Herzog M, Shelley AJ, Bradley JE, Luna-Dias APA, Calvao RHS, Lowry C, Camargo M, Rubio JM, Post RJ, Coelho GE. Discovery of a new focus of human onchocerciasis in central Brazil. Trans R Soc Trop Med Hyg 1999; 93:235-9.

Mans BJ, Calvo E, Ribeiro JM, Andersen JF. The crystal structure of D7r4, a salivary biogenic amine-binding protein from the malaria mosquito *Anopheles gambiae*. J Biol Chem 2007; 282(50):36626-33.

Mans BJ, Francischetti IMB. Sialomic perspectives on the evolution of blood-feeding behavior in arthropods: Future therapeutics by natural design. Springer Press; 2010.

Marchler-Bauer A, Panchenko AR, Shoemaker BA, Thiessen PA, Geer LY, Bryant SH. CDD: a database of conserved domain alignments with links to domain three-dimensional structure. Nucleic Acids Res 2002; 30(1):281-3.

Marinotti O, James AA. An alpha-glucosidase in the salivary glands of the vector mosquito *Aedes aegypti*. Insect Biochem 1990; (6):619-23.

Markwardt F. Inventory of coagulation inhibitors from animals feeding on blood. Thromb Haemost (Stuttgart) 1994; 72:477-80.

Medeiros JF, Py-Daniel V, Barbosa UC, Farias ES. Epidemiological studies of *Mansonella ozzardi* (Nematoda, Onchocercidae) in indigenous communities of Pauini municipality, Amazonas, Brazil. *Acta Amaz* 2007; 37: 241-246.

Medeiros JF, Py-Daniel V, Barbosa UC, Izzo TJ. *Mansonella ozzardi* in Brazil: prevalence of infection in riverine communities in the Purus region, in the state of Amazonas. *Mem Inst Oswaldo Cruz* 2009; 104: 74-80.

Megahed MM. Anatomy and histology of the alimentary tract of the female biting midge *Culicoides nubeculosus* (Diptera:Heleidae=Ceratopogonidae). *Parasitology* 1956; 46:22-47.

Mende K, Petoukhova O, Koulitchkova V, Schaub GA, Lange U, Kaufmann R, Nowak G. Dipetalogastin, a potent thrombin inhibitor from the blood-sucking insect. *Dipetalogaster maximus* cDNA cloning, expression and characterization. *Eur J Biochem*. 1999 Dec;266(2):583-90.

Miles PW. The saliva of Hemiptera. *Adv Anat Pathol* 1972; 9:183-255.

Misumi Y, Ogata S, Ohkubo K, Hirose S, Ikehara Y. Primary structure of human placental 5'-nucleotidase and identification of the glycolipid anchor in the mature form. *Eur J Biochem* 1990; 191(3):563-9.

Mohamed HA, Ingram GA, Molyneux DH, Sawyer BV. Use of fluorescein-labelled lectin binding of salivary gland to distinguish between *Anopheles stephensi* and *Anopheles albimanus* species and strains. *Insect Biochem* 1991; 21:767-73.

Molyneux DH, Okolo CJ, Lines JD. Variation in fluorescein-labelled lectin staining of salivary glands in the *Anopheles gambiae* complex. *Med Vet Entomol* 1990; 4:459-62.



Moraes MAP. Oncocercose entre os índios Yanomami. Cad Saúde Pública 1991; 7:503-14.

Moraes MAP, Chaves GM. Um caso de oncocercose no território de Roraima, Brasil. Rev Inst Med Trop Sao Paulo 1974; 16:110-3.

Moraes MAP, Dias LB. Oncocercose no Estado do Amazonas. Rev Inst Med Trop São Paulo 1972; 14:330-3.

Moreira-Ferro CK, Daffre S, James AA, Marinotti O. A lysozyme in the salivary glands of the malaria vector *Anopheles darlingi*. Insect Mol Biol 1998; 7(3):257-64.

Moreira-Ferro CK, Marinotti O, Bijovsky AT. Morphological and biochemical analyses of the salivary glands of the malaria vector, *Anopheles darlingi*. Tissue & Cell 1999; 31(3):264-73.

Moreira CK, Marrelli MT, Lima SP, Marinotti O. Analysis of salivary gland proteins of the mosquito *Anopheles darlingi* (Diptera: Culicidae). J Med Entomol 2001; (38):763-7.

Mulla MS, Lacey LA. Biting Flies the Lower Colorado River Basin: Economic and public health implications of *Simulium* (Diptera: Simuliidae). National Center for Infections Disease Traveler's Health 1976; 2003-4.

Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 1997; 10(1):1-6.

Nobile M, Noceti F, Prestipino G, Possani LD. Helothermine, a lizard venom toxin, inhibits calcium current in cerebellar granules. Exp Brain Res 1996; 110(1):15-20.

Noeske-Jungblut C, Haendler B, Donner P, Alagon A, Possani L, Schleuning W. Triabin, a highly potent exosite inhibitor of thrombin. J Biol Chem 1995; (270):28629-34.

Novak MG, Ribeiro JM, Hildebrand JG. 5-hydroxytryptamine adult female *Aedes aegypti* and its role in regulation of salivation. J Exp Biol 1995; (198):167-74.

Ogata S, Hayashi Y, Misumi Y, Ikehara Y. Membrane-anchoring domain of rat liver 5'-nucleotidase: identification of the COOH-terminal serine-523 covalently attached with a glycolipid. Biochem 1990; 29(34):7923-7.

Okolo CJ, Jenni L, Molyneux DH, Wallbanks KR. Surface differences of *Glossina* salivary glands and infectivity of *Trypanosoma brucei gambiense* to *Glossina*. Ann Soc belge Med Tropicale 1990; (70):39-47.

Oliveira F, Kamhawi S, Seitz AE, Pham VM, Guigal PM, Fischer L, Ward J, Valenzuela JG. From transcriptome to immunome: identification of DTH inducing proteins from a *Phlebotomus ariasi* salivary gland cDNA library. Vaccine 2006; 24(3):374-90.

Orr CW, Hudson A, West AS. The salivary glands of *Aedes aegypti*. Can J Zool 1961; (39):265-72.

Paesen GC, Adams PL, Harlos K, Nuttall PA, Stuart DI. Tick histamine-binding proteins: isolation, cloning, and three- dimensional structure. Molecular Cell 1999; 3(5):661-71.

Pereira MH, Souza MEVAP, Martins MS, Penido CM, Diotaiuti L. Anticoagulant activity of *Triatoma infestans* and *Panstrongylus megistus* saliva (Hemiptera/Triatominae). Acta Trop 1996; 61:255-61.

Perez de Leon AA, Lloyd JE, Tabachnick WJ. Sexual dimorphism and development change of the salivary glands in adult *Culicoides variipennis* (Diptera: Ceratopogonidae). J Med Entomol 1994; (31):898-902.

Perez de Leon AA, Valenzuela JG, Tabachnick WJ. Anticoagulant activity in salivary glands of the insect vector *Culicoides variipennis* sonorensis by an inhibitor of factor Xa. Exp Parasitol 1998; 88(2):121-30.

Perrone JB, De Maio J, Spielman A. Regions of mosquito salivary gland distinguished by surface lectin binding characteristics. Insect Biochem Mol Biol 1986;16:313-8.

Pimenta PFP, De Souza W. Fine structure and cytochemistry of endoplasmatic reticulum and its association with the plasma membrane of *Leishmania mexicana amazonensis*. J Submicrosc Cytol Pathol 1985; (17):413-9.

Pimenta PFP, De Souza W. Structural observations on the attachment of promastigotes of the *Leishmania mexicana amazonensis* to the surface of macrophages. Microsc Electron Biol Celular 1986; (10):65-75.

Pimentel GE, Rossignol PA. Age dependence of salivary bacteriolytic activity in adult mosquitoes. Comp Biochem Physiol B 1990; 96:549-51.

Pinheiro FP, Bensabath G, Costa D, Jr., Maroja OM, Lins ZC, Andrade AH. Haemorrhagic syndrome of Altamira. Lancet 1974; 1(7859):639-42.

Poehling HM. Distribution of specific proteins in the salivary gland lobes of culicidae and their relation to age and blood sucking. J Insect Physiol 1979; (25):3-8.

Py-Daniel V. Algumas considerações sobre o Programa de Eliminação da Oncocercose para as Américas. Mem Del CAICET IV 1994; 1,2:157-61.

Py-Daniel V. Oncocercose, uma endemia focal no Hemisfério norte da Amazônia. INPA Press; 1997.

- Py-Daniel V, Andreazze, R., and Medeiros, J.F. Py-Daniel, V., Andreazze, R., and Medeiros, J.F. Projeto piloto Xitei/Xidea (Roraima). Índices epidemiológicos da transmissão de *Onchocerca volvulus* (Leuckart, 1893) para os anos de 1995-1996. Entomol Vectores 2000; 7:389-444.
- Py-Daniel V, Darwich SM, Mardini LB, Strieder MN, Coscarón S. Simuliidae (Diptera, Culicomorpha) no Brasil – XII sobre o gênero *Thyrsopelma* Enderlein 1934: descrição de *T. jeteri* Sp. N., considerações sobre o cibário e chaves para as pupas. Acta Amaz 2005; 35:463-72.
- Py-Daniel V, Moreira-Sampaio RT. Jalacingomyia Gen. (Culicomorpha); A ressurreição de Gymnopauidinae; a eliminação do nível tribal, Apresentação de novos caracteres e a redescritção dos estágios larval e pupal de *Simulium colombaschense* (Fabricius, 1787) (Diptera, Simuliidae). Mem Del CAICET IV 1994; 1,2:101-48.
- Racioppi JV, Spielman A. Secretory proteins from the salivary glands of adult *Aedes aegypti* mosquitoes. Insect Biochem 1987; 17:503.
- Ribeiro JM. Blood-feeding arthropods: live syringes or invertebrate pharmacologists? Infect Agents Dis 1995; 4(3):143-52.
- Ribeiro JM. Blood-feeding in mosquitoes: probing time and salivary gland anti- haemostatic activities in representatives of three genera (*Aedes*, *Anopheles*, *Culex*). Med Vet Entomol 2000; 14(2):142-8.
- Ribeiro JM, Alarcon-Chaidez F, Francischetti IM, Mans BJ, Mather TN, Valenzuela JG, Wikel SK. An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. Insect Biochem Mol Biol 2006; 36(2):111-29.

Ribeiro JM, Arca B, Lombardo F, Calvo E, Phan VM, Chandra PK, Wikel SK. An annotated catalogue of salivary gland transcripts in the adult female mosquito, *Aedes aegypti*. BMC genomics 2007; 8(1):6.

Ribeiro JM, Charlab R, Pham VM, Garfield M, Valenzuela JG. An insight into the salivary transcriptome and proteome of the adult female mosquito *Culex pipiens quinquefasciatus*. Insect Biochem Mol Biol 2004; 34(6):543-63.

Ribeiro JM, Charlab R, Rowton ED, Cupp EW. *Simulium vittatum* (Diptera: Simuliidae) and *Lutzomyia longipalpis* (Diptera: Psychodidae) salivary gland hyaluronidase activity. J Med Entomol 2000; 37(5):743-7.

Ribeiro JM, Endris TM, Endris R. Saliva of the soft tick, *Ornithodoros moubata*, contains anti-platelet and apyrase activities. Comp Biochem Physiol A 1991; 100(1):109-12.

Ribeiro JM, Francischetti IM. Role of arthropod saliva in blood feeding: sialome and post-sialome perspectives. Annu Rev Entomol 2003; 48:73-88.

Ribeiro JM, Katz O, Pannell LK, Waitumbi J, Warburg A. Salivary glands of the sand fly *Phlebotomus papatasi* contain pharmacologically active amounts of adenosine and 5'-AMP. J Exp Biol 1999; 202(Pt 11):1551-9.

Ribeiro JM, Mans BJ, Arca B. An insight into the sialome of blood-feeding Nematocera. Insect Biochem Mol Biol 2010; 40(11):767-84.

Ribeiro JM, Mather TN. *Ixodes scapularis*: salivary kininase activity is a metallo dipeptidyl carboxypeptidase. Exp Parasitol 1998; 89(2):213-21.

Ribeiro JM, Modi G. The salivary adenosine/AMP content of *Phlebotomus argentipes* Annandale and Brunetti, the main vector of human kala-azar. J Parasitol 2001; 87(4):915-7.

Ribeiro JM, Valenzuela JG. The salivary purine nucleosidase of the mosquito, *Aedes aegypti*. Insect Biochem Mol Biol 2003; 33(1):13-22.

Ribeiro JM, Valenzuela JG, Pham VM, Kleeman L, Barbian KD, Favreau AJ, Eaton DP, Aoki V, Hans-Filho G, Rivitti EA, Diaz LA. An insight into the sialotranscriptome of *Simulium nigrimanum*, a black fly associated with fogo selvagem in South America. Am J Trop Med Hyg 2010; 82(6):1060-75.

Ribeiro JMC. The antiserotonin and antihistamine activities of salivary secretion of *Rhodnius prolixus*. J Insect Physiol 1982; 28:69-75.

Ribeiro JMC. Role of arthropod saliva in blood feeding. Ann Rev Entomol 1987; 32:463-78.

Ribeiro JMC, Arca B. From sialomes to the sialoverse: An insight into the salivary potion of blood feeding insects. Adv Insect Physiol 2009; 37:59-118.

Ribeiro JMC, Hazzard JMH, Nussenzveig RH, Champagne D, Walker FA. Reversible binding of nitric oxide by a salivary nitrosylhemeprotein from the blood sucking bug, *Rhodnius prolixus*. Science 1993; 260:539-41.

Ribeiro JMC, Nussenzveig RH. Nitric oxide synthase activity from a hematophagous insect salivary gland. FEBS Let 1993; 330:165-8.

Ribeiro JMC, Rossignol PA, Spielman A. Blood finding strategy of a capillary feeding sandfly, *Lutzomyia longipalpis*. Comp Biochem Physiol 1986; 83A:683-6.

- Ribeiro JMC, Sarkis JJF, Rossignol PA, Spielman A. Salivary apyrase of *Aedes aegypti*: Characterization and secretory fate. *Comp Biochem Physiol* 1984; 79B:81-6.
- Ribeiro JMC, Vachereau A, Modi GB, Tesh RB. A novel vasodilator peptide from the salivary glands of the sand fly *Lutzomyia longipalpis*. *Science* 1989; 243:212-4.
- Rossignol PA, Lueders AM. Bacteriolytic factor in the salivary glands of *Aedes aegypti*. *Comp Biochem Physiol* 1986; 83B:819-22.
- Russell CL, Heesom KJ, Arthur CJ, Helps CR, Mellor PS, Day MJ, Torsteinsdottir S, Bjornsdottir TS, Wilson AD. Identification and isolation of cDNA clones encoding the abundant secreted proteins in the saliva proteome of *Culicoides nubeculosus*. *Insect Mol Biol* 2009; 18(3):383-93.
- Sá MR, Maia-Herzog M. Overseas disease: comparative studies of onchocerciasis in latin America and Africa. *Hist Cienc Saude Manguinhos* 2003; 10:251-8.
- Sais TC, deMoraes RM, Ribolla PE, de Bianchi AG, Marinotti O, Bijovsky AT. Morphological aspects of *Culex quinquesfasciatus* salivary glands. *Arthropod Struct Dev* 2003; 32:219-26.
- Santos A, Ribeiro JM, Lehane MJ, Gontijo NF, Veloso AB, Sant'Anna MR, Nascimento Araujo R, Grisard EC, Pereira MH. The sialotranscriptome of the blood-sucking bug *Triatoma brasiliensis* (Hemiptera, Triatominae). *Insect Biochem Mol Biol* 2007; 37(7):702-12.
- Schultz J, Copley RR, Doerks T, Ponting CP, Bork P. SMART: a web-based tool for the study of genetically mobile domains. *Nucleic Acids Res* 2000; 28(1):231-4.
- Serrao JE, Castrillon MI, Santos-Mallet JR, Zanuncio JC, Goncalves TC. Ultrastructure of the salivary glands in *Cimex hemipterus* (Hemiptera: Cimicidae). *J Med Entomol* 2008; 45(6):991-9.

Service MW. Medical Entomology. London: London Press; 1997.

Shelley AJ. Biosystematics and distribution of simuliid vectors of human onchocerciasis in South America. Mem Inst Oswaldo Cruz 1988; 83(4):399-403.

Shelley AJ, Hernandez LM, Maia-Herzog M, Luna Dias APA, Garritano PR. The blackflies (Diptera:Simuliidae) of Brazil. Sofia-Moscow: Pensoft Publisher; 2010.

Shelley AJ LC, Maia-Herzog M, Luna Dias ADA, Moraes MAP. Biosystematic studies on the Simuliidae (Diptera) of the Amazonia Onchocerciasis Focus. Bulletin of the British Museum (Natural History) Entomology 1997; 66:1-121.

Soliman MA, Abdel-Hamid ME, Mansour MM. Total salivary gland proteins of female *Culex pipiens* and *Aedes caspius* (Diptera: Culicidae) and their fractionation during adult development and after blood sucking. J Egypt Soc Parasit 1999; 29:9008-9014.

Stark KR, James AA. A factor Xa-directed anticoagulant from the salivary glands of the yellow fever mosquito *Aedes aegypti*. Exp Parasitol 1995; 81:321-31.

Stark KR, James AA. Salivary gland anticoagulants in culicine and anopheline mosquitoes (Diptera: Culicidae). J Med Entomol 1996; 33:645-50.

Stark KR, James AA. Isolation and characterization of the gene encoding a novel factor Xa-directed anticoagulant from the yellow fever mosquito, *Aedes aegypti*. J Biol Chem 1998; 273(33):20802-9.

Sun D, McNicol A, James AA, Peng Z. Expression of functional recombinant mosquito salivary apyrase: A potential therapeutic platelet aggregation inhibitor. Platelets 2006; 17(3):178-84.



Sutcliffe JF, McIver SB. Mechanics of blood-feeding in black flies (Diptera: Simuliidae). *J Morphol* 1987; (180):125-44.

Szymanski MM, Lovette IJ. High lineage diversity and host sharing of malarial parasites in a local avian assemblage. *J Parasitol* 2005; 91(4):768-74.

Takaoka H, Suzuki H, Noda S, Tada I, Basañez M, Narbaiza I, Yarzabal I. Studies on blackflies in relation to onchocerciasis transmission in the Southern Venezuela focus. *Mem Del CAICET* 1985; 3:97-110.

Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, Rao BS, Smirnov S, Sverdlov AV, Vasudevan S, Wolf YI, Yin JJ, Natale DA. The COG database: an updated version includes eukaryotes. *BMC bioinformatics* 2003; 4:41.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. The CLUSTAL\_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 1997; 25(24):4876-82.

Valenzuela J, Walker FA, Ribeiro JMC. The salivary vasodilator of the bed bug, *Cimex lectularius*. *J Exp Biol* 1995; 198:1519-26.

Valenzuela JG, Belkaid Y, Garfield MK, Mendez S, Kamhawi S, Rowton ED, Sacks DL, Ribeiro JM. Toward a defined anti-*Leishmania* vaccine targeting vector antigens: characterization of a protective salivary protein. *J Exp Med* 2001; 194(3):331-42.

Valenzuela JG, Belkaid Y, Rowton E, Ribeiro JM. The salivary apyrase of the blood-sucking sand fly *Phlebotomus papatasi* belongs to the novel *Cimex* family of apyrases. J Exp Biol 2001; 204(Pt 2):229-37.

Valenzuela JG, Charlab R, Galperin MY, Ribeiro JM. Purification, cloning, and expression of an apyrase from the bed bug *Cimex lectularius*. A new type of nucleotide-binding enzyme. J Biol Chem 1998; 273(46):30583-90.

Valenzuela JG, Charlab R, Mather TN, Ribeiro JM. Purification, cloning, and expression of a novel salivary anticomplement protein from the tick, *Ixodes scapularis*. J Biol Chem 2000; 275(25):18717-23.

Valenzuela JG, Francischetti IM, Pham VM, Garfield MK, Ribeiro JM. Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. Insect Biochem Mol Biol 2003; 33(7):717-32.

Valenzuela JG, Francischetti IM, Ribeiro JM. Purification, cloning, and synthesis of a novel salivary anti-thrombin from the mosquito *Anopheles albimanus*. Biochemistry 1999; 38(34):11209-15.

Valenzuela JG, Pham VM, Garfield MK, Francischetti IM, Ribeiro JMC. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. Insect Biochem Mol Biol 2002; 32:1101-22.

Valenzuela JGRibeiro JM. Purification and cloning of the salivary nitrophorin from the hemipteran *Cimex lectularius*. J Exp Biol 1998; 201(Pt 18):2659-64.

Valenzuela JG, Walker FA, Ribeiro JM. A salivary nitrophorin (nitric-oxide-carrying hemoprotein) in the bedbug *Cimex lectularius*. J Exp Biol 1995; 198(Pt 7):1519-26.

van de Locht ALD, Bauer MHR, Friedrich TKB, Hoffken W, Bode W. Two heads are better than one: crystal structure of the insect derived double domain Kazal inhibitor fhodniin in complex with thrombin. Embo J 1995; 14:5149-57.

Volf P, Tesarova P, Nohynkova E. Salivary proteins and glycoproteins in phlebotomine sandflies of various species, sex and age. Med Vet Entomol 2000; 14:251-6.

Volfova V, Hostomska J, Cerny M, Votypka J, Volf P. Hyaluronidase of bloodsucking insects and its enhancing effect on leishmania infection in mice. PLoS Neglected Trop Dis 2008; 2(9):e294.

Wächtler VK, Rühm W, Welsch U. Histologische und histochemische Untersuchungen an den Speicheldrüsen der Imagines von *Boopthora erythrocephala* de Geer und *Odagmia ornata ornata* (Meig.) (Dipt., Simuliidae). Z Angew Entomol 1971; (67):189-201.

Wang X, Ribeiro JM, Broce AB, Wilkerson MJ, Kanost MR. An insight into the transcriptome and proteome of the salivary gland of the stable fly, *Stomoxys calcitrans*. Insect Biochem Mol Biol 2009; Jul 2.

Waxman L, Smith DE, Arcuri KE, Vlasuk GP. Tick anticoagulant peptide (TAP) is a novel inhibitor of blood coagulation Factor Xa. Science 1990; 248:593-6.

Wei A, Alexander RS, Duke J, Ross H, Rosenfeld SA, Chang CH. Unexpected binding mode of tick anticoagulant peptide complexed to bovine factor Xa. J Mol Biol 1998; 283(1):147-54.

Workman EFJr, Uttheg LC, Kingdon HS, Lundblad RL. Isolation and characterization of bovine thrombin. In Chemistry and Biology and Biology of Thrombin. Lundblad RL, Fenton JW, Mann KG editors. Ann Arbor Science, Ann Arbor, MI; 1977. P. 23-42.

Wright KA. The anatomy of salivary glands of *Anopheles stephensi* liston. Can J Zool 1969; (47):579-87.

Xu X, Yang H, Ma D, Wu J, Wang Y, Song Y, Wang X, Lu Y, Yang J, Lai R. Toward an understanding of the molecular mechanism for successful blood feeding by coupling proteomics analysis with pharmacological testing of horsefly salivary glands. Mol Cell Proteomics 2008; 7(3):582-90.

Yamazaki Y, Koike H, Sugiyama Y, Motoyoshi K, Wada T, Hishinuma S, Mita M, Morita T. Cloning and characterization of novel snake venom proteins that block smooth muscle contraction. Eur J Biochem 2002; 269(11):2708-15.

Yeh RW, Jang IK. Argatroban: update. Am Heart J 2006; 151(6):1131-8.

Yoshida S, Sudo T, Niimi M, Tao L, Sun B, Kambayashi J, Watanabe H, Luo E, Matsuoka H. Inhibition of collagen-induced platelet aggregation by anopheline antiplatelet protein, a saliva protein from a malaria vector mosquito. Blood 2008; 111(4):2007-14.

Zaitz C, Campbell I, Alves GF. Commentary/millennium: endemic pemphigus foliaceus (fogo selvagem). Int J Dermatol 2000; 39(11):812-4.